

Jörg Franke & Peter Schuderer (Hrsg.)

SIMULATION

IN PRODUKTION UND LOGISTIK

2021



19. Fachtagung Erlangen, 15. - 17. September 2021



Cuvillier Verlag Göttingen
Internationaler wissenschaftlicher Fachverlag





Simulation in Produktion und Logistik 2021





Jörg Franke & Peter Schuderer (Hrsg.)

Simulation in Produktion und Logistik 2021

Erlangen, 15.-17. September 2021



Cuvillier Verlag, Göttingen



Kontakt:

Prof. Dr.-Ing. Jörg Franke
FAU Erlangen-Nürnberg, FAPS
Egerlandstraße 7-9
91058 Erlangen
joerg.franke@faps.fau.de

Prof. Dr. rer. pol. Peter Schuderer
TH Ingolstadt, Business School
Esplanade 10
85049 Ingolstadt
Peter.Schuderer@thi.de

Simulation in Produktion und Logistik 2021
Hrsg. Jörg Franke & Peter Schuderer
Göttingen: Cuvillier Verlag, 2021.
zgl. Tagungsband 19. ASIM-Fachtagung Simulation in Produktion und Logistik,
Erlangen, 15.-17. September 2021
zgl. ASIM-Mitteilung Nr. 177

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

1. Aufl. - Göttingen: Cuvillier, 2021

© CUVILLIER VERLAG, Göttingen 2021

Nonnenstieg 8, 37075 Göttingen

Telefon: 0551-54724-0

Telefax: 0551-54724-21

www.cuvillier.de

Alle Rechte vorbehalten

Dieses Werk ist einschließlich aller seiner Teile urheberrechtlich geschützt. Jede Verwertung, die über die engen Grenzen des Urheberrechtsgesetzes hinausgeht, ist ohne schriftliche Zustimmung des Verlages unzulässig und strafbar. Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen sowie die Speicherung in elektronischen Systemen.

Die Wiedergabe von Warenbezeichnungen und Handelsnamen in diesem Buch berechtigt nicht zu der Annahme, dass solche Bezeichnungen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und deshalb von jedermann benutzt werden dürften.

Soweit in diesem Werk direkt oder indirekt auf Gesetze, Vorschriften oder Richtlinien (z.B. DIN, VDI) Bezug genommen oder aus ihnen zitiert worden ist, kann der Verlag keine Gewähr für Richtigkeit, Vollständigkeit oder Aktualität übernehmen.

Umschlaggestaltung: Graphik & Design | Nürnberg, Ruth Schmidhammer

1. Auflage, 2021

Gedruckt auf umweltfreundlichem, säurefreiem Papier aus nachhaltiger Forstwirtschaft.

ISBN 978-3-7369-7479-1

eISBN 978-3-7369-6479-2



Vorwort

Die Simulation spielt in der Produktentstehung eine stetig wachsende Bedeutung. Aktuelle Technologien wie der Digitale Zwilling und der Digitale Schatten, Augmented und Mixed Reality, Künstliche Intelligenz und synthetische Lernumgebungen, wissensbasierte Konfiguratoren und webbasierte Simulationssysteme sowie die Erweiterung der Verhaltensmodellierung auf alle physikalischen Domänen sind spannende Themen, die auf der ASIM-Fachtagung Simulation in Produktion und Logistik 2021 präsentiert und diskutiert werden.

Die ASIM präsentiert die größte europäische Fachtagung zur Simulation in Produktion und Logistik nur alle zwei Jahre. Wissenschaftliche Forschungsbeiträge und interessante Anwendungsberichte aus der Industrie zeigen aktuelle Entwicklungen und zukunftsweisende Trends zu simulationsgestützten Ansätzen zur Optimierung der Markteinführungszeiten, der operativen Exzellenz und der Ressourceneffizienz. Diskutiert werden technische Möglichkeiten und organisatorische Voraussetzungen zur Nutzung digitaler Modelle in der Planung und im Betrieb von manuellen, automatisierten und hybriden Fertigungs- und Logistikprozessen.

Der vorliegende Tagungsband umfasst die Beiträge der 19. ASIM Fachtagung „Simulation in Produktion und Logistik“ (SPL 2021), die aufgrund der Coronapandemie als digitale Tagung durchgeführt wird.

Kernthemen der Konferenz bilden neue und weiterentwickelte Simulationswerkzeuge und deren fortschrittliche Nutzung zur Vorhersage und zur Rückverfolgbarkeit des Verhaltens sowie zur Verbesserung der Leistungsfähigkeit von Maschinen, Anlagen und komplexen Systemen. Zunehmende Schwerpunkte sind die tragende Rolle der Modellierung und Simulation für die Digitalisierung sowie der Einsatz von datenbasierten Methoden, der Künstlichen Intelligenz und des Maschinellen Lernens.

Wir bedanken uns herzlich bei allen AutorInnen für die Teilung ihres Wissens, den Mitgliedern des Programmkomitees für die Sicherung der Qualität und den Verantwortlichen in der Organisation für den reibungslosen Ablauf der Tagung. Besonderer Dank gilt den Sponsoren und den Ausstellern für die Gestaltung der Ausstellung und die finanzielle Unterstützung der Tagung.

Wir sind überzeugt, dass auch die digitale Form der ASIM-Fachtagung Simulation in Produktion und Logistik 2021 eine effektive Plattform bietet, um neue technologische Ansätze zu diskutieren, aufschlussreiche Einblicke in neue Anwendungen der Simulation in Produktion und Logistik zu erhalten, Fachkontakte zu knüpfen und alte Bekannte in modernen Online-Tools zu treffen. Bei der Lektüre des vorliegenden Tagungsbandes wünschen wir Ihnen viel Vergnügen und Anregungen für Ihre eigenen Tätigkeiten in Forschung und Entwicklung.

Prof. Dr.-Ing. Jörg Franke und Prof. Dr. rer. Pol. Peter Schuderer

Erlangen, September 2021



Foreword

Simulation plays an ever-increasing role in product development. Current technologies such as the Digital Twin and the Digital Shadow, Augmented and Mixed Reality, Artificial Intelligence and synthetic learning environments, knowledge-based configurators and web-based simulation systems as well as the extension of behavioral modeling to all physical domains are exciting topics that will be presented and discussed at the ASIM Dedicated Conference Simulation in Production and Logistics 2021.

ASIM presents the largest European symposium on simulation in production and logistics only every two years. Scientific research contributions and interesting application reports from industry will show current developments and forward-looking trends on simulation-based approaches to optimize time-to-market, operational excellence and resource efficiency. Technical possibilities and organizational requirements for the use of digital models in planning and operation of manual, automated and hybrid manufacturing and logistics processes are discussed.

This conference volume comprises the contributions of the 19th ASIM Symposium "Simulation in Production and Logistics" (SPL 2021), which is being held as a digital conference due to the Corona pandemic.

Core topics of the conference are new and advanced simulation tools and their advanced use for predicting and tracing behavior and improving the performance of machines, plants and complex systems. Increasing emphasis is being placed on the supporting role of modeling and simulation for digitization, as well as the use of data-based methods, artificial intelligence, and machine learning.

We would like to express our sincere gratitude to all authors for sharing their knowledge, to the members of the program committee for ensuring quality, and to those responsible in the organization for the efficient running of the conference. Special thanks go to the sponsors and exhibiting companies for the presentation of the exhibition and the financial support of the conference.

We are convinced that the digital form of the ASIM Symposium Simulation in Production and Logistics 2021 will also provide an effective platform to discuss new technological approaches, gain insightful insights into new applications of simulation in production and logistics, establish expert contacts and meet old acquaintances in modern online tools. While reading the present conference proceedings, we hope you will enjoy the conference and find inspiration for your own activities in research and development.

Prof. Dr.-Ing. Jörg Franke und Prof. Dr. rer. Pol. Peter Schuderer
Erlangen, September 2021



Die ASIM-Fachtagung „Simulation in Produktion und Logistik“ ist die regelmäßige Tagung der Fachgruppe „Simulation in Produktion und Logistik“ der Arbeitsgemeinschaft Simulation (ASIM). Die ASIM ist zugleich der Fachausschuss 4.5 der Gesellschaft für Informatik (GI).

Dieses Buch ist registriert als ASIM-Mitteilung Nummer 177.

Programmkomitee

- Chairmen: Jörg Franke, FAU Erlangen-Nürnberg (Erlangen, Germany)
Peter Schuderer, TH Ingolstadt (Ingolstadt, Germany)
- Hans-Peter Barbey (Germany)
Thomas Bauernhansl, Universität Stuttgart (Germany)
Eva Brandmeier, HAW Coburg (Germany)
Thorsten Claus, TU Dresden - IHI Zittau (Germany)
Uwe Clausen, TU Dortmund (Germany)
Jochen Deuse, TU Dortmund (Germany)
Toni Donhauser, Siemens Mobility GmbH (Germany)
Klaus Dröder, TU Braunschweig (Germany)
Johannes Fottner, TU München (Germany)
Jörg Franke, FAU Erlangen-Nürnberg (Germany)
Frank Herrmann, OTH Regensburg (Germany)
Steffen Ihlenfeldt, TU Dresden (Germany)
János Jósvai, Széchenyi István University (Hungary)
Angel A. Juan, Universitat Oberta de Catalunya (Spain)
Markus Klug, SSI Schäfer IT Solutions GmbH (Austria)
Bernd Kuhlenkötter, Ruhr-Universität Bochum (Germany)
Gisela Lanza, Karlsruher Institut für Technologie (Germany)
Christoph Laroque, Westsächsische Hochschule Zwickau (Germany)
Mathias Liewald, Universität Stuttgart (Germany)
Lothar März, STREMLER AG (Germany)
Jochen Merhof, HAW Coburg (Germany)
Joachim Metternich, TU Darmstadt (Germany)
Hans-Christian Möhring, Universität Stuttgart (Germany)
Alexander Pflaum, Otto-Friedrich Universität Bamberg (Germany)
Carsten Pöge, Volkswagen AG (Germany)
Bastian Prell, TH Wildau (Germany)
Markus Rabe, TU Dortmund (Germany)
- Andreas Rinkel, HSR Hochschule für Technik Rapperswil (SUISSE)
Oliver Rose, Universität der Bundeswehr München (Germany)
Jürgen Roßmann, RWTH Aachen (Germany)
Johannes Schilp, Universität Augsburg (Germany)
Andreas Schlegel, Fraunhofer IWU (Germany)
Robert Schmitt, WZL der RWTH Aachen (Germany)
Monika Schneider, RSE Fabrik-/Logistik-Planungsgesellschaft mbH (Germany)
Peter Schuderer, Technische Hochschule Ingolstadt (Germany)
Sven Spieckermann, SimPlan AG (Germany)
Dirk Steinhauer, SimPlan AG (Germany)
Patricia Stock, REFA-Institut e.V. (Germany)
Johannes Stoldt, Fraunhofer IWU (Germany)
Steffen Straßburger, TU Ilmenau (Germany)
Sebastian Thiede, University of Twente (Netherlands)
Margit Thomsen, INCONTROL Simulation Solutions (Germany)
Alexander Verl, Universität Stuttgart (Germany)
Markus Vorderwinkler, University of Applied Sciences Upper Austria (Austria)
Matthias Wenk, Ostbayerische Technische Hochschule Amberg-Weiden (Germany)
Sigrid Wenzel, Universität Kassel (Germany)



Contents

DIGITALER ZWILLING UND DIGITALER SCHATTEN

DIGITAL TWIN AND DIGITAL SHADOW

Die neue VDI-Richtlinie zur Prognose von Umgebungseinflüssen in der Digitalen Fabrik <i>Zülch, Gert; Karlsruher Institut für Technologie; Keller, Volker; Stadler Rail Service Deutschland</i>	1
Digital-physische Verbundkonzepte: Gegenüberstellung, Nutzeffekte und kritische Hürden <i>Scheer, Richard; Robert Bosch GmbH; Straßburger, Steffen; TU Ilmenau; Knapp, Marc; Robert Bosch GmbH</i>	11
Simulationsbasierte Bewertung von kombinierten Lean und Industrie 4.0 Maßnahmen in wandlungsfähigen Produktionssystemen <i>Martin, Niels L.; Dér, Antal; Langer, Adrian; Henningsen, Nadja; Ortmeier, Christian; Abraham, Tim; Herrmann, Christoph; TU Braunschweig</i>	21
Der Digitale Zwilling als echtzeitnahes Fertigungsabbild <i>Selmaier, Andreas; Herbert, Meike; Sjarov, Martin; FAU Erlangen-Nürnberg; Distler, Johannes; Fürst, Jens; Siemens Healthcare GmbH; Franke, Jörg; FAU Erlangen-Nürnberg</i>	31
Von der Simulation zum Experimentierbaren Digitalen Zwilling und zurück <i>Schluse, Michael; Roßmann, Jürgen; RWTH Aachen</i>	41

ENERGIEEFFIZIENZ, -FLEXIBILITÄT UND NACHHALTIGKEIT

ENERGY EFFICIENCY, ENERGY FLEXIBILITY AND SUSTAINABILITY

Modulare Simulation bei der energieorientierten Planung industrieller Heiz- und Kühlsysteme <i>Moog, Daniel; Borst, Fabian; Weigold, Matthias; TU Darmstadt</i>	51
Simulationsbasierte Optimierung zur Konfiguration von ökoeffizienten Supply Chains <i>Schreiber, Lucas; Fraunhofer Institut für Materialfluss und Logistik (IML); Niehus, Christian; TU Dortmund; Moroff, Nikolas; Fraunhofer Institut für Materialfluss und Logistik (IML)</i>	61
Analyse von konventionellen Prioritätsregeln zur Reduktion von CO ₂ -Emissionen durch den Einsatz von Photovoltaikanlagen <i>Terbrack, Hajo; TU Dresden; Claus, Thorsten; IHI Zittau; Götz, Matthias; TU Dresden; Herrmann, Frank; Ostbayerische Technische Hochschule Regensburg; Selmair, Maximilian; BMW Group</i>	75



Material- und Energieflusssimulation zur prädiktiven Bestimmung von Ofenreinigungsintervallen <i>Dettelbacher, Johannes; Schlüter, Wolfgang; Hochschule Ansbach</i>	85
Eine Datenbank für klassifizierte Forschungs- und Anwendungsberichte zu energieorientierter Simulation in Produktion und Logistik <i>Stoldt, Johannes; Fraunhofer IWU; Prell, Bastian; TH Wildau; Rabe, Markus; TU Dortmund; Wenzel, Sigrid; Universität Kassel; Thiede, Sebastian; University of Twente</i>	93
Simulation-based assessment of energy demand and costs associated with production scrap in the battery production <i>Ventura Silva, Gabriela; Thomitzek, Matthias; Abraham, Tim; Herrmann, Christoph; TU Braunschweig</i>	103
EREIGNISDISKRETE SIMULATION	
DESCRETE EVENT SIMULATION	
A modular, discrete-event simulation framework for modelling free ranging transportation vehicles in intralogistics <i>Reith, Karl-Benedikt; Rank, Sebastian; Schmidt, Thorsten; TU Dresden</i>	113
Referenzmodell zur wertstrombasierten Simulation von Unstetigförderern in der Grobplanungsphase von Produktionssystemen <i>Rabe, Markus; TU Dortmund; Wincheringer, Walter; Sohny, Tobias; Hochschule Koblenz</i>	123
Application of discrete-event simulation for factory planning - A case study <i>Herbert, Meike; Selmaier, Andreas; Mühlmann, Felix; FAU Erlangen-Nürnberg; Fürst, Jens; Siemens Healthcare GmbH; Franke, Jörg; FAU Erlangen-Nürnberg</i>	133
Einsatz der ereignisdiskreten Materialflusssimulation bei Methoden aus der kontinuierlichen Verbesserung im Fertigungsumfeld: Erfolg und Auswirkungen der Speedweek 4.0 <i>Sitz, Sarah; Zerreis, Maximilian; Robert Bosch GmbH; Lechler, Tobias; Franke, Jörg; FAU Erlangen-Nürnberg</i>	143
Methoden und Konzepte des Digitalen Logistikzwillings der AG der Dillinger Hüttenwerke (Dillinger) <i>Herzog, Alexander; TU Clausthal; Necil, Jan; Pollet, Martin; Busch, Heike; AG der Dillinger Hüttenwerke</i>	157
Einsatz ereignisdiskreter Simulation zur simultanen Allokation von Bestands-, Kapazitäts- und Zeitpuffern in variabilitätsbeeinflussten Produktionssystemen <i>Lenze, David; TU Dortmund; Schallow, Julian; RIF e.V. - Institut für Forschung und Transfer; Deuse, Jochen; TU Dortmund</i>	167

INTRALOGISTIK, LIEFERKETTEN UND PRODUKTIONSNETZWERKE

INTRALOGISTICS, SUPPLY CHAINS AND PRODUCTION NETWORKS

- Simulationsbasierte Validierung eines automatisierten Produktions- und Materialflusssteuerungssystems
Lichtenstern, Isabella; Wucherer, Stefanie; Hochschule für angewandte Wissenschaften Augsburg;
Klarmann, Steffen; Valeo Schalter und Sensoren GmbH;
Kerber, Florian; Hochschule für angewandte Wissenschaften Augsburg..... 177
- Einfluss von Störungen auf den Umgang mit Deadlocks in einem fahrerlosen Transportsystem
Müller, Marcel; Ulrich, Jan Hendrik; Otto-von-Guericke-Universität Magdeburg;
Reyes Rubiano, Lorena Silvana; University of La Sabana;
Reggelin, Tobias; Zadek, Hartmut;
Otto-von-Guericke-Universität Magdeburg 187
- Simulationsbasierte Untersuchung der Grenzproduktivität von Robotern in einem AutoStore-Lagersystem
Galka, Stefan; Scherbarth, Christoph; OTH Regensburg 197
- Simulationsmodell mit 3D-Animation zur schnellen Bewertung von Ablaufplänen in der Produktion
Rolf, Benjamin; Reggelin, Tobias; Lang, Sebastian; Müller, Marcel;
Prehm, Johann; Otto-von-Guericke-Universität Magdeburg..... 207
- Demand Driven Resupply of Offshore Components by Cascading Simulation and Linear Optimization
Rippel, Daniel; Lütjen, Michael; Universität Bremen;
Szczerbicka, Helena; Universität Hannover;
Freitag, Michael; Universität Bremen 217
- Utilising Relations between Actions to Improve the Performance of Optimisation Procedures for Distribution Networks
Rabe, Markus; TU Dortmund;
Ammouriova, Maja; Universitat Oberta de Catalunya..... 227
- Analytical and Empirical Study of Proper Parameters for TOC under Uncertainty
Rükgauer, Andreas; FHWS Würzburg-Schweinfurt 237

PRODUKTIONSPROZESSE

PRODUCTION PROCESSES

- Die Materialflusssimulation im Kontext eines Produktlebenszyklusmanagements
Pöge, Carsten; Meyer, Torben; Volkswagen AG..... 247
- Optimierung der Auslastung eines Matrix-Montage-Systems durch Konzeptionierung und Implementierung eines Genetischen Algorithmus für Maschinenbelegungsplanung
Purgander, Jana; Hinckeldeyn, Johannes; TU Hamburg 257



Modularer Simulationsbaukasten zur Modellierung der Instandhaltungsabläufe in Zugdepots <i>Donhauser, Toni; Siemens Mobility GmbH;</i> <i>Natterer, Niklas; Siemens Advanta Consulting;</i> <i>Genc, Emin; Siemens Mobility GmbH</i>	267
Collaborative Virtual Reality System for Industrial based Assembly Training <i>Müller, Rainer; Hörauf, Leenhard; Eichenwald, Max; ZeMA gGmbH</i>	275
Simulationsbasierte Konfiguration der Stationen von Matrix-Produktionssystemen <i>Völker, Sven; Verbeet, Richard; Technische Hochschule Ulm</i>	285
Rückwärtssimulation als Instrument zur Produktionsplanung - Erkenntnisse aus einer praxisbezogenen Fallstudie <i>Schneider, Germar; Scholl, Wolfgang;</i> <i>Infineon Technologies Dresden GmbH & Co. KG;</i> <i>Leissau, Madlene; Laroque, Christoph;</i> <i>Westfälische Hochschule Zwickau</i>	295
SIMULATION MIT DATA ANALYTICS UND MACHINE LEARNING	
SIMULATION WITH DATA ANALYTICS AND MACHINE LEARNING	
Scalable cooperative Multi-Agent-Reinforcement-Learning for order-controlled on schedule manufacturing in flexible manufacturing systems <i>Denkena, Berend; Dittrich, Marc-André; Fohlmeister, Silas; Kemp, Daniel;</i> <i>Palmer, Gregory; Leibniz Universität Hannover</i>	305
A review on simulation metamodeling for decision support systems using artificial neural networks <i>Billiet, Harold; Stark, Rainer; TU Berlin</i>	315
Towards Production-Ready Reinforcement Learning Scheduling Agents: A Hybrid Two-Step Training Approach Based on Discrete-Event Simulations <i>Kemmerling, Marco; Samsonov, Vladimir; Janke, Tim; Lütticke, Daniel;</i> <i>Cybernetics Lab IMA & IfU;</i> <i>Gützlaß, Andreas; Schmidhuber, Matthias; WZL der RWTH Aachen;</i> <i>Meisen, Tobias; Cybernetics Lab IMA & IfU;</i> <i>Schuh, Günther; WZL der RWTH Aachen</i>	325
Entscheidungsbäume und bestärkendes Lernen zur dynamischen Auswahl von Reihenfolgeregeln in einem flexiblen Produktionssystem <i>Heger, Jens; Abdine, Mazhar Zein El; Sekar, Sughanthan;</i> <i>Voß, Thomas; Leuphana Universität Lüneburg</i>	337
An Approach for Deep Reinforcement Learning for Production Program Planning in Value Streams <i>West, Nikolai; Hoffmann, Florian; Schulte, Lukas; TU Dortmund;</i> <i>Hernandez Moreno, Victor; University of Technology;</i> <i>Deuse, Jochen; TU Dortmund</i>	347



Simulative dispatching optimization of maintenance resources in a semiconductor use-case using reinforcement learning <i>Altenmüller, Thomas; Hoffmann, Clara; Infineon Technologies AG; May, Marvin Carl; Kuhnle, Andreas; Lanza, Gisela; Karlsruher Institut für Technologie</i>	357
Automatisierung im Prozess der Wissensentdeckung in Simulationsdaten - Charakterisierung der Ergebnisdaten <i>Genath, Jonas; Bergmann, Sören; Feldkamp, Niclas; Straßburger, Steffen; TU Ilmenau</i>	367
Entwicklung einer integrierten Lösung für das Data Farming und die Wissensentdeckung in Simulationsdaten <i>Genath, Jonas; Bergmann, Sören; TU Ilmenau; Spieckermann, Sven; Stauber, Stephan; SimPlan AG; Feldkamp, Niclas; TU Ilmenau</i>	377
Simulationsgestützter Ansatz zur Verbesserung der Leistungsprognose maschineller Tunnelvortriebe mithilfe baubegleitender Prozessdaten <i>Jodehl, Annika; Salloum, Yara; König, Markus; Thewes, Markus; Ruhr-Universität Bochum</i>	387
Simulation-based demand forecast generation to analyze forecast accuracy and its influence on logistical performance <i>Felberbauer, Thomas; St. Pölten UAS; Seiringer, Wolfgang; Altendorfer, Klaus; UAS Upper Austria</i>	399
Combining Engineering Data, Sensor Data and Artificial Intelligence for automated edge network infrastructures <i>Hürkamp, André; Aschersleben, Franziska; Czarski, Marvin; Gellrich, Sebastian; Herrmann, Christoph; Dröder, Klaus; TU Braunschweig</i>	409
SIMULATION UND OPTIMIERUNG	
SIMULATION AND OPTIMIZATION	
Entwicklung einer dynamischen Routenfindung in hybriden Mensch-Roboter-Kommissioniersystemen <i>Kauke, Dirk; Rett, Angelika; Fottner, Johannes; TU München</i>	421
Improving Blood Supply Chain CrisisManagement by Simulation-based Optimization <i>Horstkemper, Dennis; WWU Münster; Reuter-Oppermann, Melanie; TU Darmstadt; Middelhoff, Michael; Widera, Adam; Hellingrath, Bernd; WWU Münster</i>	433



System concept for semi-automated generation of layouts for simulation models based on point clouds <i>Zeiser, Reinhard; Fraunhofer IGCV;</i> <i>Ullmann, Felix; Hörmann Rawema Engineering & Consulting GmbH;</i> <i>Neuhäuser, Thomas; Fraunhofer IGCV;</i> <i>Hohmann, Andrea; Fraunhofer IGCV;</i> <i>Schilp, Johannes; Universität Augsburg</i>	443
Exploiting the potential of human-machine work systems: Cycle-time reduction through simulation-based analysis and optimized task allocation <i>Komenda, Titanilla; Schelle, Maximilian; Kamhuber, Felix;</i> <i>Fraunhofer Austria Research GmbH;</i> <i>Schlund, Sebastian; TU Wien</i>	453
Modelling of Microstructures during in-situ Alloying in Additive Manufacturing for efficient Material Qualification Processes <i>Zimbrod, Patrick; Schilp, Johannes; Universität Augsburg</i>	463
Simulationsbasierte Optimierung von Bestelllosgrößen in der verbrauchsgesteuerten Materialdisposition der Investitionsgüterindustrie <i>Schmid, Alexander; Fraunhofer Austria Research GmbH;</i> <i>Lielacher, Magdalena; TU Wien;</i> <i>Sobotka, Thomas; Fraunhofer Austria Research GmbH;</i> <i>Sihn, Wilfried; TU Wien</i>	475
Zuweisung bester Abstellplätze im LIFO-Lager unter Berücksichtigung der Kommissionierungsreihenfolge mittels genetischem Algorithmus <i>Kuhn, Dominik; Adelsbach, Jan; Bashir, Attique; Hörauf, Leenhard;</i> <i>Müller, Rainer; ZeMA gGmbH</i>	485
Interoperabilität realer und simulierter Produktionssysteme mittels OPC UA <i>Schmidl, Elisabeth; Wenk, Matthias; OTH Amberg-Weiden;</i> <i>Russwurm, Eva; Franke, Jörg; FAU Erlangen-Nürnberg</i>	495
TRANSPORTLOGISTIK UND -SYSTEME	
TRANSPORT LOGISTICS AND -SYSTEMS	
Ein unternehmensübergreifender Standard für Logistiksimulation mit grafischer Prozessmodellierung <i>Sokoll, Kristina; AUDI AG;</i> <i>von Braunschweig, Philipp; Volkswagen AG;</i> <i>Gustafsson-Ende, Linda; BMW Group;</i> <i>Hilmer, Frank; SimPlan AG</i>	505
A Simulation Study on Electric Last Mile Delivery with Mobile Smart Cargo Boxes <i>Davidsson, Paul; Johansson, Emil; Lorig, Fabian; Persson, Jan;</i> <i>Malmö University</i>	515
Entwicklung eines anwendungsorientierten Bausteinkastens zur Simulation kombinierter Transportmodelle mittels autonomer Fahrzeuge <i>Staritz, Johannes; Auf der Landwehr, Marvin; Trott, Maik;</i> <i>von Viebahn, Christoph; Hochschule Hannover</i>	525



Direktumschlag an der Kaikante - Eine Machbarkeitsstudie für hafenerinterne Containertransporte <i>Nellen, Nicole; Lange, Ann-Kathrin; Jahn, Carlos; TU Hamburg</i>	535
Simulation von kombinierter Stau- und Routenplanung in Multi-Terminal Häfen <i>Franzkeit, Janna; TU Hamburg;</i> <i>Voß, Thomas; Leuphana Universität Lüneburg;</i> <i>Pache, Hannah; TU Hamburg;</i> <i>Heger, Jens; Leuphana Universität Lüneburg;</i> <i>Jahn, Carlos; TU Hamburg</i>	545
Erarbeitung einer Prozesssteuerungsstrategie für zwei Transportmittel mit gemeinsamen Aktionsbereich am Beispiel eines Prozesskransystems zur Kommissionierung von Schüttgut <i>Betker, Vincent; Völker, Michael; Schmidt, Thorsten; TU Dresden</i>	555
VERIFIKATION, VALIDIERUNG UND VIRTUELLE INBETRIEBNAHME VERIFICATION, VALIDATION AND VIRTUAL COMMISSIONING	
Generierung realitätsnaher Testdaten für die Simulation von Produktionen <i>Krockert, Martin; Matthes, Marvin; Munkelt, Torsten;</i> <i>Hochschule für Technik und Wirtschaft Dresden;</i> <i>Völker, Sven; Hochschule Ulm</i>	565
Remote-Lernfabrik - Simulationsmodelle im Anlagenentstehungsprozess und deren Integration in die Virtuelle Inbetriebnahme <i>Massow, Benjamin Bastian; Hausberger, Thomas; Geiger, Dominique Mathäus;</i> <i>Klotz, Matthias; Die Unternehmerische Hochschule MCI;</i> <i>Schlegel, Andreas; Putz, Matthias; Fraunhofer IWU</i>	575
Konzept zur Integration eines kontinuierlichen Materialflussmodells in die Virtuelle Inbetriebnahme durch Signalgenerierung <i>Kienzlen, Annika; Verl, Alexander; Universität Stuttgart</i>	585
Virtuelle Inbetriebnahme eines Leitsystems für die roboterbasierte automatische Kommissionierung in der Automobilindustrie <i>Xie, Xinyi; Büttner, Thomas; Friedemann, Marco; Wenzel, Ken;</i> <i>Fraunhofer IWU</i>	595
Continuous validation and precise updating for high accuracy of digital twins of production system <i>Overbeck, Leonard; Le Louarn, Arthur; Brützel, Oliver; Stricker, Nicole; Lanza, Gisela; Karlsruher Institut für Technologie</i>	609
Investigation and evaluation of 3D recording methods for use cases in production planning <i>Metzner, Maximilian; Schiessel, Svenja; Siemens AG;</i> <i>Grünhöfer, Lukas; Franke, Jörg; FAU Erlangen-Nürnberg</i>	619
Physikalische Sensorsimulation zur Verifikation und Validierung von optischen Systemen <i>Thieling, Jörn; Roßmann, Jürgen; RWTH Aachen</i>	629



Entwicklung und Erprobung eines Vorgehens zur Validierung von
Simulationsbausteinen komplexer Werkzeugmaschinen
Tjaden, Greta; Mieth, Carina;
TRUMPF Werkzeugmaschinen GmbH & Co. KG..... 641

WIRTSCHAFTLICHKEIT DURCH SIMULATION

ECONOMIC EFFICIENCY THROUGH SIMULATION

Simulation der Einplanung der terminkritischen Montage von Abfüllanlagen am
Beispiel der Krones AG
Herrmann, Frank; OTH Regensburg;
Savasci, Faruk; KRONES AG 653

Notations in the simulation development: A state-of-the-art Literature Research
Trott, Maik; Auf der Landwehr, Marvin; von Viebahn, Christoph;
Hochschule Hannover 663



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Die neue VDI-Richtlinie zur Prognose von Umgebungseinflüssen in der Digitalen Fabrik

The new VDI guideline for the prognosis of environmental influences in the digital factory

Gert Zülch, Karlsruher Institut für Technologie, Karlsruhe (Germany),
gert.zuelch@gefora-beratung.de

Volker Keller, Stadler Rail Service Deutschland, Berlin (Germany),
volker.keller@stadlerail.com

Abstract: In order to regard working humans in simulation models and tools of the digital factory, three main aspects can be distinguished: One of them looks at humans as resources of the work system and analyses their performance in terms of production logistics in connection with the technical facilities. The second aspect includes human models with regard to their anthropometry and work physiology. These two aspects have already been reflected in former VDI guidelines. A new guideline deals with the influences of the indoor work environment. According to a fundamental ergonomic concept, on the one hand the physical influences are to be considered, on the other hand their effects on a working human. In addition, a distinction must be made between point-in-time related and time-period related analyses. For the latter in particular, the coupling to discrete event-driven simulation is necessary. The new VDI guideline will be explained here in more detail.

1. Problemstellung und Bezug zur Simulation

Der vorliegende Beitrag greift ein Thema auf, das bereits während der ASIM-Fachtagung 2019 in Chemnitz behandelt wurde, nämlich die Personaleinsatzsimulation im Rahmen der Digitalen Fabrik (Zülch 2019). Bei diesem Beitrag ging es um den Stand der Technik zur ergonomischen Analyse des arbeitenden Menschen in der Digitalen Fabrik (siehe Zülch 2020). Dazu wurde in der VDI-Richtlinie 3633 Blatt 6 aufgezeigt, wie die zeitliche Arbeitsbelastung und weitere makro-ergonomische Aspekte in existierenden Simulationsverfahren behandelt werden können. In mikro-ergonomischer Hinsicht wurden Verfahren zur anthropometrischen und arbeitsphysiologischen Arbeitsgestaltung angesprochen, die in der VDI-Richtlinie 4499 Blatt 4 behandelt werden (siehe zum Überblick Zülch 2013).

Die beiden genannten VDI-Richtlinienblätter behandeln gemäß dem Belastungs-Beanspruchungs-Konzept nach Rohmert (1984) nur einen ergonomischen Gesichtspunkt, nämlich die Effekte auf den arbeitenden Menschen durch die Arbeitsaufgabe selbst. Es fehlt also demnach noch, wie Belastungen und Beanspruchungen durch Einflüsse aus der (vorwiegend physikalischen) Arbeitsumgebung (siehe Abb. 1) in Werkzeugen der Digitalen Fabrik aufgegriffen werden können. Dieser Fragestellung wird im vorliegenden Beitrag anhand der neuen VDI-Richtlinie 4499 Blatt 5 nachgegangen (siehe VDI 2020). Da sich dieses Richtlinienblatt zurzeit in der Endphase seiner Veröffentlichung befindet, kann hier nur übersichtsweise auf einige methodische Punkte eingegangen werden.

2. Abgrenzung zu bisherigen Arbeiten

Das neue Richtlinienblatt ergänzt somit die vorigen, die sich mit der Einbeziehung des arbeitenden Menschen in die Digitale Fabrik befassen. Dazu lässt sich feststellen, dass es zwar – auch aufgrund des Arbeits- und Gesundheitsschutzes (siehe die geforderte Gefährdungsbeurteilung in § 5 f. ArbSchG) – umfangreiche Erkenntnisse zu den Effekten der Arbeitsumgebung auf den Menschen und zu rechtlichen Vorgaben hierzu gibt. Für deren Prognose existieren auch in Deutschland verfügbare Softwareverfahren, diese sind aber bis auf eine unten näher bezeichnete Ausnahme nicht in Verfahren der Digitalen Fabrik eingebunden. Die vorhandenen Verfahren stammen dabei nicht nur aus dem arbeitswissenschaftliche Umfeld, sondern zu einem großen Teil auch aus der Bauphysik und der Gebäudeplanung. Insofern stellt die Prognose von Umgebungseinflüssen auf den arbeitenden Menschen ein neues Feld der Digitalen Fabrik dar.

Aus dem internationalen Bereich sind kaum entsprechende Prognoseverfahren bekannt geworden. In der International Ergonomics Association gibt zwar ein Technical Committee „Digital Human Modeling and Simulation“, dieser Fachkreis befasst sich aber vorrangig mit der Menschmodellierung (siehe z.B. Briceno und Paul 2019). Auch in internationalen Ergonomiekongressen (z.B. Nunes 2020) und in neueren englischsprachigen Veröffentlichungen (z.B. Neumann et al. 2021) findet sich kaum ein Hinweis hierauf. Im nationalen Bereich gibt es ebenfalls nur wenige Veröffentlichungen, die auf die Bedeutung der Prognose von Umgebungseinflüssen hinweisen (siehe zu einer Ausnahme Illmann et al. 2015).

3. Vorgehensweise und Ergebnisse der Richtlinienerstellung

Den Anstoß zur Entwicklung der neuen VDI-Richtlinie 4499 Blatt 5 gab der übergeordnete VDI-Fachausschuss „Digitale Fabrik“. Dieser beauftragte dann einen speziell dafür eingerichteten Richtlinienausschuss mit der Erarbeitung. Die Entwicklung des Richtlinienblattes erstreckte sich seit deren Beginn im Jahre 2013 bisher über acht Jahre und umfasste in diesem Zeitraum eine Vielzahl von Expertentreffen.

Gegen Ende des Jahres 2020 legte der beauftragte Richtlinienausschuss eine Entwurfsfassung vor, die dann vom Fachausschuss genehmigt wurde (Abb. 1). Danach erfolgt nun eine Bearbeitung durch die VDI-Redaktion. Die daraus

resultierende vorläufige Fassung, der sog. Gründruck, wird voraussichtlich im Jahre 2021 veröffentlicht und damit für etwaige Einwände freigegeben.

VDI-RICHTLINIEN		
VEREIN DEUTSCHER INGENIEURE	Digitale Fabrik Prognose von Umgebungseinflüssen auf den arbeitenden Menschen	VDI 4499 Blatt 5 Entwurf
Vorbemerkung		
Einleitung		
Gliederung dieses Richtlinienblattes		
1	Anwendungsbereiche	
2	Begriffe	
3	Grundlagen	
3.1	Belastungs-Beanspruchungs-Konzept	
3.2	Bewertungs- und Beurteilungsmethoden	
3.3	Rechtliche Rahmenbedingungen	
4	Analyse von Umgebungseinflüssen	
4.1	Arbeitsraum (Geometrie, Reflexion, Absorption)	
4.2	Raumluft (Klima, Luftströmung, Gefahrstoffausbreitung)	
4.3	Mechanische Schwingungen (Lärm, Vibrationen)	
4.4	Elektromagnetische Schwingungen (Beleuchtung, Elektromagnetische Felder, Radioaktive Strahlung)	
5	Anwendungsbeispiele	
5.1	Prognose kombinierter Umgebungsbelastungen	
5.2	Visualisierung von Umgebungseinflüssen in einer CAVE	
5.3	Resümee der Prognose von Umgebungseinflüssen	

Abbildung 1: Gründruck-Entwurf der VDI-Richtlinie 4499 Blatt 5

Erst nach Behandlung etwaiger Einwände erfolgt dann die endgültige Veröffentlichung im Beuth-Verlag in Berlin. In jüngerer Zeit werden die VDI-Richtlinien in Deutsch und Englisch veröffentlicht. Da sich der VDI wesentlich durch seine Richtlinien finanziert, sind die dann zur Verfügung stehenden Print- und Internetversionen kostenpflichtig.

Zu jedem der als wesentlich betrachteten Umgebungseinflüsse wird im Richtlinienblatt zumindest ein Anwendungsbeispiel beschrieben, das auf einem in Deutschland verfügbaren Softwareverfahren basiert. An dieser Stelle werden im vorliegenden Beitrag – unter Rückgriff auf frühere Arbeiten – einige grundsätzliche Probleme bei der Prognose von Umgebungseinflüssen behandelt.

4. Besondere Aspekte der Prognose von Umgebungseinflüssen

4.1 Arten der Prognose von Umgebungseinflüssen

In der VDI-Richtlinie 4499 Blatt 5 werden die auf den arbeitenden Menschen in einem umschlossenen Arbeitsraum wirkenden Umgebungseinflüsse behandelt.

Dabei wird gemäß dem Belastungs-Beanspruchungs-Konzept bei jedem der behandelten Einflüsse nach Belastung und Beanspruchung sowie nach deren Bewertung bzw. Beurteilung anhand normativer Vergleichswerte unterschieden (Abb. 2). Weiterhin wird eine Unterscheidung getroffen in personen- bzw. arbeitsplatzbezogene sowie nach zeitpunkt- und zeitraumbezogenen Prognosen. Für die zuletzt genannte Prognoseart ist die Verbindung mit einem ereignisdiskreten Simulationsverfahren erforderlich.

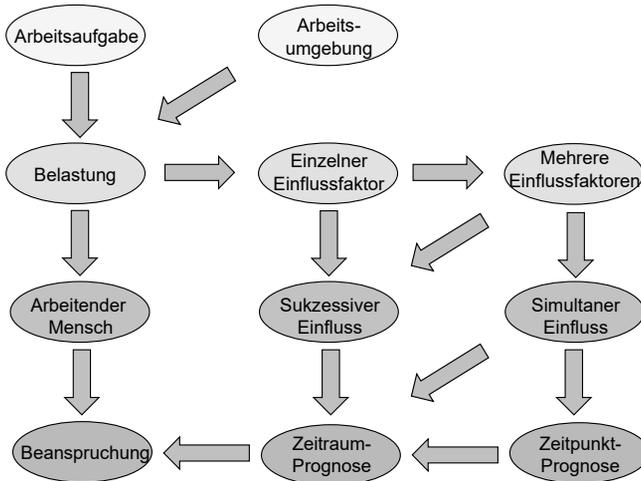


Abbildung 2: Arten der Belastungs-Beanspruchungs-Prognosen

Bei den Verfahren zeigen sich im praktischen Anwendungsfall oftmals Schwierigkeiten. Nachfolgend wird auf einige typische Problembereiche hingewiesen, die so auch bei Analysen in der Realität auftreten können. Bei den hier erörterten Anwendungsfällen handelt es sich um Zeitpunkt-bezogene Prognosen von Umgebungsbelastungen.

4.2 Validierung des Prognoseverfahrens

Eine wichtige Fragestellung betrifft die Validität des verwendeten Prognoseverfahrens, also im vorliegenden Fall die Übereinstimmung von gemessenen und berechneten Werten eines einzelnen Umgebungseinflusses an einem Arbeitsplatz. Da in der Literatur keine entsprechenden Untersuchungen für Umgebungseinflüsse gefunden werden konnten, wird hier auf eine eigene frühere Studie zurückgegriffen (Keller 2002, S. 160 ff.).

Gegenstand dieser Validierungsstudie ist eine Werkstatt für die Produktion von Werkzeugen, in der neben fünf Maschinen für die spanende Bearbeitung auch ein Arbeitsplatz für die Montage der Werkzeuge vorhanden ist. Abbildung 3 zeigt das digitale Modell der Werkstatt und beispielhaft die Zeitpunkt-bezogene Berechnung des Schalldruckpegels am Montageplatz.

Zur Berechnung der Beleuchtungsstärken wurde das Radiosity-Verfahren (Ashdown 1994) verwendet. Der Vergleich zwischen den gemessenen und berechneten Werten

zeigte im Grundsatz eine gute Übereinstimmung (Tab. 1). Am Montageplatz und an den Maschinenbedienplätzen ergaben sich Abweichungen von 10 bis 27 lx, die vom menschlichen Auge kaum wahrgenommen werden können.

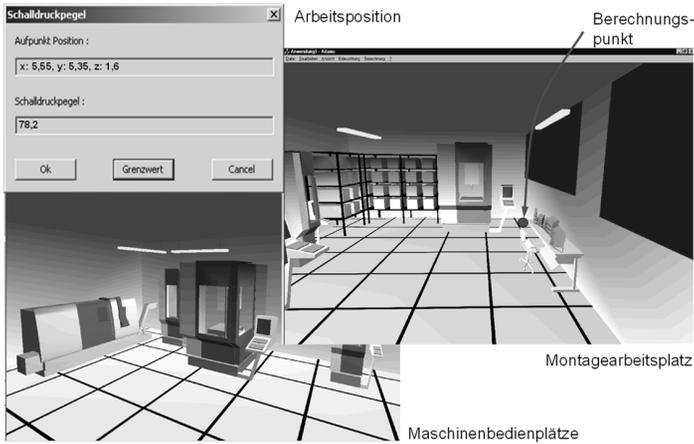


Abbildung 3: Werkzeugbau mit Berechnung des Schalldruckpegels an einem Arbeitsplatz (Quelle: Keller 2002, S. 159, ergänzt)

Tabelle 1: Gemessene und berechnete Umgebungsbelastungen in der betrachteten Werkstatt (nach Keller 2002, S. 161 f.)

Nr.	Arbeitsplatz	Beleuchtungsstärke, in lx		
		Messwert	Berechn.Ist	Abweichung
1	Drehmaschine	299	313	14
2	Sägemaschine	231	204	27
3	Fräsmaschine	310	328	18
4	Bohrmaschine	211	221	10
5	Schleifmaschine	95	83	12
6	Montageplatz	303	317	14

Nr.	Arbeitsplatz	Schalldruckpegel, in dB(A)		
		Messwert	Berechn.Ist	Abweichung
	Immission	jede Maschine einzeln im Leerlauf außer Betrieb		
1	Drehmaschine			
2	Sägemaschine	65,8	68,0	2,2
3	Fräsmaschine	80,3	77,3	3,0
4	Bohrmaschine	65,4	67,4	2,0
5	Schleifmaschine	67,9	65,8	2,1
6	Montageplatz Masch. 2 bis 5 in Betrieb	80,2	78,4	1,8

Die Schallausbreitung wurde für jede einzelne Maschine im Leerlauf mit dem Spiegelquellen-Verfahren nach VDI 3760:1992-02 berechnet. Beim Schalldruckpegel zeigten sich deutliche Abweichungen, die maximal 3,0 dB(A) erreichten. Die Ungenauigkeit liegt damit im Bereich orientierender Messgeräte und ist daher im Rahmen der Digitalen Fabrik als akzeptabel anzusehen. Die Abweichungen hängen aber auch von der verwendeten Berechnungsmethode ab. Eine zusätzliche Ungenauigkeit ist darin begründet, dass zur Vermeidung hoher Rechenzeiten sämtliche Einrichtungen und Betriebsmittel in der Werkstatt vereinfachend als Quader angenommen werden mussten (Keller 2002, S. 154).

4.3 Anwendungsbeispiel zur Umplanung einer Werkstatt

Die typische Anwendung einer Prognose von Umgebungseinflüssen in der Digitalen Fabrik besteht darin, alternative Gestaltungslösungen in einem Arbeitsraum zu analysieren. Abbildung 4 zeigt dazu – wiederum am Beispiel des obigen Werkzeugbaus - die Ausgangssituation und eine der möglichen Layout-Alternativen zur Verbesserung der Umgebungsbelastungen (Keller 2002, S. 156 ff.).

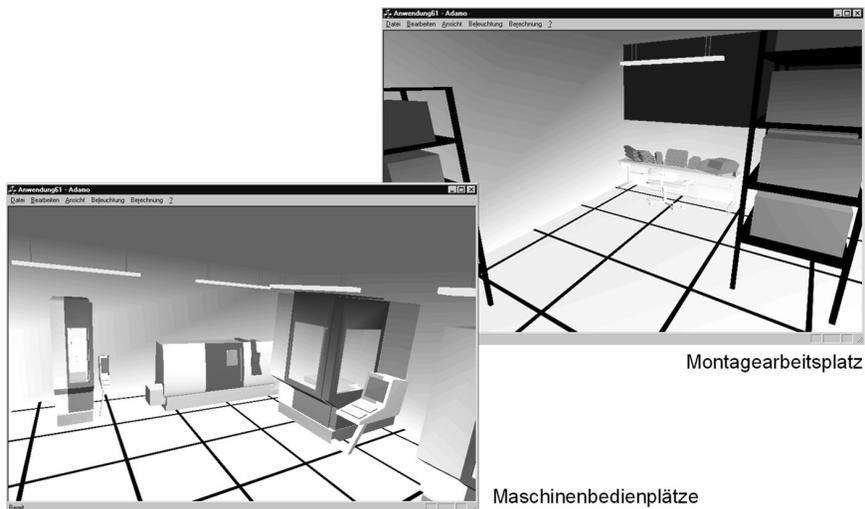


Abbildung 4: Layout-Alternative in der betrachteten Werkstatt zur Verbesserung der Umgebungsbelastungen (Keller 2002, S. 173)

Anhand dieser Planungsstudie wurden zur Verbesserungen des Ausgangszustands die Beleuchtungssituation und die Schallausbreitung an den Arbeitsplätzen analysiert und die Effekte in verschiedenen Layout-Alternativen prognostiziert (Tab. 2). Mit der Radiosity-Methode ergaben sich Beleuchtungsstärken, die an allen Arbeitsplätzen Werte erreichten, welche die Vorgaben nach DIN EN 12464-1:2003-03 voll erfüllten. Vor allem an der Schleifmaschine konnte die ungünstige Beleuchtungssituation von 83 lx auf 378 lx wesentlich verbessert werden.

Table 2: Planungsstudie zur Verbesserung der Umgebungsbelastungen in der betrachteten Werkstatt (nach Keller 2002, S. 170 ff.)

Nr.		Arbeitsplatz	Beleuchtungsstärke, in lx		
	Immission	Berechn.Ist	Berechn.Plan bei Layoutänderung	Bemerkung	
1	Drehmaschine	313	413		
2	Sägemaschine	204	349		
3	Fräsmaschine	328	356		
4	Bohrmaschine	221	386		
5	Schleifmaschine	83	378		
6	Montageplatz	317	466	Abschirmwand	

Nr.		Arbeitsplatz	Schalldruckpegel, in dB(A)		
	Immission	Berechn.Ist einzel im Leerlauf	Berechn.Plan bei Stahlbearb.	Bemerkung	
1	Drehmaschine	(82,5)	[89,0]	(Tabellenwert)	
2	Sägemaschine	68,0	[89,0]		
3	Fräsmaschine	77,3	[83,5]		
4	Bohrmaschine	67,4	[74,0]		
5	Schleifmaschine	65,8	[90,0]		
6	Montageplatz Masch. 1 bis 5 in Betrieb	78,4	(81,0) [90,1]	[Stahl Ck 45 N]	

Hierbei zeigten sich einige praktische Schwierigkeiten der Prognose: Die Schalldruckpegel an den Maschinenbedienplätzen wurden einzeln im Leerlauf der jeweiligen Maschine ermittelt, ohne dass zunächst die übrigen Maschinen in Betrieb waren (Tab. 1). Diese Annahme wurde für das relativ kleine Arbeitssystem mit nur wenigen Bedienpersonen als angemessen angesehen. Außerdem war aus betrieblichen Gründen die Drehmaschine bei der Berechnung ihres Immissionswertes ausgenommen. Hierfür wurde der Tabellenwert nach VDI 3742-1:1981-02 mit 82,5 dB(A) eingesetzt (Tab. 2). Unter diesen Voraussetzungen ergab sich speziell für den Montageplatz ein Schalldruckpegel von 81,0 dB(A).

Nimmt man aber an, dass an allen Werkzeugmaschinen Stahl bearbeitet wird, so erhält man dort einen Schalldruckpegel von 90,1 dB(A), was nach der Berufsgenossenschaftlichen Vorschrift VBG 121 (1990; jetzt LärmVibrationsArbSchV 2017) das Tragen persönlicher Gehörschutzmitteln erforderlich machen würde (ebd. § 8). Bei dieser Berechnung wurden die Schalldruckpegel der einzelnen Werkzeugmaschinen für die Bearbeitung der Stahlsorte Ck 45 N angesetzt (siehe Richtlinienreihe VDI 3742).

5. Weiterentwicklung von Umgebungsprognosen in der Digitalen Fabrik

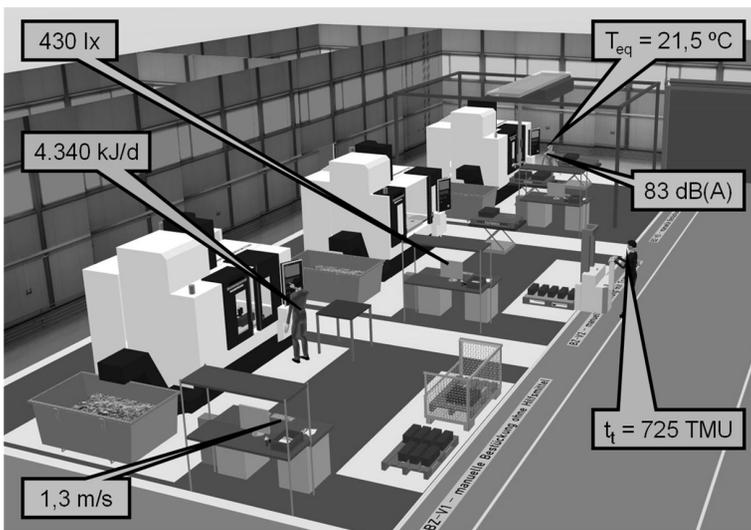
Für die Prognose von Umgebungseinflüssen ist bereits eine Vielzahl von Berechnungsmethoden vorhanden. Diese beziehen sich aber vorwiegend auf

einzelne Belastungsarten; die zugehörigen Softwareverfahren sind (bis auf wenige Ausnahmen) nicht in Verfahren der Digitalen Fabrik integriert. Bezüglich der resultierenden Beanspruchungen des arbeitenden Menschen gibt es erste Methoden und Softwareverfahren.

Außerdem handelt es sich bei den verfügbaren Verfahren in aller Regel um Zeitpunkt-bezogene Bewertungen, für Zeitraum-bezogene fehlen vielfach geeignete Methoden. Für die sukzessive Belastung durch einen einzelnen Umgebungseinfluss gibt es nur wenige Beispiele; diesbezüglich kann beispielhaft auf die Berechnungsmethode zur Bewertung eines schwankenden Schalldruckpegels hingewiesen werden (personenbezogener Tages-Lärmexpositionspegel nach § 2 Abs. 2 LärmVibrationsArbSchV).

Erst recht fehlt es an Bewertungsmethoden und -verfahren für simultan wirkende Umgebungseinflüsse, und dies dann auch für Kombinationen aus unterschiedlichen Belastungsarten durch die Arbeitsaufgabe und die Arbeitsumgebung. Ein Gegenbeispiel ist die Bewertung von energetisch belastender Hitzearbeit (siehe z.B. BGHM 2013, S. 14 ff.). Diesbezüglich fehlt es an integrierenden Methoden; ein hierzu entwickelter lexikographischer Ansatz bei gleichzeitig wirkenden Einflüssen (Zülch und Zülch 2017, S. 42) kann nur eine Behelfslösung darstellen.

Dies alles zeigt, dass es eine Vielzahl offener Fragestellungen zur Prognose von Umgebungseinflüssen in Werkzeugen der Digitalen Fabrik gibt. Bis zu einer umfassenden Prognose bedarf es also noch vielfacher Forschungsanstrengungen (Abb. 5).



Legende der Bewertungskriterien:

Produktionslogistik Arbeitsaufgabe Arbeitsumgebung

Abbildung 5: Vision der Prognose ergonomischer Belastungen in der Digitalen Fabrik (Layout-Graphik: Spitzhirn 2020)

Literatur

- ArbSchG: Gesetz über die Durchführung von Maßnahmen des Arbeitsschutzes zur Verbesserung der Sicherheit und des Gesundheitsschutzes der Beschäftigten bei der Arbeit (Arbeitsschutzgesetz - ArbSchG). Vom 7. August 1996 (BGBl. I S. 1246), zuletzt geändert durch Artikel 1 des Gesetzes vom 22. Dezember 2020 (BGBl. I S. 3334). <https://www.gesetze-im-internet.de/arbSchG/ArbSchG.pdf>, letzter Zugriff 28.01.2021.
- Ashdown, I.: Radiosity. New York, NY u.a.: John Wiley & Sons 1994.
- BGHM – Berufsgenossenschaft Holz und Metall: Hitzearbeit erkennen – beurteilen – schützen. Mainz: BGHM 2013 (BGI 579).
- Briceno, L.; Paul, G.: MakeHuman: A Review of the Modelling Framework. In: Bagnara, S.; Tartaglia, R.; Albolino, S.; Alexander, T.; Fujita, Y. (Hrsg.): Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018). Cham: Springer 2019, S. 224-232. (Advances in Intelligent Systems and Computing, Vol. 822). https://rd.springer.com/chapter/10.1007%2F978-3-319-96077-7_23, letzter Zugriff 28.01.2021.
- Brinkmeier, B.: Persönliche Mitteilung. Fellbach: Siemens Industry Software. Düsseldorf: VDI, Präsentation 13.01.2017.
- DIN EN 12464-1:2003-03: Licht und Beleuchtung - Beleuchtung von Arbeitsstätten – Teil 1: Arbeitsstätten in Innenräumen. Berlin: Beuth Verlag.
- Illmann, B.; Fritzsche, L.; Ullmann, S.; Leidholdt, W.: Ganzheitliche Gefährdungsbeurteilung mit digitalen Menschmodellen – Die Integration von Umgebungsbedingungen in die Digitale Fabrik. In: Gesellschaft für Arbeitswissenschaft (Hrsg.): VerANTWORTung für die Arbeit der Zukunft. Dortmund: GfA Press 2015, 6 S.
- IBP – Fraunhofer-Institut für Bauphysik: BNB-Tool Thermischer Komfort. Holzkirchen: Fraunhofer IBP, 2014. <https://www.ibp.fraunhofer.de/content/dam/ibp/ibp-neu/de/dokumente/produktblaetter/ee/produktblatt-bnb-tool.pdf>, letzter Zugriff 08.03.2021.
- Keller, V.: Ansatz zur objektorientierten Modellierung betrieblicher Arbeitsschutzdaten. Aachen: Shaker Verlag 2002 (ifab-Forschungsberichte aus dem Institut für Arbeitswissenschaft und Betriebsorganisation der Universität Karlsruhe, Band 25).
- Keller, V.; Zülch, G.: Repräsentation von Arbeitsschutzdaten und ihre Visualisierung in rechnerunterstützten Managementsystemen. In: Zülch G, Brinkmeier, B. (Hrsg.): Arbeitsschutz-Managementsysteme. Aachen: Shaker 2000, S. 219–229.
- LärmVibrationsArbSchV: Lärm- und Vibrations-Arbeitsschutzverordnung. Vom 06.03.2007 (BGBl. I S. 261); zuletzt geändert durch Art. 5 Abs. 5 der Verordnung vom 18.10.2017, BGBl. I Nr. 69 vom 23.10.2017, S. 3584. https://www.gesetze-im-internet.de/l_rmVibrationsarbschv/L%C3%A4rmVibrationsArbSchV.pdf, letzter Zugriff 21.04.2021.
- Neumann, W. P.; Winkelhaus, S.; Grosse, E. H.; Glock, C. H.: Industry 4.0 and the human factor – A systems framework and analysis methodology for successful development. International Journal of Production Economics, Amsterdam, 233 (2021), 18 S. <https://doi.org/10.1016/j.ijpe.2020.107992>, letzter Zugriff 29.01.2021.

- Nunes, I. L. (Hrsg.): *Advances in Human Factors and Systems Interaction. Proceedings of the AHFE 2020 Virtual Conference on Human Factors and Systems Interaction, July 16-20, USA.* Heidelberg & Berlin: Springer 2020. <https://rd.springer.com/book/10.1007%2F978-3-030-51369-6>, letzter Zugriff 29.01.2021.
- Rohmert, W.: Das Belastungs-Beanspruchungs-Konzept. *Zeitschrift für Arbeitswissenschaft* 38 (1984), S. 193–200.
- Schneck, M.: Persönliche Mitteilung. Fellbach: Delmia Systemes, 11.07.2009.
- Spitzhirm, M.: Persönliche Mitteilung. Chemnitz: imk automotive, 13.12.2020.
- VBG 212: Berufsgenossenschaftliche Vorschriften, Unfallverhütungsvorschrift Lärm, Januar 1990.
- VDI 3633-6:2001-10: Digitale Fabrik - Ergonomische Abbildung des Menschen in der Digitalen Fabrik. Berlin: Beuth. <https://www.beuth.de/de/technische-regel/vdi-3633-blatt-6/44890290>, letzter Zugriff 21.04.2021.
- VDI 3742 Richtlinienreihe Blatt 1 bis 6, 1981-2 bis 1983-6: Emissionskennwerte technischer Schallquellen; Spanende Werkzeugmaschinen. Berlin: Beuth.
- VDI 3742-1:1981-02: Emissionskennwerte technischer Schallquellen; Spanende Werkzeugmaschinen; Drehmaschinen. Berlin: Beuth. <https://www.beuth.de/de/technische-regel/vdi-3742-blatt-1/547180>, letzter Zugriff 21.04.2021.
- VDI 3760:1992-02: Berechnung und Messung der Schallausbreitung in Arbeitsräumen. Berlin: Beuth. <https://www.google.com/search?client=firefox-b-d&q=Beuth+VDI+3742-1%3A1981-02>, letzter Zugriff 21.04.2021.
- VDI 4499-4:2015-03: Digitale Fabrik - Ergonomische Abbildung des Menschen in der Digitalen Fabrik. Berlin: Beuth. <https://www.beuth.de/de/technische-regel/vdi-4499-blatt-4/222813009>, letzter Zugriff 21.04.2021.
- VDI – Verein Deutscher Ingenieure (2020): VDI 4499-5 - Projekt. Digitale Fabrik: Prognose von Umgebungseinflüssen auf den arbeitenden Menschen. Berlin: Beuth. <https://www.vdi.de/richtlinien/details/vdi-4499-blatt-5-digitale-fabrik-prognose-von-umgebungseinflussen-auf-den-arbeitenden-menschen>, letzter Zugriff 26.01.2021.
- Zülch, G.: Ergonomische Abbildung des Menschen in der Digitalen Fabrik – Die neue VDI-Richtlinie 4499-4. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): *Simulation in Produktion und Logistik 2013*. Paderborn: HNI-Verlagsschriftenreihe 2013 S. 53-60 (ASIM-Mitteilung Nr. 147).
- Zülch, G.: Entwicklung und Stand der Personaleinsatzsimulation und deren Einbindung in die Digitale Fabrik. In: Putz, M.; Schlegel, A. (Hrsg.): *Simulation in Produktion und Logistik 2019*. Auerbach: Verlag Wissenschaftliche Scripten 2019, S. 183-192 (ASIM-Mitteilung Nr. 172).
- Zülch, G.: Development and Status of Personnel Deployment Simulation and its Inclusion into the Digital Factory. *SNE Simulation Notes Europe* 30 (2020) 4, S. 197-202.
- Zülch, M.; Zülch, G.: Production logistics and ergonomic evaluation of U-shaped assembly systems. *International Journal of Production Economics* 190 (2017), S. 37-44. <http://dx.doi.org/10.1016/j.ijpe.2017.01.004>, letzter Zugriff 05.07.2017.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Digital-physische Verbundkonzepte: Gegenüberstellung, Nutzeffekte und kritische Hürden

Concepts for Digital-Physical Connection: Comparison, Benefits and Critical Issues

Richard Scheer, Steffen Straßburger, TU Ilmenau, Ilmenau (Germany),
richard.scheer@de.bosch.com, steffen.strassburger@tu-ilmenau.de

Marc Knapp, Robert Bosch GmbH, Stuttgart (Germany),
marc.knapp@de.bosch.com

Abstract: Several concepts for digital-physical connection exist in literature and practice. This paper provides an overview over prevalent concepts. It characterises their specific attributes and places them in contrast with each other. Furthermore, it describes the major benefits as well as the most critical issues in the implementation of these concepts. These potential benefits and issues might then also serve as indicators for further research. From a practical perspective, this paper introduces a straightforward procedure to indicate the appropriate and most efficient concept for any specific implementation of a digital-physical connection system. It bases this indication on the specific requirements of the application.

1 Einführung

Im Rahmen industrieller Entwicklungsstrategien, wie z.B. Industrie 4.0, wird wiederholt die Notwendigkeit von Verbundsystemen zwischen physischen Komponenten der Produktion und computergestützter Simulation betont. Es mangelt jedoch an klarer Differenzierung zwischen den verschiedenen digital-physischen Verbundkonzepten. Prävalente Konzepte wie Online-Simulation, Symbiotische Simulation, Digitale Zwillinge, Cyberphysische Systeme (CPS) oder Dynamic Data Driven Application Systems (DDDAS) müssen zunächst gegenübergestellt und voneinander abgegrenzt werden, zumal dies in der Literatur bisher nicht genügend erfolgt ist.

Im Kontext von Produktion und Logistik verspricht ein digital-physischer Verbund stellenweise signifikante Vorzüge gegenüber einem klassischen System. Gleichzeitig erfordern jene Verbundsysteme enorme Konzipierungs-, Implementierungs- sowie Betriebsaufwendungen. Folglich ist es für den effizienten Einsatz von digital-physischen Verbundsystemen notwendig, Aufwand und Nutzen apriorisch

abzuwägen und die tauglichste Methodik zu identifizieren. Eine Erfassung der wichtigsten Nutzeffekte und Aufwände soll im Laufe dieser Abhandlung angerissen werden, um eine strukturiertere Vorgehensweise jener Effizienzbetrachtung zu unterstützen. Es wird jedoch keine tiefgehende Analyse einzelner Aufwendungen und einzelner Nutzeffekte von digital-physischen Verbundsystemen erfolgen. Vielmehr sollen diese Aspekte aufgezeigt werden, um potenzielle Nutzer zu sensibilisieren und weitergehende Forschungspotenziale aufzuzeigen. Aus einer praktischen Perspektive soll dieser Beitrag helfen, ein geeignetes Verbundsystem für einen Anwendungsfall auszuwählen.

2 Begriffliche Klärung und Abgrenzung

Bevor eine Analyse der Nutzeffekte und Implementierungshürden einzelner Verbundkonzepte durchgeführt werden kann, müssen diese zunächst einzeln definiert und voneinander abgegrenzt werden. Die Konzepte der deskriptiven Fertigungsdatenanalyse (auch Data Analytics oder Data Mining im Produktionskontext) und der Offline-Simulation (Ablaufsimulation bzw. Materialflusssimulation) sollen zwar Teil der Abgrenzung sein, werden jedoch an dieser Stelle nicht näher erläutert, da sie keine Verbundkonzepte sind. Für nähere Information kann man Meister et al. 2019 respektive Gutenschwager et al. 2017 konsultieren.

2.1 Online-Simulation

Die erste signifikante Untersuchung des Konzeptes Online-Simulation erfolgte durch Davis (1998), in welcher er den Grundgedanken äußerte, dass ein Simulationsmodell in Echtzeit die gleichen Umwelteinflüsse erfahren müsse wie sein physisches Gegenstück. Laut Hanisch et al. (2003) sowie Cardin und Castagna (2011) müsse in jedem Fall eine Initialisierung des Simulationsmodells anhand des Zustands des Realsystems durchgeführt werden, da innerhalb dieser Methodik der Anfangszustand der Simulation besonders bedeutend sei. Ebenso müssen Simulationsergebnisse ausreichend schnell zur Verfügung gestellt werden, um nützlich zu sein (Hanisch et al. 2003). Obwohl ein allgemeines Verständnis des Konzeptes besteht, existiert keine allgemeingültige Begriffsdefinition (Weigert et al. 2017).

2.2 Dynamic Data Driven Application System

Laut Darema (2005) handle es sich bei einem DDDAS um die Integration von historisch oder gegenwärtig aufgenommenen Daten in eine ausführbare Anwendung, wobei jene Anwendung den Aufnahme- bzw. Messprozess der Daten dynamisch steuern könne. Alle Stellgrößen, die ein DDDAS beeinflussen kann, sind rein virtuell, d. h. das physische System wird ausdrücklich nicht beeinflusst (Blasch et al. 2018). Schlussendlich handelt es sich bei einem DDDAS um einen Regelkreis zwischen Simulation und Messung bzw. Datenaufnahme (Darema 2005). Im Rahmen eines DDDAS lassen sich Open- und Closed-Loop-Strategien differenzieren. Ein Open-Loop DDDAS verwendet Realdaten, um den Simulationsprozess zu beeinflussen, wobei bei einem Closed-Loop DDDAS zusätzlich die Simulation den Aufnahmeprozess der Realdaten beeinflusst (Knight et al. 2007).

2.3 Symbiotische Simulation

Im Rahmen des Dagstuhl-Seminars im Jahr 2002 wurden Symbiotische Simulationen als eine der größten Zukunftsherausforderungen im Themengebiet von Modellierung und Simulation identifiziert (Fujimoto et al. 2002). Symbiotische Simulation besitzt stellenweise große Ähnlichkeit zur Online-Simulation und dem DDDAS (Aydt et al. 2009). Ein symbiotisches Simulationssystem besteht aus physischer und simulierter Komponente, wobei die Simulation anhand von realen Sensordaten initialisiert wird und das physische System einen Nutzen aus den Simulationsergebnissen bezieht (Aydt et al. 2008). Ähnlich wie beim DDDAS wird dabei zwischen Open-Loop und Closed-Loop unterschieden. Während in einem Open-Loop System nur die simulative Komponente Informationen aus dem physischen System bezieht, so findet in einem Closed-Loop System ein bidirektionaler Informationsaustausch statt (Aydt et al. 2008). Ein wichtiges Instrument, um in diesem Kontext alternative Handlungsmöglichkeiten zu bewerten, ist die What-If-Analyse (Aydt et al. 2009).

2.4 Digitale Zwillinge

Die erste bedeutende Nennung des Begriffs erfolgte durch Shafto et al. im Jahr 2012 im Kontext der Simulation von Luft- und Raumfahrtvehikeln und deren Subsystemen. Seit einiger Zeit wird der Digitale Zwilling als ein wichtiges Konzept des Industrie 4.0 Paradigmas verwendet. Kuehner et al. (2021) führten jüngst einen Meta-Review (ein Review von Reviews) zum Thema durch und konnten dabei feststellen, dass zu Fragen von einer Definition, dem wahrgenommenen Entwicklungsstand, den potenziellen Anwendungen sowie den wichtigsten Forschungsfragen wenig wissenschaftliche Einigkeit vorliegt. Anhand wiederkehrender Aspekte definierten sie den Digitalen Zwilling anhand von vier Punkten:

- Digitale Repräsentation des physischen Gegenstücks
- Dient als Simulationsbasis oder ist selbst ein Simulationsmodell
- Hat eine bidirektionale (Daten-)Verbindung mit dem physischen Gegenstück
- Bidirektionale Verbindung kann über mehrere Systemlebensphasen bestehen

2.5 Cyberphysische Systeme

Im Jahr 2006 wurde der Begriff des CPS durch die National Science Foundation der USA verwendet (Lee 2015). Laut Lee (2015) handle es sich bei einem CPS um eine Erweiterung von einem physischen System um eingebettete Computer, welche physische Prozesse überwachen und steuern könnten (hauptsächlich in Form von Regelkreisen). Gleichzeitig sei auch das physische System in der Lage, die Berechnung der Computer zu beeinflussen. Letztlich wird damit ein physisches System um eine integrierte Cyber-Komponente erweitert. Diese Cyber-Komponente, in der Regel ein Modell, kann in Komplexität dabei von einer verhaltensbeschreibenden Differentialgleichung bis zu einem komplexen Simulationsmodell rangieren (Lee 2015). Bei der Ausweitung des CPS-Konzeptes auf ein ganzes Produktionssystem spricht man von einem Cyberphysischen Produktionssystem. Laut Monostori (2014) bestehe ein solches System aus autonomen, kooperativen Elementen, die über verschiedene Ebenen (z.B. Maschine, Produktion, Logistiknetzwerk) miteinander verbunden sind.

2.6 Abgrenzung

Die fünf genannten Konzepte voneinander abzugrenzen wird dadurch erschwert, dass nicht für alle dieser Konzepte eine allgemein akzeptierte Definition existiert. Dies führt zu mangelnder Trennschärfe und führt u.U. zu begrifflicher Verwirrung. Zudem kann es vorkommen, dass ein Verbundsystem die Definition von mehr als einem dieser Konzepte erfüllt. So könnte eine Symbiotische Simulation bspw. auch eine Online-Simulation oder DDDAS sein (Aydt et al. 2009). Demnach kann an dieser Stelle nur eine Abgrenzung anhand der hier selektierten Definitionen und Literatur erfolgen.

Zunächst haben alle Verbundkonzepte eine physische sowie eine digitale Komponente, die bspw. durch Simulation zum Gesamtsystem beiträgt. Zudem existiert in allen Konzepten zumindest eine monodirektionale Verbindung vom physischen zum digitalen System. In Tabelle 1 sind wesentliche Unterscheidungsmerkmale zwischen den fünf Verbundsystemen sowie Offline-Simulation und der Fertigungsdatenanalyse aufgeführt. Die hauptsächlichen Unterscheidungsmerkmale der digital-physischen Verbundsysteme untereinander sind die Art der digital-physischen Kopplung, sowie der Integrationsgrad der Recheneinheit. Um Übersichtlichkeit zu gewährleisten, wurde bewusst auf den Vergleich weiterer Verbundsystem-spezifischer Nuancen verzichtet.

Tabelle 1: Unterschiede zwischen prävalenten digital-physischen Verbundkonzepten

	P → D Verbindung	D → P Verbindung	Simulation	Recheneinheit integriert
Fertigungsdatenanalyse	+	-	-	-
Offline-Simulation	-	-	+	-
Online-Simulation	+	-	+	-
DDDAS	+	-(Messung)	+	-
Symbiotische Simulation	+	+/-	+	-
Digitaler Zwilling	+	+	+	-
CPS	+	+	+	+

P – Physisch, D - Digital

3 Nutzeffekte digital-physischer Verbundsysteme

Der grundsätzliche Vorteil von digital-physischen Verbundsystemen ist die Verfügbarkeit von Realdaten im Gegensatz zu den lediglich hinreichend repräsentativen Annahmen oder Zufallsverteilungen in Offline-Systemen. Im Kontext der Simulation bedeutet dies eine Reduktion des Simulationsfehlers (Weigert et al. 2017), wie Abbildung 1 veranschaulicht. Es ist zu sehen, dass wiederholte Initialisierung den Simulationsfehler reduziert. Gleichzeitig ist zu erkennen, dass Simulationen in Verbundsystemen nur dann Vorteile gegenüber klassischen Simulationen erbringen, wenn der Zeithorizont klein ist (d. h. operativ, ggf. taktisch), da diese auf erneute Initialisierungen angewiesen sind. Damit ist über strategische

Zeiträume der Gütegewinn einer Online-Simulation gegenüber einer Offline-Simulation relativ gering.

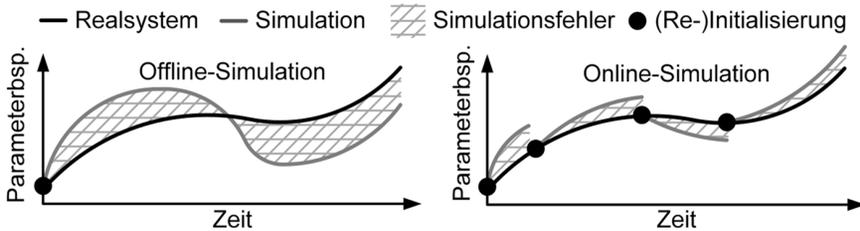


Abbildung 1: Vergleich Offline-/Onlinesimulation, angelehnt an Weigert et al. 2017

Angesichts der großen Heterogenität der Anwendungsfelder von Simulation im Bereich von Produktion und Logistik ist eine Auflistung aller Nutzeffekte von Verbundsystemen unrealistisch, da spezifische Anwendungsfälle äußerst zahlreich sind. Jedoch kann die allgemeine Nützlichkeit bzw. Funktionalität eines Verbundkonzeptes eingeschätzt werden, indem man die Verantwortungsteilung zwischen Mensch und Verbundsystem als Indikator verwendet.

Der Funktionsumfang eines spezifischen Verbundsystems soll an den gelisteten vier Hürden gemessen werden. Die Definition dieser Hürden ist durch die Reifegradabstufung aus der Fertigungsdatenanalyse (Meister et al. 2019) sowie durch die Klassifizierung von Symbiotischen Simulationen in Open- und Closed-Loop Varianten (Aydt et al. 2008) inspiriert.

- System zur Deskription
- System zur Prognose
- System zur indirekten Präskription
- System zur direkten Präskription

Tabelle 2: Vergleich Funktionsumfänge von betrachteten Konzepten

	Deskription	Prognose	Indirekte Präskription	Direkte Präskription
Fertigungsdatenanalyse	+	-	-	-
Offline-Simulation	-	+	-	-
Online-Simulation	+	+	+	-
DDDAS	+	+	+	- (Messung)
Symbiotische Simulation	+	+	+/-	+/-
Digitaler Zwilling	+	+	+	+
CPS	+	+	+	+

Tabelle 2 zeigt eine Übersicht des Funktionsumfangs der hier behandelten Verbundkonzepte.

Ein System zur Deskription hat einen informativen und ggf. diagnostischen Charakter und bezieht sich auf den gegenwärtigen Zustand, d. h., alle Interpretationen und Aktionen auf Basis dieser Interpretationen obliegen dem Nutzer.

Ein System zur Prognose stellt einen meist probabilistischen Blick auf den zukünftigen Systemzustand zur Verfügung. Alle Aktionen zur Sicherstellung oder Vermeidung dieser Zukunft obliegen dem Nutzer.

Indirekte Präskription erfolgt dann, wenn ein Verbundsystem anhand von gegenwärtigen Informationen Handlungsoptionen zur Verfügung stellt, die gewünschte Zustände hervorbringen können. Eine Plausibilisierung, Auswahl und Umsetzung der Handlungsoptionen erfolgen durch den Nutzer.

Ein System zur direkten Präskription evaluiert ein System anhand von gegenwärtigen sowie prognostizierten Daten und erstellt Strategien sowie Handlungsmaßnahmen, um einen Zielzustand zu erreichen oder zu halten. Die Auswahl und Durchführung dieser Maßnahmen obliegen ebenso dem Verbundsystem. Der Nutzer hat im operativen Betrieb keine aktive Rolle.

4 Kritische Implementierungshürden

Die Implementierungshürden von digital-physischen Verbundsystemen sind abhängig von deren Funktionen. Dementsprechend sollen diese anhand der in Kapitel 3 beschriebenen Funktionsumfänge strukturiert werden.

Ein System, welches der Deskription dient, benötigt eine Infrastruktur, um Daten aus dem physischen System zu beziehen. Kuehner et al. (2021) stellten bspw. fest, dass die Erzeugung und der Betrieb einer robusten Dateninfrastruktur eine der größten Herausforderungen im Kontext von Digitalen Zwillingen sind. Daten könnten z.B. periodisch aus Datenbanken bezogen werden (Aydt et al. 2009), sollten jedoch echtzeitnah verbunden sein, damit z.B. Simulationen mit dem aktuellen Zustand des physischen Systems initialisiert werden können (Hanisch et al. 2003). In einem konkreten Anwendungsfall kann die Implementierung einer solchen Infrastruktur sehr aufwändig sein. Moderne Fertigungs- und Montageanlagen bieten jedoch häufig Schnittstellen zur Datenaufnahme und -übermittlung, die diesen Prozess unterstützen können.

Ein weiterer kritischer Punkt ist der Datenschutz und die Datensicherheit, wie Barricelli et al. (2019) bspw. im Bezug auf Digitale Zwillinge feststellen. Die Aufnahme und Verarbeitung von personenbezogenen Daten sind besonders prekär und in den meisten Situationen unerwünscht. In breitflächiger, ungerichteter Datenaufnahme durch verschiedene Sensoren ist eine Herstellung von einem Personenbezug häufig möglich und muss daher behutsam verarbeitet werden. Des Weiteren stellen diverse Komponenten von digital-physischen Verbundsystemen potenzielle Ziele für Cyberattacken dar (Singh und Jain 2018) und müssen daher aktiv gesichert werden.

Ein System, das zur Prognose dienen soll, muss einen Mechanismus beinhalten, welcher Aussagen über die Zukunft generieren kann. In der Regel handelt es sich dabei um ein Simulationsmodell. Die Erstellung von geeigneten Modellen ist häufig

zeitaufwändig und erfordert entsprechende Kompetenzen. So identifizierten bspw. Kuehner et al. (2021) Aufwände für Modellierung und Simulation als eine der größten Hürden bei der Implementierung von Digitalen Zwillingen. Ein besonderes Augenmerk muss auf die Fähigkeit des Modells gelegt werden, mit dem physischen System synchronisiert werden zu können. Es ist zudem zu beachten, dass hochfrequente Synchronisierung evtl. einen erhöhten Rechenaufwand bedeuten kann (Modoni et al. 2018).

Ein präskriptives System erfordert Algorithmen bzw. Heuristiken, die aus dem gegenwärtigen und aus dem prognostizierten zukünftigen Zustand Handlungsstrategien entwickeln können. Diese Handlungsalternativen stellen sicher, dass gewünschte Verhaltensmuster eintreten oder vermieden werden. Beispielsweise kann eine solche Generierung und Bewertung von Handlungsalternativen über What-If-Analysen (Aydt et al. 2009) realisiert werden. In indirekt präskriptiven Systemen müssen die Handlungsalternativen dem Nutzer in geeigneter Form zur Verfügung gestellt werden.

In direkt präskriptiven Systemen erfolgen die Validierung und Umsetzung der Handlungsalternativen vollautomatisiert. Jedoch trägt eine solche Konfiguration die gleichen Risiken, die alle autonomen Systeme wie z.B. autonomes Fahren mit sich ziehen. Dies umfasst rechtlich-regulatorische Aspekte wie z.B. strafrechtliche Verantwortung im Schadensfall (Beck 2020) oder ethisch-moralische Aspekte, wie sie im wissenschaftlichen Feld der Maschinenethik behandelt werden (Winfield et al. 2019).

5 Auswahl eines geeigneten Verbundkonzeptes

Eine Effizienzbetrachtung kann an dieser Stelle nur qualitativ erfolgen. In diesem Kontext muss ein Verbundsystem so gewählt werden, dass der minimale Aufwand für die gewünschte Funktionalität anfällt. Ohne näher auf konzeptspezifische Nuancen einzugehen, die potenziell im einzelnen Anwendungsfall interessant sind, kann eine grundsätzliche Strategie zunächst nur durch Vermeidung unnötiger Aufwände strukturiert werden. Abbildung 2 zeigt eine solche Strategie. Dabei stellt sich ein Nutzer Fragen zum beabsichtigten Nutzen bzw. den notwendigen Fähigkeiten des Verbundsystems und wird daran in der Auswahl eines geeigneten Konzeptes unterstützt. In der Abbildung wird eine positive Antwort auf eine Frage durch ein Häkchen gekennzeichnet, während eine negative Antwort durch ein Kreuzchen symbolisiert wird. Möchte man bspw. ein System entwickeln, welches Realdaten extrahiert und auf deren Basis ein Simulationsmodell betreibt, so würde man die ersten zwei Fragen positiv und die dritte Frage negativ beantworten. Somit wäre das effizienteste System in diesem Fall eine Online-Simulation.

Diese Strategie ersetzt nicht die anschließende detaillierte Auseinandersetzung mit dem ausgewählten oder ähnlichen Verbundkonzepten. Vielmehr dient sie der inhaltlichen Initialorientierung und der ersten Eingrenzung von konzeptionellen Optionen.

Es ist zudem zu beachten, dass in der Kontrastierung der Konzepte das zu verbindende physische System in Dimension, Detailgrad und zu einem gewissen Grad auch Funktionalität vergleichbar sein muss. Beispielsweise ist der

Implementierungsaufwand eines Maschinen-CPS nicht mit dem eines Digitalen Zwillings einer ganzen Produktionslinie vergleichbar.

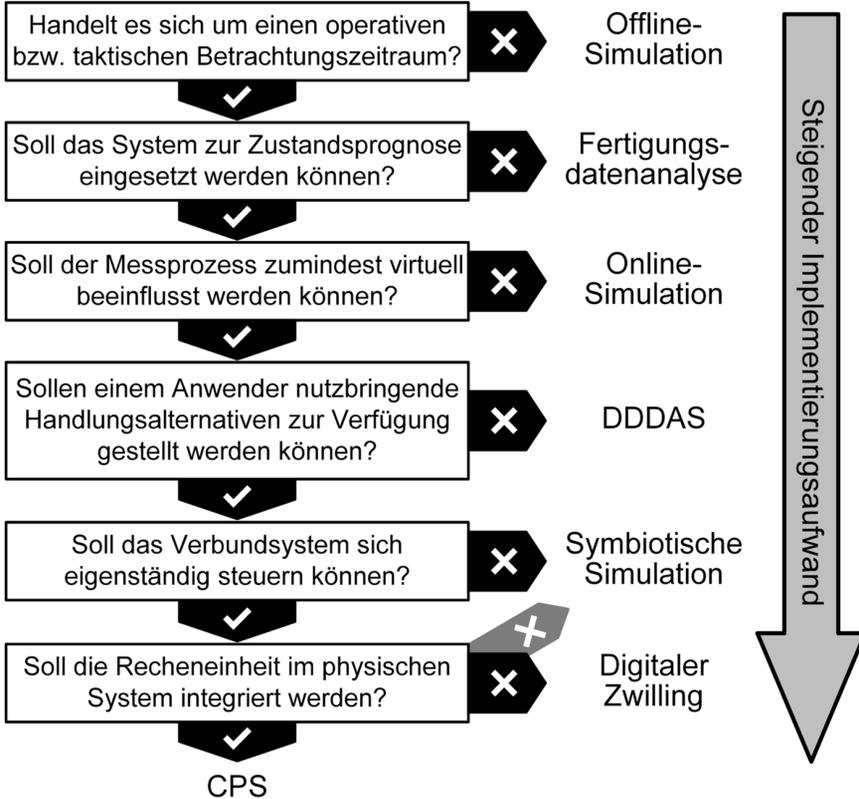


Abbildung 2: Strategie zur effizienten Verbundkonzeptauswahl

6 Fazit

Dieser Beitrag identifiziert prävalente digital-physische Verbundkonzepte, präsentiert deren Kernaspekte und grenzt diese voneinander ab. Zudem erfolgt eine Kategorisierung der Verbundkonzepte anhand ihrer Funktionsumfänge, um anschließend eine generelle Aussage über deren wahrscheinliche Implementierungshürden und damit -kosten zu treffen. Schließlich wird eine übersichtliche Vorgehensweise zur Auswahl eines Verbundsystems präsentiert. Die Selektion erfolgt dabei anhand des benötigten Funktionsumfangs und kann so den Implementierungsaufwand im Anwendungsfall minimieren.

Dieser Beitrag bewertet weder die Validität noch die Reife der einzelnen Konzepte. Im Rahmen der Übersichtsgenerierung kann auch nicht in eine signifikante Detailtiefe vorgedrungen werden, sodass evtl. konzeptspezifische Nuancen nicht betrachtet werden. Zudem reicht die hier präsentierte Vorgehensweise nicht als alleinige

Vorbereitung für eine Implementierung aus, sondern muss um eine konzeptspezifische Recherche ergänzt werden. Schlussendlich eignet sich die entwickelte Vorgehensweise somit hauptsächlich als erster Aufsatzzpunkt.

Zukünftig ist es notwendig, die hier aufgezeigten Verbundkonzepte stärker in den Fokus zu stellen, um eine Trennschärfe zwischen den Konzepten herstellen zu können. Dies wirkt auch etwaigen Falschklassifizierung entgegen. Des Weiteren müssen die Punkte, die in Kapitel 3 und 4 aufgeführt wurden, näher untersucht werden. Zum einen müssen die Nutzeffekte von digital-physischen Verbundsystemen stärker ausgenutzt werden und zum anderen müssen deren Implementierungsbarrieren abgebaut werden.

Literatur

- Aydt, H.; Turner, S.; Cai, W.; Low, M.Y.H.: Symbiotic Simulation Systems: An Extended Definition Motivated by Symbiosis in Biology. In: 22nd International Workshop on Principles of Advanced and Distributed Simulation, Rom (Italien), 3.-6. Juni 2008, S. 109-116.
- Aydt, H.; Turner, S.; Cai, W.; Low, M.Y.H.: Research issues in symbiotic simulation. In: Rosetti, M.D.; Hill, R.R.; Johansson, B.; Dunkin, A.; Ingalls, R.G. (Hrsg.): Proceedings of the 2009 Winter Simulation Conference, Austin (USA), 13.-16. Dezember 2009, S. 1213-1222.
- Barricelli, B.R.; Casiraghi, E.; Fogli, D.: A Survey on Digital Twin: Definitions, Characteristics, Applications and Design Implications. IEEE Access 7 (2019), S. 167653-167671.
- Beck, S.: Die Diffusion strafrechtlicher Verantwortlichkeit durch Digitalisierung und Lernende Systeme. Zeitschrift für Internationale Strafrechtsdogmatik 2 (2020), S. 41-50.
- Blasch, E.; Bernstein, D.; Rangaswamy, M.: Introduction to Dynamic Data Driven Application Systems. In: Blasch, E.; Ravela, S.; Aved, A. (Hrsg.): Handbook of Dynamic Data Driven Application Systems. Cham (Schweiz): Springer 2018, S. 1-25.
- Cardin, O.; Castagna, P.: Proactive production activity control by online simulation. International Journal of Simulation and Process Modelling 6 (2011) 3, S. 177-186.
- Darema, F.: Grid Computing and Beyond: The Context of Dynamic Data Driven Applications Systems. Proceedings of the IEEE 93 (2005) 3, S. 692-697.
- Davis, W.J.: On-Line Simulation: Need and Evolving Research Requirements. In: Banks, J. (Hrsg.): Handbook of Simulation: Principles, methodology, applications and practice. New York (USA): John Wiley & Sons 1998, S. 465-516.
- Fujimoto, R.; Lunceford, D.; Page, E.; Uhrmacher, A.M. (Hrsg.): Grand Challenges for Modelling and Simulation – Dagstuhl Seminar Report 350. Dagstuhl, 26.-30. August 2002.
- Gutenschwager, K.; Rabe, M.; Spieckermann, S.; Wenzel, S.: Simulation in Produktion und Logistik – Grundlagen und Anwendungen. Berlin, Heidelberg: Springer 2017.
- Hanisch, A.; Tolujew, J.; Raape, U.; Schulze, T.: Online-Simulation für Personenströme in einem Frühwarnsystem. In: Hohmann, R. (Hrsg.): Tagungsband ASIM 2003 – 17. Symposium Simulationstechnik, Magdeburg, 16.-19. September 2003, S. 221-226.

- Knight, D.; Ma, Q.; Rossmann, T.; Jaluria, Y.: Evaluation of Fluid-Thermal Systems by Dynamic Data Driven Application Systems – Part II. In: Shi, Y.; van Albada, G.D.; Dongarra, J.; Sloot, P.M.A. (Hrsg.): Proceedings of the 2007 International Conference on Computational Science, Beijing (China), 27.-30. Mai 2007, S. 1189-1196.
- Kuehner, J.; Scheer, R.; Strassburger, S.: Digital Twin: Finding Common Ground – A Meta-Review. Angenommen zur Veröffentlichung in Proceedings of the 54th CIRP Conference on Manufacturing Systems, 2021.
- Lee, E. A.: The past, present and future of cyber-physical systems: A focus on models. *Sensors* 15 (2015) 3, S. 4837-4869.
- Meister, M.; Beßle, J.; Cviko, A.; Böing, T.; Metternich, J.: Manufacturing Analytics for problem-solving processes in production. In: Butala, P.; Govekar, E.; Vrabec, R. (Hrsg.): *Procedia CIRP* 81 – Proceedings of the 52nd Conference on Manufacturing Systems, Ljubljana (Slowenien), 12.-14. Juni 2019, S. 1-6.
- Modoni, G. E.; Caldarola, E. G.; Sacco, M.; Terkaj, W.: Synchronizing physical and digital factory: benefits and technical challenges. In: Teti, R. (Hrsg.): *Procedia CIRP* 79 – Proceedings of the 12th CIRP Conference on Intelligent Computing in Manufacturing Engineering, Golf von Neapel (Italien), 18.-20. Juli 2018, S. 472-477.
- Monostori, L.: Cyber-physical production systems: Roots, expectations and R&D challenges. In: ElMaraghy, H. (Hrsg.): *Procedia CIRP* 17 - Proceedings of the 47th CIRP Conference on Manufacturing Systems, Windsor (Kanada), 28.-30. April 2014, S. 9-13.
- Shafto, M.; Conroy, M.; Doyle, R.; Glaessgen, E.; Kemp, C.; LeMoigne, J.; Wang, L.: Modeling, simulation, information technology & processing roadmap. Washington DC (USA): National Aeronautics and Space Administration, 2012.
- Singh, A.; Jain, A.: Study of Cyber Attacks on Cyber-Physical System. In: Proceedings of 3rd International Conference on Internet of Things and Connected Technologies, Jaipur (Indien), 26.-27. März 2018, S. 686-690.
- Weigert, D.; Reggelin, T.; Tolujew, J.: Material Flow Simulation of Logistics Processes – An Approach of Online Analysis, Planning and Control of Logistics Processes of Supply Chains. In: Proceedings of MultiScience – XXXI. MicroCAD International Scientific Conference, Miskolc (Ungarn), 20.-21. April 2017, (ohne Seitenzahlen).
- Winfield, A.F.; Michael, K.; Pitt, J.; Evers, V.: Machine Ethics: The Design and Governance of Ethical AI and Autonomous Systems. *Proceedings of the IEEE* 107 (2019) 3, S. 509-517.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Simulation Based Assessment of Lean and Industry 4.0 Measures in Changeable Production Systems

Simulationsbasierte Bewertung von Lean und Industrie 4.0 Maßnahmen in wandlungsfähigen Produktionssystemen

Niels L. Martin, Antal Dér, Adrian Langer, Nadja Henningsen, Christian Ortmeier, Tim Abraham, Christoph Herrmann, TU Braunschweig, Braunschweig (Germany),
n.martin@tu-bs.de, a.der@tu-bs.de, adrian.langer@tu-bs.de, n.henningsen@tu-bs.de, c.ortmeier@tu-bs.de, t.abraham@tu-bs.de, c.herrmann@tu-bs.de

Abstract: Increasing demand for individualized products as well as the need for flexible and variable manufacturing leads to an increase in the complexity of modern production systems (PS). This complexity requires the implementation of digitalised lean as well as Industry 4.0 (I4.0) measures. Even though these measures might enable the handling of this complexity in the operation phase, their planning and implementation is complex. This paper presents a simulation framework that targets the planning and assessment of digitalised lean and I4.0 measures in a systemic context. The framework consists of the main PS components and additional I4.0 elements. These have been modelled using an agent-based simulation approach and is applicable to multiple manufacturing scenarios. An exemplary application to a changeable PS for individualized pharmaceutical products concludes the paper.

1 Introduction

Shorter product and process innovation cycles, a high demand for individualized products and the transformation to a sustainable economy create a turbulent market environment. Production planners and operators are challenged to cope with the resulting internal complexity and to meet customer demands with the implementation of changeable PSs. The inherent complexity of changeable PSs can be handled with the implementation of digitalised lean methods or I4.0 measures, so called cyber-physical production systems (CPPS) (Wagner et al., 2017). CPPS contain four elements: i) a physical world, to which the CPPS is referenced, ii) a data acquisition and treatment with technical components such as sensors and data storages, iii) a cyber world applying digital methods on the data allowing an iv) feedback to the physical world. Those systems are considered enablers for I4.0 (Veile et al., 2020; Inkermann et al., 2019). Even though I4.0 measures might enable the operation of complex PS,

their planning and implementation further increases the already existing complexity of the PS. Consequently, production planners, especially in small and medium sized enterprises, struggle with the implementation of I4.0 measures. The challenges can be traced back to several hurdles, such as the unknown impacts of single measures on PS level, combined with high investments and uncertain payback periods (Orzes et al., 2020; Veile et al., 2020; Wischmann et al., 2015). Therefore, a particular need for generic approaches that are able to plan and evaluate I4.0 applications in specific use cases exists. In this context, simulation approaches can provide an effective means for a planning support. While respecting process chain inherent dynamics, they can enable the prospective evaluation of I4.0 applications in complex production systems and their investigation in different scenarios (Martin et al., 2020). Against this background, this paper proposes a generic I4.0 simulation framework for planning and assessing digitalised lean and I4.0 measures (further only called I4.0 measures). The framework is applied in the context of individualized pharmaceutical production.

2 Digitalised Lean and Generic Simulation

2.1 The digitalisation of lean measures

Transparent and structured processes as well as the interrelation with relevant indicators are indispensable to simulate the impacts of the implementation of I4.0 measures. An important prerequisite for the targeted implementation of I4.0 measures is to view production from a lean perspective - as a holistic PS (Bick, 2014; Dombrowski and Richter, 2016). The fundamental principles of holistic PSs, such as standardisation and continuous improvement, still apply even with the introduction of I4.0. Furthermore, the principles can be used to enhance existing methods and tools through the functions of I4.0 (Deuse et al., 2020). In this way, I4.0 can complement existing lean approaches and raise new potentials.

Within the Verein Deutscher Ingenieure (VDI - The Association of German Engineers) guideline 2870, lean methods are already evaluated regarding the target indicators quality, cost and time (VDI-2870). Further and new developed lean methods considering I4.0 were also evaluated with regard to these indicators and described with method sheets ("GaProSys 4.0", 19804 N). Based on this guideline, a structured catalogue of I4.0 and lean measures was developed within the scope of this research work (Figure 1). As an example of a digitalised lean application, this work examines a digital assistant system (DAS), which supports assembly processes as a human-machine interface by providing assisting information for the employee. Through the integrated testing of individual process steps up to quality control, specific process and product characteristics can be depicted. DASs use advanced sensor technologies to allow imaging processes or pick-by-light for example.

The aforementioned target variables quality, cost and time were supplemented by sustainability to address the growing challenges of sustainable production. These are integrated as dimensions and divided into categories. Thus, the dimension quality is subdivided into the categories product and process quality. According to VDI guideline 2870, product quality is composed of the product characteristics lifetime, safety and functionality. Since these indicators are not directly related to the production logistic processes, they are not considered in this approach.

Dimension		Quality						Cost		Time			Sustainability										
Category		Product			Process					Production													
Lean principles		Lifetime	Safety	Functionality	Input errors	Measurement errors	Employee-rel. dev.	Machine-rel. dev.	Method-rel. dev.	Material-rel. dev.	Environment-rel. dev.	Fixed costs	Variable costs	Post-process waiting	Transport	Pre-process waiting	Set-up	Processing	Rework	Emission	Pollution	Resource consumption	
Visual management	DAS	o	o	o	+	o	++	o	o	o	o	o	o	o	o	o	+	++	+	o	o	o	
	Shop floor management																						
	Andon																						

Legend Indicator Simulation-relevant indicator

Influence rating:
 (-) High negative; (-) Negative; (o) No; (+) Positive; (++) High positive

Figure 1: Extract for a proposed structured catalogue of I4.0 and lean measures

The decisive influencing variables for process quality result from possible process discrepancies. Discrepancies are caused by input or measurement errors or due to machines, employees, material, environmental influences or methods (VDI-2870). For the dimension time, the throughput element according to Wiendahl was supplemented by the rework time (Wiendahl, 2010). The sustainability indicators for sustainable manufacturing according to Joung are used (Joung et al., 2013). Only indicators relevant to the simulation were considered. To enable decision support, a more granular evaluation of the indicators is required. The target variables summarised generically in VDI guideline 2870 must therefore be converted into simulation-specific key indicators.

2.2 Generic simulation approaches for the planning

With respect to the manifold performance indicators and the high planning complexity, Albrecht et al. present an integrated planning and optimisation methodology for designing the changeability of PSs (Albrecht et al., 2014). The approach combines system dynamics (SD) and discrete event simulation (DE) modelling for the evaluation of changeability measures on a production system. Aspects of integrating I4.0 measures are however beyond the scope of the approach. The dynamic relationships within changeable manufacturing systems can also be modelled by combining agent-based simulation (AB) and DE. In this way, product-specific flows as well as specific elements of PSs (e.g. storage, workstation) can be modelled. A generic approach allows adaptations of the model elements and enables the focus on specific key performance indicators (KPIs) (Schönemann et al., 2019).

By now, simulation models for changeable PS use sub-models to model the existing elements of this system. Those elements can be connected via material, energy and information flows (Florescu and Barabas, 2020; Filz et al., 2019; Kirchner and März, 2002). The used agent-based approach of Filz et al. allows the flexible rearrangement of the system. Each of the generic agents consist of parameters, a specific behaviour and interconnections to other agents (Filz et al., 2019). The change of parameters within the sub-models allows the investigation of different sub-model types, e.g. a transportation agent can be either a tigger train or an automated guided vehicle (Filz et al., 2019). This enables the investigation of several scenarios.

The planning and assessment of I4.0 measures is difficult due to the case specific complexity of PSs. None of the above models investigates the impact of I4.0 measures

in a systemic context and provides decision support prior to implementation based on quantifiable performance indicators. Hence, there is a strong research need for the planning and assessment of I4.0 measures in changeable production systems.

3 Generic I4.0 simulation framework

Core of the framework is an agent-based production system model that is supplemented with a generic module for I4.0 measures (Figure 2). The module contains a structured catalogue of I4.0 measures and characterizes them according to quality, cost, time and sustainability dimensions (see also Figure 1). Within the production system model, the I4.0 measures are evaluated in a systemic context. Therefore, potential cross-impacts between measures and inside the process chain (e.g. shifting of bottlenecks) are evaluated with regard to relevant KPIs. A generic model library contains PS elements covering a wide variety of real-life scenarios. These are used to flexibly set up production systems on a case specific basis. The input and output module provide an interface for parametrizing the models and investigating the assessment results. In the following, the production system model and the generic model library will be described with a special focus on the integration of I4.0 measures.

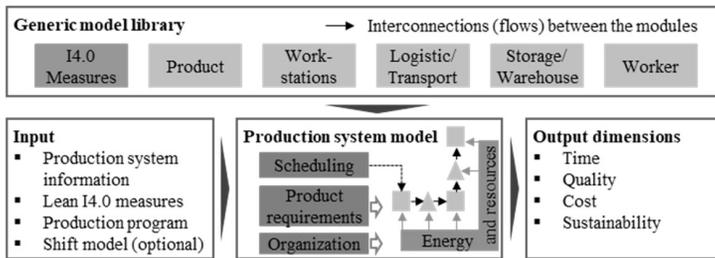


Figure 2: Overview of generic I4.0 simulation framework

3.1 Production system model

This module coordinates the material flow inside the production system and provides functionality to calculate KPIs, e.g. lead times or utilization. The module is a generic process chain modelling environment. Therefore, it can be used to model any process chain in discrete manufacturing according to the user’s input. The underlying logic applies agent-based modelling to allow for a multidirectional interaction between the generic production system elements from the model library. Consequently, the system performance evolves in a bottom-up manner. For example, product agents store their individual processing sequence based on the incoming orders. During simulation, material supply agents and workstations respond to the specific processing sequence and execute the transformation process.

3.2 Generic model library

The model library consists of generic agent types, which represent fundamental elements of a PS. As modelling of generic elements of a PS has already been discussed

in literature, a reference should be made to models that served as inspiration for the model library: logistic/ transport (Filz et al., 2019), product (Schönemann et al., 2019), workstations (manual and automatic) (Langer et al., 2021) and human worker (Halubeck and Herrmann, 2011; Schönemann et al., 2019). The agents have interfaces to interconnect different flows. Material flows determine the processing route of the (individualised) products through a PS. Energy flows allow the forecast of energy demands and are mainly modelled as inputs into the system boundaries. The data flows are mainly used for the interconnection between the agents. This is especially useful for the implementation of I4.0 measures. From a generic point of view, there are two different groups of I4.0 measures – process and process chain related measures. Both are depicted in Figure 3 within an exemplary process chain of two workstations and two warehouses connected via transport agents. The process related I4.0 measure is shown with its system boundary influencing only a single process step (here, the first workstation). The process chain related I4.0 measures either influence and/or receive data from several agents in the process chain.

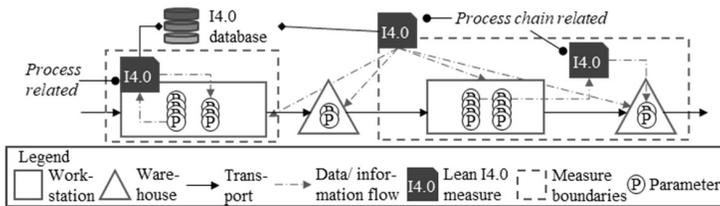


Figure 3: Schematic I4.0 measures implementation into the production system model

The behaviour of the agents is determined by the agents’ parameters as well as state charts. For a better description of how I4.0 measures interact with the generic elements, the agent storage supplemented with a I4.0 measure, the aforementioned digital assistant system (DAS), is illustrated in Figure 4.

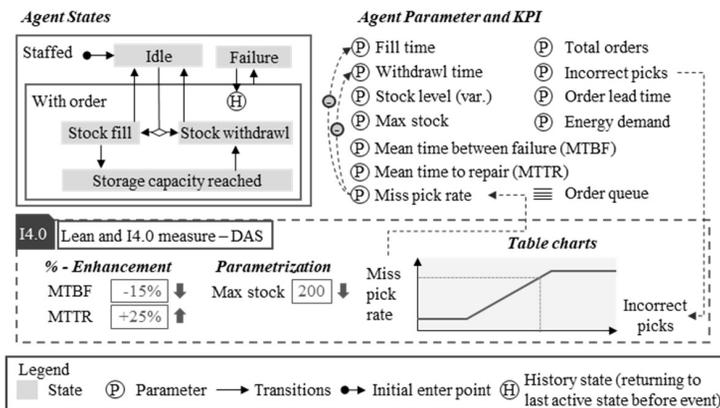


Figure 4: Exemplary illustration of storage agent with integrated I4.0 measure

This agent is able to fill and withdraw stock. The orders, which are interconnected with the product agents, reaching the storage agent are inserted into an order queue and consequently worked off. If the stock level reaches maximum stock, then the stock has to be withdrawn first before it can be filled again. The parameters on the right trace the process quality being able to calculate KPIs as the order lead time or picking accuracy. In order to allow for the investigation of process KPIs, the parameters total orders, incorrect picks and order lead time are raised continuously. These parameters are aggregated and can be extended, depending on the KPIs to be assessed (compare section 2).

As an I4.0 measure exists of the elements of a CPPS (compare section 1), the data acquisition of the physical world can be simulated by utilising the agent's parameters. The cyber world and the feedback element of a CPPS can be simulated by modelling the change of an agent's parameter according to the influence of the I4.0 measure. As illustrated in Figure 4, the influence of an I4.0 measure on the agent is modelled using table charts, a percentage enhancement or parametrisation.

The choice of modelling depends on the available data and relevance of the process step. The values can be taken from a database or individually parametrised on a theoretical or empirical basis. Since the specific impact of I4.0 measures is only available for a few applications on usually specific production systems, reference cases for similar I4.0 measures can also be used to allow a first assessment. When more I4.0 measures are implemented on an agent, the beneficial and disadvantageous impacts between the measures need to be addressed. For this, a cross-impact analysis is suggested (Martin et al., 2020). In the exemplary case in Figure 4 the integration of a pick-by-vision DAS is shown, but also further measures could be added. Since the integrated DAS requires further maintenance and will cause failures, the MTBF is decreased by 15 %, while the mean time to repair is increased by 25 %. Furthermore, the integration of the DAS will reduce the maximal stock capacity to 200 products. Additionally, depending on the already existing incorrect picks, the missed pick rate will be adjusted according to a given table function. The fill and withdrawal times are bounded to the missed pick rate and will be reduced according to their calculation within the simulation.

4 Framework application

The framework was applied on a changeable PS for the individualised primary and secondary packaging of pharmaceutical solid dosage forms, such as tablets or capsules. Figure 5 illustrates a changeable PS with one of the possible material flow routes in the system. Within the primary packaging, a foil is deep drawn into a blister, then filled with solid dosage forms and in the next step sealed. After printing individual information on each sealed blisters' cavity, the separation process perforates the blister and concludes the fully automated primary packaging process chain. In the secondary packaging step, the blister cavities as well as a pharmaceutical supplement are manually packed into a secondary package.

The technical system of the changeable PS consists of a software component, master and service modules. Those modules are connected by material, energy and information flows. The master modules use robots for the systems material flow, are used as small buffer and support the service modules within their process if necessary. The service modules offer a specific manufacturing service and are connected to the

master modules via multi-functional, fast-switch plugs allowing all necessary flows. Hence, the service modules can therefore be changed quickly, allowing to change the PS. The centralised software component is used as a broker between products and service modules and as an interface to the user of the system.

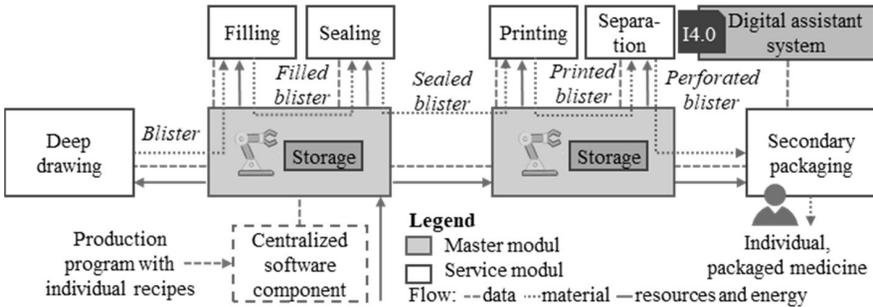


Figure 5: Overview of changeable production system with I4.0 implementation

The production program consists of highly individualised pharmaceutical products. Each order can have varying dosage forms, formats and secondary packaging as well as an individualised print on the blisters to name only a few. In this use case, the limiting factor for the overall performance of the system is the quality at the manual process step of secondary packaging. This has been previously assessed with an extended value stream methodology based on (Martin et al., 2020). The core question therefore is how to increase the quality rate by implementing an I4.0 measure and which other performance indicators are affected as well. The above described DAS shall be simulated within this changeable PS, forecasting quality KPIs.

Since rearrangements of flows are necessary, a flexible modelling approach is needed to adapt I4.0 measures. Model libraries of common simulation software are mostly not able to model the impact of I4.0 measures. The required elements for this framework are therefore modelled within the simulation software AnyLogic® from the AnyLogic Company. Allowing a decision support, two scenarios are simulated using the generic I4.0 simulation framework. The first scenario simulates the system without and the second scenario with the described DAS for the secondary packaging step. Within the secondary packaging process step, the DAS guides the user to correctly chose the packaged drugs within the perforated blisters. In addition to the correct selection of the blisters, the quality of the perforated blisters is controlled when reaching the service module. In case of detected quality problems at the secondary packaging, the products are newly produced. For the simulation, the different processes are modelled with a special focus on quality parameters. Since there is no data of the DAS for this specific use case available, empirical data was acquired and transferred from another assembly station from a learning factory with similar basic features and implemented DAS. The parameter changes for the scenario with implemented DAS are shown in Table 1. It can be seen that the (pre-/post-) processing times are significantly decreasing with the implementation of the DAS. The quality rate can be set from 90 % to 100 % and the DAS has a failure detection rate of 95 % instead of 80 % when only the worker checks the incoming quality.

Table 1: Parameter changes of the DAS for the workstation secondary packaging

Pre-processing time	Processing time	Post-processing time	Quality rate	Failure detection rate
- 80 %	- 25 %	- 100 %	100 %	95 %

In this use case, the simulation results regarding the quality and lead time of the products are of relevance for the planner’s decision. Therefore, Monte Carlo simulation (n=10) has been applied. The resulting quality indicators as well as the lead time is illustrated in Figure 6. In comparison to the reference scenario without the implementation of the DAS,

- the Not-right-first-time (NRFT) indicator has increased on average by 1400 ppm products (+15 %),
- the overall process quality has increased on average by 11 % points reaching 100% compared to 89 % for the reference scenario,
- and the orders’ average lead time is reduced by 53 sec (-2,5%).

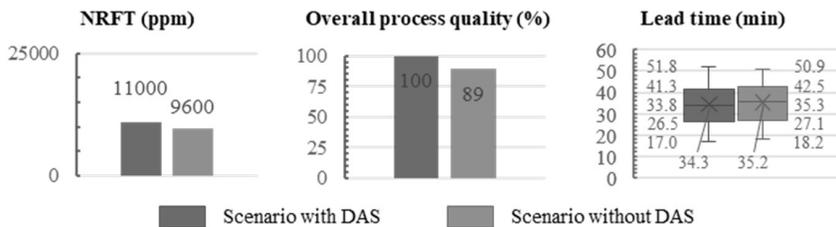


Figure 6: Simulation results comparing simulation scenarios with and without DAS

Since the detected failure probability with the DAS increases at the secondary packaging process step, the NRFT indicator increases as well. However, the final overall process quality increases in comparison to the scenario without the DAS. With increasing simulation runs, it is only stochastically justified that the overall process quality will decrease to 95 % as this is the probability for the failure detection with the implementation of a DAS. The lead time is mainly influenced by the (pre-/post-) processing times and the failure detection rate. The first is shorter for the DAS and the second results in longer lead times per product, due to the fact that the products have to be produced again. The shorter (pre-/post-) processing times have a higher influence on the lead time and therefore the lead time is reduced slightly.

It has been shown that with the implementation of the DAS in the changeable PS, the overall process quality and the lead time improves, while the NRFT has worsened. This information is particularly of interest for the production planner and helps to decide on the implementation of the DAS.

5 Conclusion and outlook

In order to cope with the increasing complexity of changeable PSs, production planners are faced with the challenge of successfully implementing lean and I4.0

measures. This paper presents a simulation-based framework for planning and assessing lean and I4.0 measures in changeable production systems. The framework evaluates I4.0 measures in a systemic context. Therefore, not only direct impacts but also potential cross-impacts between measures and inside the process chain (e.g. shifting of bottlenecks) are evaluated with regard to relevant performance indicators. Based on VDI 2870, simulation-relevant indicators were determined that address the different dimensions of quality, cost, time and sustainability. While the approach can be flexibly applied to evaluate KPIs from all dimensions, the case study primarily focused on the benefits of employing a digital assistant system in individualised pharmaceutical production on quality indicators. The results indicate that the implementation of the DAS not only increased the quality but also reduced the lead time of production, which together leads to a better customer satisfaction. Based on the above findings, further studies can be conducted to quantify the effects of I4.0 measures on other KPIs. Likewise, consideration of other I4.0 measures is of great importance in order to obtain an overall view, especially when combining different measures. Finally, the extension of the approach to include the parameters for determining product quality is useful in order to take into account other influences in addition to the process and production. Further research should be made on the impact assessment of individual I4.0 measures to allow building an I4.0 data base.

Acknowledgement

Funded by European Regional Development Fund (EFRE | ZW 6-85012454) within the research project “Synus” (Methods and Tools for the synergetic conception and evaluation of Industry 4.0-solutions). Furthermore, the authors acknowledge the personal effort of Mr. Bastian Thiede, who was strongly involved in the conceptualisation and installation of the changeable production system – the xCell.

References

- Albrecht, F.; Kleine, O.; Abele, E.: Planning and Optimization of Changeable Production Systems by Applying an Integrated System Dynamic and Discrete Event Simulation Approach. *Procedia CIRP* 17 (2014), pp. 386–391.
- Bick, W., 2014: Warum Industrie 4.0 und Lean zwingend zusammengehören. Hg. v. VDI-Z (11). Online verfügbar unter https://www.roi.de/fileadmin/user_upload/presse/2014_11_VDI-ZB880_ROI-Management.pdf, zuletzt geprüft am 03.05.2021.
- Deuse, J.; Dombrowski, U.; Nöhring, F.; Mazarov, J.; Dix, Y.: Systematic combination of Lean Management with digitalization to improve production systems on the example of Jidoka 4.0. *International Journal of Engineering Business Management* 12 (2020), 184797902095135.
- Dombrowski, U.; Richter, T.: Ganzheitliche Produktionssysteme und Industrie 4.0. *ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb* 111 (2016) 12, pp. 771–774.
- Filz, M.-A.; Gerberding, J.; Herrmann, C.; Thiede, S.: Analyzing different material supply strategies in matrix-structured manufacturing systems. *Procedia CIRP* 81 (2019), pp. 1004–1009.
- Florescu, A.; Barabas, S.A.: Modeling and Simulation of a Flexible Manufacturing System—A Basic Component of Industry 4.0. *Applied Sciences* 10 (2020) 22, pp. 8300.

- Halubeck, P.; Herrmann, C.: Design of Mixed Model Assembly Lines–Simulation based Planning Support 2011.
- Inkermann, D.; Schneider, D.; Martin, N.L.; Lembeck, H.; Zhang, J.; Thiede, S.: A framework to classify Industry 4.0 technologies across production and product development. *Procedia CIRP* 84 (2019), pp. 973–978.
- Joung, C.B.; Carrell, J.; Sarkar, P.; Feng, S.C.: Categorization of indicators for sustainable manufacturing. *Ecological Indicators* 24 (2013), pp. 148–157.
- Kirchner, S.; März, L.: Towards self-adaptive production systems: Modular generic simulation models for continuous replanning and reconfiguration. *International Journal of Production Research* 40 (2002) 15, pp. 3627–3640.
- Langer, A.; Ortmeier, C.; Martin, N.L.; Abraham, T.; Herrmann, C.: Combining Process Mining And Simulation In Production Planning. In: Nyhuis, P.; Herberger, D.; Hübner, M. (Hrsg.): *Proceedings of the 2nd Conference on Production Systems and Logistics (CPSL 2021)*, Hannover, 2021, 2021,
- Martin, N.L.; Dér, A.; Herrmann, C.; Thiede, S.: Assessment of Smart Manufacturing Solutions Based on Extended Value Stream Mapping. *Procedia CIRP* 93 (2020), pp. 371–376.
- Orzes, G.; Poklemba, R.; Towner, W.T.: Implementing Industry 4.0 in SMEs: A Focus Group Study on Organizational Requirements. In: Matt, D.T.; Modrák, V.; Zsifkovits, H. (Hrsg.): *Industry 4.0 for SMEs*. Cham: Springer International Publishing; Imprint: Palgrave Macmillan 2020, pp. 251–277.
- Schönemann, M.; Bockholt, H.; Thiede, S.; Kwade, A.; Herrmann, C.: Multiscale simulation approach for production systems. *The International Journal of Advanced Manufacturing Technology* 102 (2019) 5-8, pp. 1373–1390.
- Schönemann, M.; Herrmann, C.; Greschke, P.; Thiede, S.: Simulation of matrix-structured manufacturing systems. *Journal of Manufacturing Systems* 37 (2015), pp. 104–112.
- VDI-2870 VDI-Fachbereich Fabrikplanung und -betrieb Ganzheitliche Produktionssysteme - Grundlagen, Einführung und Bewertung, 2012: VDI-Fachbereich Fabrikplanung und -betrieb Ganzheitliche Produktionssysteme - Grundlagen, Einführung und Bewertung, Lean production systems - Basic principles, introduction and review.
- Veile, J.W.; Kiel, D.; Müller, J.M.; Voigt, K.-I.: Lessons learned from Industry 4.0 implementation in the German manufacturing industry. *Journal of Manufacturing Technology Management* 31 (2020) 5, pp. 977–997.
- Wagner, T.; Herrmann, C.; Thiede, S.: Industry 4.0 Impacts on Lean Production Systems. *Procedia CIRP* 63 (2017), pp. 125–131.
- Wiendahl, H.-P.: *Betriebsorganisation für Ingenieure*. München: Hanser 2010.
- Wischmann, S.; Wangler, L.; Botthof, A. *Industrie 4.0, 2015: Industrie 4.0: Volkswirtschaftliche Faktoren für den Standort Deutschland, Eine Studie im Rahmen der Begleitforschung zum Technologieprogramm AUTONOMIK für Industrie 4.0.*



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Der Digitale Zwilling als echtzeitnahes Fertigungsabbild

Digital Twin-based real time representation of production facilities

Andreas Selmaier, Martin Sjarov, Meike Herbert, Jörg Franke, Lehrstuhl für
Fertigungsautomatisierung und Produktionssystematik/ FAU Erlangen-Nürnberg,
Erlangen (Germany), andreas.selmaier@faps.fau.de, martin.sjarov@faps.fau.de,
meike.herbert@faps.fau.de, joerg.franke@faps.fau.de

Johannes Distler, Jens Fürst, Technology Center for Power and Vacuum
Components/ Siemens Healthcare GmbH, Erlangen (Germany),
johannes.distler@siemens-healthineers.com, jens.fuerst@siemens-healthineers.com

Abstract: In the course of the fourth industrial revolution, numerous digital applications are being introduced along the entire life cycle of production systems. The structures of the classic automation pyramid are being softened and replaced by hybrid and decentralized structures in order to allow data and information to flow flexibly directly to where they are needed. For decision making processes in the manufacturing environment, a comprehensive overview of the factory and its current state is crucial. For this, specialised solutions are required to provide an intuitive access to the digital factory. The digital representation including enrichment with actual data falls under the paradigm "Digital Twin", which can be systematized by means of the Digital Twin Structure Model. In this context, the article discusses a corresponding solution as a "proof-of-concept". The underlying concept as well as a real-world application are presented accordingly. Furthermore, the added value and future potentials are laid out.

1 Einführung

Ein gebräuchliches Konzept zur Strukturierung der Automatisierungsarchitektur von Produktionssystemen ist die Automatisierungspyramide, welche die Teilsysteme gekoppelt, jedoch mit einer hierarchischen Verantwortlichkeitstrennung darstellt. Die starren Grenzen dieses hierarchischen Konzepts gehen mit Limitationen hinsichtlich der Flexibilität und Erweiterbarkeit einher. Im Zuge der aktuellen Digitalisierung im Rahmen von Industrie 4.0 werden daher unter Verwendung neuer Informations- und Kommunikationstechnologien die Grenzen der Automatisierungspyramide zunehmend aufgelöst, indem dezentrale Systeme Daten über die Grenzen zuvor getrennter Ebenen hinweg austauschen (vgl. Abb. 1) (Bettenhausen und Kowalewski, 2013).

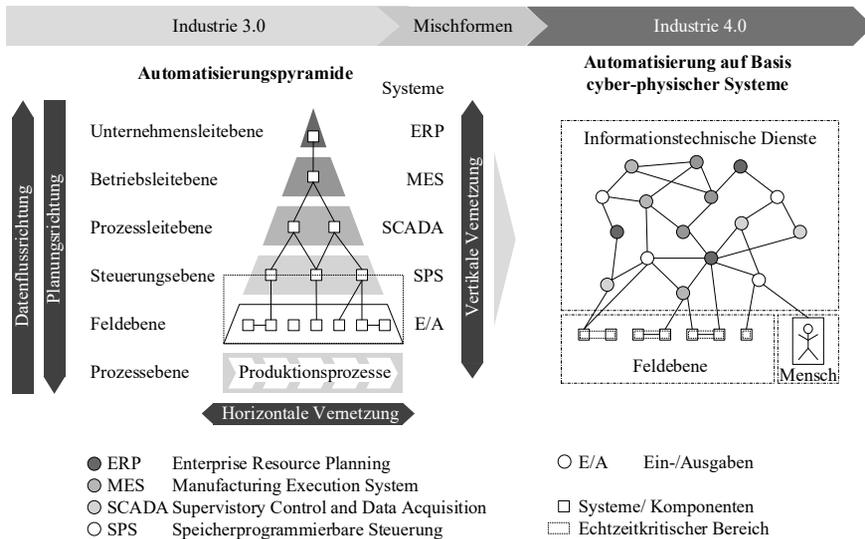


Abbildung 1: Auflösung starrer Strukturen der Automatisierungspyramide durch dezentrale und verteilte Dienste, i.A.a. (Fleischmann, 2019)

Dieser Vorgang, der auch als vertikale Integration bezeichnet wird, bietet Vorteile hinsichtlich der Erweiterbarkeit, Wiederverwendbarkeit und Wandlungsfähigkeit von Produktionssystemen. Während die Digitalisierung auf der einen Seite als Befähiger zur Erschließung von Optimierungspotentialen dient, führt die Hinzunahme neuer Softwarelösungen auch zu einem Komplexitätsanstieg des (Plattform I4.0). Die Realisierung eines Informationssystems (IS), welches diese Komplexität mittels eines echtzeit-nahen Abbilds der gesamten Fabrik in übersichtlicher Weise begreifbar macht und Inhalt des folgenden Beitrags ist, bringt auf strategischer, taktischer und operativer Ebene erhebliche Mehrwerte mit sich. Auf strategischer Ebene schaffen Leistungskennzahlen Transparenz und ermöglichen durch einen entsprechenden Drilldown Ursachen und Wirkzusammenhänge zu erkennen (Klein, 2015). Auf taktischer und operativer Ebene stellen Auftragsstatus sowie aktuelle Prozess- und Anlagenzustände wichtige Echtzeitinformationen da, und sind ferner eine Voraussetzung für betriebsbegleitende Simulationsaufgaben (Selmaier et al., 2019). Das Verknüpfen von Anlagen- und Prozessdokumentationen sowie Arbeits- und Wartungsplänen erhöht die funktionsübergreifende Zugänglichkeit und Transparenz. Durch die Nutzung aktueller 3D-Visualisierungsmöglichkeiten ermöglicht das vorgestellte IS einem breiten Anwenderkreis den Zugang zur digitalen Fabrik und erleichtert die Verortung von Informationen der realen Welt.

2 Stand der Technik und Forschung

2.1 Informationssysteme im Kontext I4.0

IS sind soziotechnische Systeme, die der Erfassung, Verarbeitung und Bereitstellung von Information und Kommunikation dienen. Sie werden zur Erreichung von Sachzielen unter der Einhaltung von Formalzielen eingesetzt (Eller und Riedl, 2016).

Auf Basis einer Studie fassen (Eller und Riedl, 2016) die strategischen Formalziele von IS sowie deren Beziehungen untereinander in einem Zielsystem zusammen, welches Abbildung 2 zeigt.

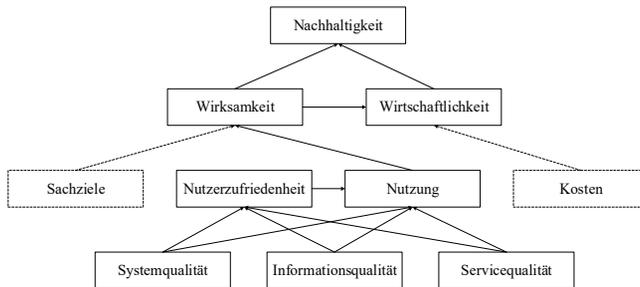


Abbildung 2: Zielsystem strategischer Formalziele von Informationssystemen i.A.a. (Eller und Riedl, 2016)

Dieses Zielsystem beschreibt Nachhaltigkeit als höchstes Ziel, welches durch die Wirksamkeit sowie die Wirtschaftlichkeit bedingt wird. Die Wirksamkeit wird wiederum beeinflusst durch den Erfüllungsgrad der Sachziele (vgl. Tab. 1) sowie die Nutzung. An dieser Stelle sei insbesondere das Formalziel der Systemqualität hervorgehoben, welches wiederum eine Vielzahl an Teilzielen wie Fehlerfreiheit, Sicherheit, Wartbarkeit, Flexibilität, Durchdringung, Zuverlässigkeit sowie die Benutzbarkeit beinhaltet (Eller und Riedl, 2016).

Ergänzend zu den Formalzielen, welche erfolgsentscheidende Kriterien für den Einsatz und Betrieb des IS darstellen, beschreiben Sachziele die zu realisierende Aufgabe in Abhängigkeit des Einsatzgebietes. Im Gegensatz zu Systemen, die der Komplexitätsreduktion dienen, werden IS zur Verbesserung der Beherrschbarkeit von komplexen Systemen (hier: Produktionssystemen) eingesetzt (Schoeneberg, 2014). Im Produktionskontext kommen so eine Vielfalt an klassischen IS (Industrie 3.0) zum Einsatz, deren Vielfalt im Zuge von Industrie 4.0 durch dezentrale Vernetzungsmöglichkeiten noch einmal erweitert wird. Diese IS lassen sich nach ihrer Aufgabenart wie in Tabelle 1 dargestellt kategorisieren. Zusammenfassend geht hervor, dass IS zur Steigerung der Prozessqualität und -effizienz von administrativen, dispositiven, planenden und ausführenden Tätigkeiten eingesetzt werden. Bei der Umsetzung moderner IS wird dabei bereits in Teilen das Entwicklungsparadigma des Digitalen Zwillings angewandt, vgl. (Lee et al., 2020; Fleischmann, 2019; Donhauser, 2020).

2.2 Digitaler Zwilling

Der „Digitale Zwilling“ ist als Schlagwort und aufkommendes Paradigma in den letzten Jahren verstärkt in den Fokus wissenschaftlicher Veröffentlichungen gerückt. Das Fazit eines diesbezüglichen Reviews ist, dass verschiedene, teils widersprüchliche Auffassungen und Definitionen zum Wesen des Digitalen Zwillings existieren, jedoch bestimmte wiederkehrende Muster identifizierbar sind (Sjarov et al., 2020). Um dem zu begegnen, wurde das Digital Twin Structure Model (DTSM) als Metamodell zusammen mit einer Definition des Digitalen Zwillings erarbeitet, welches in modifizierter Form in Abbildung 3 dargestellt ist (Lechler et al., 2020):

Der Digitale Zwilling ist hier der Definition nach der Teil des gleichnamigen Paradigmas, in welchem zu einem physischen Asset gehörende, digital erfasste Daten zusammen mit den digitalen Engineering-Modellen genutzt werden, um verschiedene Funktionen und Services bereitzustellen. Das DTSM gliedert zunächst den Lösungsraum entlang einer Lebenszyklus-Achse in „Typ“ und „Instanz“, wobei erstere der Entwurfs- und Engineering-Phase entspricht und letztere der tatsächlichen Fertigung und anschließenden Nutzung. In der Vertikalen wird zwischen der physischen Ebene unterschieden, in der das entsprechende Asset lokalisiert ist, sowie der Cyber-Ebene, welche alle digitalen Daten, Modelle und Funktionen enthält.

Tabelle 1: Gruppierung von Informationssystemen im Produktionskontext nach Aufgabenart

Aufgabenart	Informations- und IT-Systeme	Weiterführende Literatur
Wissensmanagement	<ul style="list-style-type: none"> ■ Wiki ■ Intranet / Unternehmensinterne soziale Netzwerke ■ Ontologiebasierte Wissensmanagement-Systeme* 	(Brandmeier, 2019)
Informationsbereitstellung	<ul style="list-style-type: none"> ■ PDM-Systeme ■ Business Intelligence- / Analytics-Umgebungen* 	(Sand, 2021)
Ausführungsaufgaben	<ul style="list-style-type: none"> ■ Manufacturing Execution Systems (MES) ■ Intralogistics Execution Systems (IES)* ■ Warehouse-Management-Systeme (WMS) ■ Ein- / Verkaufssysteme ■ Digitale Plantafeln / Shopfloorboards* ■ RPA-Plattformen* ■ Low-Code-Apps* 	(Meißner et al., 2018) (Langmann, 2021)
Planungs- und Steuerungssysteme	<ul style="list-style-type: none"> ■ ERP-Systeme ■ Dispositionssysteme ■ Simulationsgestützte Auftragsregelung* ■ Predictive-Maintenance-Systeme* 	(Donhauser, 2020)** (Lee et al., 2020)**
Kontrollsysteme	<ul style="list-style-type: none"> ■ Leitstände ■ SCADA-Systeme ■ Condition-Monitoring-Systeme* ■ Energieüberwachungs-Systeme* 	(Fleischmann, 2019)** (Rackow, 2016)

* Neue Anwendungen durch I4.0; ** Einsatz von Digitalen Zwillingen

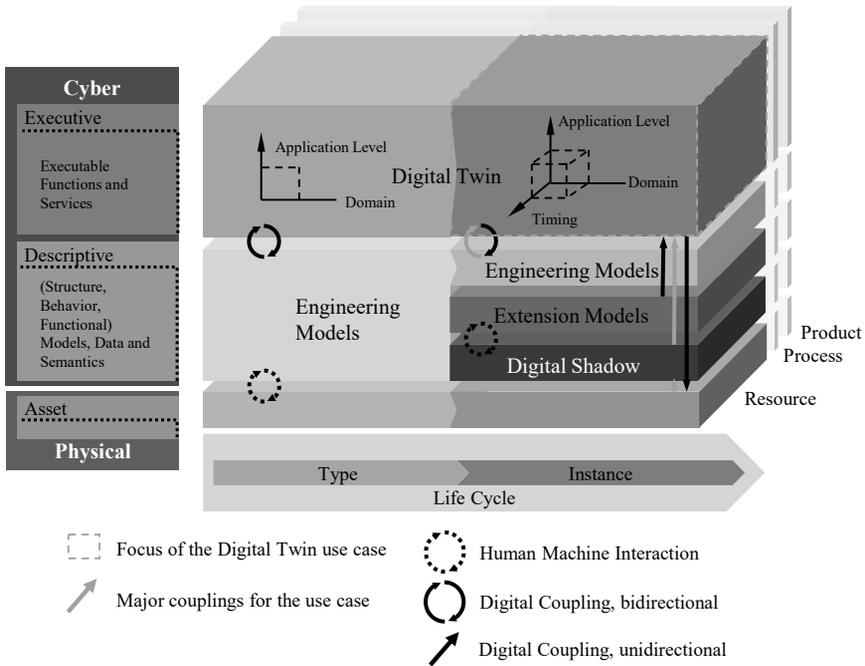


Abbildung 3: Einordnung des digitalen Fertigungsabbildes in das Digital Twin Structure Model (DTSM) i.A.a. (Lechler et al., 2020)

In der dritten Dimension wird zwischen den wesentlichen Artefakten „Produkt, Prozess und Ressource“ unterschieden, um eine weitere Granularität im Kontext der industriellen Anwendung bereitzustellen. Um nun Anwendungsfälle des Digitalen Zwillings unterscheiden zu können, sind diesem die Dimensionen Anwendungslevel, Domäne sowie Timing zugeordnet. Auch berücksichtigt das Modell etwaige Informationsflüsse auf konzeptioneller Ebene, ohne jedoch konkrete Technologie-Ausprägungen vorzuschreiben.

2.3 Zwischenfazit

Während IS auf der einen Seite die Beherrschbarkeit komplexer Systemen erleichtern, bilden sie auf der anderen Seite bereits bei mittelgroßen Unternehmen ein schwer zu überblickendes Netzwerk. Die fehlende Transparenz über den werksweiten Einsatz von IS führt zu Redundanzen und Medienbrüchen (Schoeneberg, 2014). Weiterhin erschweren mangelnde Sichtbarkeit und Zugänglichkeit von vorhandenen IS die ganzheitliche Berücksichtigung von Informationen. Somit wird der Bedarf eines IS abgeleitet, welches eine integrierende Funktion einnimmt und damit einen aktuellen Überblick über die Fertigung auf intuitive Weise gewährt. Ein solches ist der Aufgabenart „Informationsbereitstellung“ in Tabelle 1 zuzuordnen. Für die Erstellung stellt das Paradigma des Digitalen Zwillings einen vielversprechenden Ansatz dar, indem das physische Produktionssystem inklusive relevanter Informationsquellen in einem digitalen Modell abgebildet und mittels dedizierten Services unter Nutzung einer aktuellen Datenbasis den Mitarbeitern zugänglich gemacht wird.

3 Lösungsansatz

Die Lösung ist ein echtzeitnaher Digitaler Zwilling der Fertigung, welcher sich aus einem interaktiven, dreidimensionalen Modell sowie darin verorteter digitaler Assets und Realdaten zusammensetzt (vgl. Abb. 4). Als spezialisierte Systemlösung wird die webbasierte Software Intosite der Siemens AG eingesetzt.

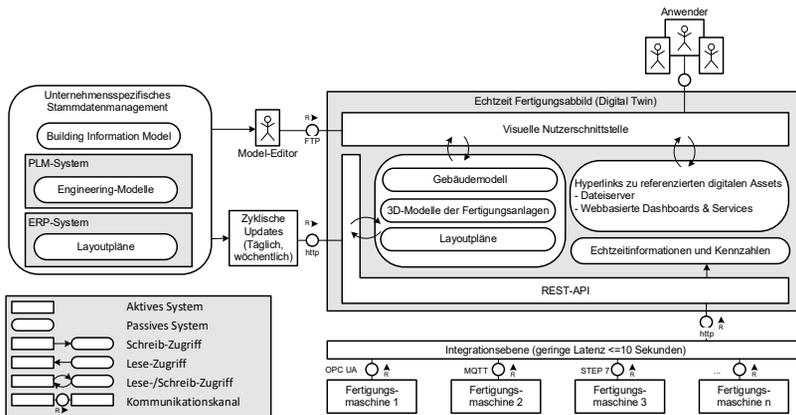


Abbildung 4: Architektur des digitalen Echtzeitabbildes der Fertigung (Vereinfachte Darstellung)

Die Erstellung des 3D-Modells der Fertigung lässt sich sowohl automatisiert über eine REST-Schnittstelle sowie manuell durch den Upload einzelner Modelle oder halbautomatisch mittels listenbasierter Imports realisieren. Da die Stammdaten in der Praxis meist in verschiedenen Systemen und Formaten vorliegen, sind hierbei entsprechende Konvertierungsvorgänge in die von Intosite unterstützten Datenformate vorzusehen. Echtzeitinformationen, wie Alarme, Kennzahlen oder aktuelle Prozess- und Anlagenzustände werden via REST-API mit geringer Latenz im Sekundenbereich aktualisiert. Zwar ermöglicht die Software einen direkten Upload von Dokumenten, doch sind diese bevorzugt über Verlinkungen einzubinden, um eine redundante Datenhaltung und Pflege zu vermeiden. Weiterführende webbasierte Inhalte, wie Dashboards oder entsprechende Abprünge in ERP- oder PLM-Systembereichen können über Rollen und Gruppen eingeschränkt werden. Während für die Aktualisierung der zugrundeliegenden Modelle ein Expertenwissen notwendig ist, können die eingebundenen Informationsquellen durch den Anwenderkreis selbst aktualisiert werden. Es entsteht ein interaktives und echtzeitnahes Fertigungsabbild, welches wie folgt durch das DTSM charakterisiert werden kann, siehe Abbildung 3 und Tabelle 2 (Lechler et al., 2020). Bei der Ausprägung des digitalen Fertigungsabbildes handelt es sich um eine Repräsentation von instanziierten Fertigungsressourcen unter Verwendung zugehöriger CAD-Engineering-Modelle. Die Applikation dient der mit Echtzeitdaten und Informationen angereicherten Visualisierung der Produktionsanlage. Die im DTSM als Digitaler Schatten bezeichneten Fertigungsdaten umfassen sowohl historische als auch echtzeitnahe Daten.

Table 2: Kategorisierung des Anwendungsfalls im DTSM

Anwendungsfall	PPR			Lebenszyklus		Anwendungs-Level				Zeitverhalten			Domäne				
	Produkt	Prozess	Ressource	Typ	Instanz	Visualisierung	Identifizierung	Vorhersage	Regelung	Asynchron	Echtzeit-nah	Echtzeit	Physisch	Logistisch	Ökonomisch	Software	Andere
Intuitives Echtzeitabbild von Produktionsanlagen			●		●	●	●			●	●		●				

4 Anwendungsfall

Der ausgewählte Standort in Erlangen für Hochleistungsrontgenstrahler der Siemens Healthcare GmbH beschäftigt einige hundert Mitarbeiter mit Fertigungsbezug, welche in einer funktionalen Aufbauorganisation eingeordnet sind. Der Fertigungstyp weist überwiegend den Charakter einer Werkstattfertigung auf. Auf den rund 200 Equipments und Fertigungsstationen werden jährlich über 15.000 Röntgenstrahler für Anwendungen der Computertomographie sowie der interventionellen Radiologie gefertigt. In Hinblick auf einen anstehenden Umzug des Werks in das neu zu errichtende High-Energy and Photonics Center (HEP) in Forchheim (InFranken, 2020), werden die zukünftigen informationsverarbeitenden Prozesse des Werks bereits in der Planung mit den Werkzeugen der Digitalisierung vorbereitet. Durch die Vielzahl an parallel stattfindenden Digitalisierungsaktivitäten und die damit einhergehenden Änderungen der IT-Landschaft, besteht ein hoher Synchronisations- und Einarbeitungsaufwand. Hiervon sind operative Tätigkeiten gleichermaßen betroffen, wie strategische IT-Konsolidierungsaufgaben. Infolgedessen wurde sowohl für das bestehende Werk in Erlangen Süd als auch für das neu entstehende HEP-Center in Forchheim ein Informationssystem nach dem in Abbildung 4 illustrierten Architekturkonzept prototypisch umgesetzt. Abbildung 5 zeigt die verschiedenen Ansichten und Navigationsmöglichkeiten des digitalen Fertigungsabbildes. Die Kartenansicht von Intosite ermöglicht einen geographischen Überblick sowie Verortung von Produktionsstätten des Unternehmens. Somit eignet sich die Lösung nicht nur für den werksweiten, sondern auch unternehmensweiten Einsatz. Im vorliegenden Anwendungsfall wurden sowohl Teile der bestehenden Werke in Erlangen, eine Testfläche für neues Fertigungsequipment in Forchheim, als auch Teile des HEP-Centers in Forchheim abgebildet. Die Modelle der einzelnen Standorte bestehen aus 2D-Layoutplänen, den 3D-CAD Modellen der Gebäude und Anlagen sowie nachfolgend erläuterten Platzierungszeichen („Placemarks“). Die Teilelemente des Gesamtmodells werden bei der Modellzusammenstellung in Hierarchie-Ebenen strukturiert, welche sich durch entsprechende An- und Abwahl ein- bzw. ausblenden lassen. Eine weitere Navigationsmöglichkeit stellen Zonen dar, welche eine zusätzliche Gruppierungslogik für zusammengehörige Fertigungsbereiche und Placemarks darstellen. In Abbildung 5c werden Zonen eingesetzt, um beispielsweise ganze Stockwerke ein- bzw. ausblenden zu können. Da die Modellstruktur eine

wesentliche Grundlage für die Bedienbarkeit des IS darstellt, wurde auf eine einheitliche Gestaltungslogik wert gelegt. Die Modellerstellung erfordert spezifisches Fachwissen und wird nach dem Anwenderkonzept in Abbildung 4 von vorgesehenen Modell-Editoren erstellt und bei Änderungen aktualisiert. Hingegen wird bei der Integration von Informationsquellen und -systemen mittels Placemarks ein Ansatz zur Pflege und Aktualisierung durch die Anwender selbst angestrebt.

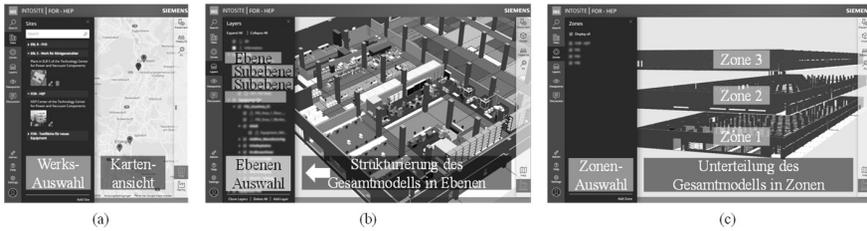


Abbildung 5: Ansichten und Navigationsmöglichkeiten des 3D-Fertigungsabbildes; (a) Kartenansicht mit Werksübersicht, (b) Strukturierung des Gesamtmodells in Hierarchie-Ebenen, (c) Zonenauswahl als Navigationshilfe und Visualisierung zusammengehöriger Modellbestandteile

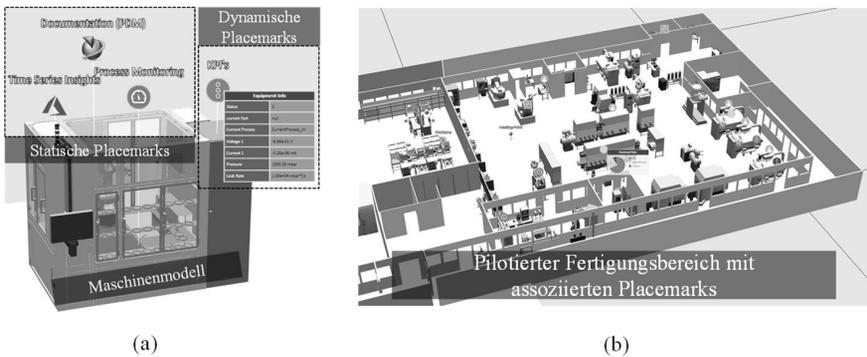


Abbildung 6: Detailansicht einer Maschine mit statischen und dynamischen Placemarks (a) sowie Momentaufnahme eines abgebildeten Fertigungsbereichs (b)

Innerhalb des Modells lassen sich durch die Anwender statische und dynamische Informationselemente mittels Placemarks platzieren. In statischen Placemarks können Verlinkungen zu anderen Informationssystemen, Adressinformationen oder auch Panoramaaufnahmen hinterlegt werden (vgl. Abb. 6a). Echtzeitnahe Informationen lassen sich in dynamischen Placemarks in das Modell integrieren und über eine REST-API automatisiert aktualisieren. Für die Darstellung der dynamischen Placemarks bietet Intosite drei verschiedene Optionen: Die Ampeldarstellung eignet sich für die Visualisierung von Anlagen- oder Prozesszuständen; das Kreisdiagramm zur Darstellung von prozentualen Verteilungen, wie Gut-, Nacharbeits-, und Ausschussmengen; in der Tabellenform lassen sich Kennzahlen und aktuelle Prozesswerte oder Auftrags-IDs abbilden. Die Verknüpfung von Informationselementen mit dem

Anlagenmodell (vgl. Abb. 6a) lässt sich analog auf einen gesamten Fertigungsbereich (vgl. Abb. 6b) übertragen. Da in der Praxis, wie auch im hier vorliegenden Anwendungsfall, die IT-Landschaft einer Fertigung stetigen Veränderungen unterliegt, ist eine fortlaufende Pflege und Aktualisierung des Modells notwendig. Während es für einen einzelnen Mitarbeiter eine mühsame Aufgabe darstellt, den aktuellen Zustand einer Fertigung zu erfassen und abzubilden, lässt sich das verteilte Detailwissen eines anwenderzentrierten Ansatzes nahezu aufwandsneutral zusammentragen.

5 Zusammenfassung und Ausblick

Das in diesem Beitrag vorgestellte Fallbeispiel zeigt, wie mit verfügbaren Tools ein dreidimensionales Fertigungsabbild realisiert werden kann, welches den aktuellen Ist-Zustand der realen Fertigung widerspiegelt. Das theoretisch stark diskutierte Paradigma des Digitalen Zwillings wird damit zunehmend operationalisiert. Die Verwendung eines dreidimensionalen Modells, welches Informationselemente in ihrem Kontext erscheinen lässt, ermöglicht eine intuitive Bedienung und erhöht zudem die Einprägbarkeit beim Wissenstransfer. Die durch das realgetreue Abbild entstehenden Mehrwerte sind insbesondere auf die gesteigerte Transparenz und Übersicht zurückzuführen. Dies ist sowohl für die Einarbeitung neuer Mitarbeiter, als auch für Konsolidierungs- und Entscheidungsprozesse in komplexen Systemen vorteilig. Mittels asynchronen Informationsaustausches erhalten Mitarbeiter die Möglichkeit über ihre Funktionsbereiche und -tools hinaus Zusammenhänge zu erkennen und Synergien zu nutzen. Weiterhin kann mit den Visualisierungsmöglichkeiten im Modell die Sichtbarkeit erfolgreicher Digitalisierungsansätze erhöht werden. Das Optimierungspotential der Lösung besteht in einer weiteren Automatisierung zur Synchronisation von Realität und digitalem Abbild sowie der Überführung des Systems von einem „Proof of Concept“-Status hin zu produktivem Einsatz auf Werks- und Unternehmensebene.

Danksagung

Die Ergebnisse wurden im Rahmen einer Forschungskooperation zwischen der Siemens Healthcare GmbH und dem Lehrstuhl für Fertigungsautomatisierung und Produktionssystematik (FAPS) der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) erarbeitet. Die Autoren möchten sich bei den beteiligten Mitarbeitern des Technology Centers for Power and Vacuum Components für die sehr gute Zusammenarbeit bedanken.

Literatur

- Bettenhausen, K.; Kowalewski, S.: Cyber-Physical Systems. Chancen und Nutzen aus Sicht der Automation. VDI (2013).
- Brandmeier, M.: Ganzheitliches ontologiebasiertes Wissensmanagement im Umfeld der industriellen Produktion 2019.
- Donhauser, T.: Ressourcenorientierte Auftragsregelung in einer hybriden Produktion mittels betriebsbegleitender Simulation, FAU University Press, Doctoralthesis, 2020.

- Eller, C.; Riedl, R.: Ziele von Informationssystemen. HMD Praxis der Wirtschaftsinformatik 53 (2016) 2, S. 224–238.
- Fleischmann, H.: Modellbasierte Zustands- und Prozessüberwachung auf Basis sozio-cyber-physischer Systeme, FAU University Press, Doctoralthesis, 2019.
- InFranken, 2020: Siemens Healthineers baut 24 Meter hohes Technologiecenter in Forchheim. inFranken.de, Dezember 2020. Online verfügbar unter <https://www.infranken.de/lk/forchheim/siemens-baut-24-meter-hohes-technologiecenter-in-forchheim-art-5136063>, zuletzt geprüft am 11.05.2021.
- Klein, A.: Unternehmenssteuerung mit Kennzahlen: Inkl. Arbeitshilfen online ; Auswahl, Ermittlung, Analyse, Kommunikation. s.l.: Haufe Verlag 2015.
- Langmann, C.: Robotic Process Automation (RPA) - Study on Characteristics of Successful RPA Implementations, 2021,
- Lechler, T.; Fuchs, J.; Sjarov, M.; Brossog, M.; Selmaier, A.; Faltus, F.; Donhauser, T.; Franke, J.: Introduction of a comprehensive Structure Model for the Digital Twin in Manufacturing. In: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria, 08.09.2020 - 11.09.2020, 2020, S. 1773–1780.
- Lee, J.; Ni, J.; Singh, J.; Jiang, B.; Azamfar, M.; Feng, J.: Intelligent Maintenance Systems and Predictive Manufacturing. Journal of Manufacturing Science and Engineering 142 (2020) 11.
- Meißner, A.; Hertle, C.; Metternich, J.: Digitales Shopfloor Management – Ihr Weg zur vernetzten Fabrik. Zeitschrift für wirtschaftlichen Fabrikbetrieb 113 (2018) 5, S. 281–284.
- Plattform I4.0: Mensch und Maschine: Neue Interaktionsformen eröffnen Chancen bei der Arbeitsorganisation und verändern die Anforderungen an Beschäftigte. Online verfügbar unter <https://www.plattform-i40.de/PI40/Redaktion/DE/Standardartikel/handlungsfelder-arbeit-40.html>, zuletzt geprüft am 19.02.2021.
- Rackow, T.: Erweiterung des Unternehmenscontrollings um die Dimension Energie. Bamberg: Meisenbach 2016.
- Sand, C.: Prozessübergreifende Analyse komplexer Montageprozessketten mittels Data Mining. Erlangen: FAU University Press 2021.
- Schoeneberg, K.-P.: Komplexitätsmanagement in Unternehmen. Wiesbaden: Springer Fachmedien Wiesbaden 2014.
- Scholz, M.: Intralogistics Execution System mit integrierten autonomen, servicebasierten Transportentitäten. Erlangen: FAU University Press 2019.
- Selmaier, A.; Donhauser, T.; Lechler, T.; Zeitler, J.; Franke, J.: Simulationsgestützte Produktionsplanung flexibler Fertigungssysteme. wt Werkstattstechnik - Online wt-online 2019 (2019), Seite 240-247.
- Sjarov, M.; Lechler, T.; Fuchs, J.; Brossog, M.; Selmaier, A.; Faltus, F.; Donhauser, T.; Franke, J.: The Digital Twin Concept in Industry – A Review and Systematization. In: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria, 08.09.2020 - 11.09.2020, 2020, S. 1789–1796.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Von der Simulation zum Experimentierbaren Digitalen Zwilling und zurück

From Simulation to Experimentable Digital Twins and back

Michael Schluse, Jürgen Roßmann, RWTH Aachen University, Aachen (Germany),
schluse@mimi.rwth-aachen.de, rossmann@mimi.rwth-aachen.de

Abstract: This contribution proposes a new methodology allowing for the in-depth analysis of complete systems of Cyber-Physical Systems in their operational environment which is the basis for various simulation-based methods in engineering and operation. Key structuring element and semantic entity is the Experimentable Digital Twin (EDT) which describes a system and its interfaces from a Model-based Systems Engineering (MBSE) perspective and wraps Simulation Models or Experimentable Models describing its behavior. EDTs bridge the gap between MBSE and simulation, leveraging the potential of state-of-the-art modeling, simulation and co-simulation approaches in a comprehensive and easy-to-use workflow.

1 Einleitung

Technische Systeme entwickeln sich immer mehr zu komplexen **Systemen von Cyber-Physischen Systemen**. Dieser Trend ist in nahezu allen Anwendungsbereichen zu beobachten, auch und insbesondere in Produktion und Logistik, und wird aktuell z.B. durch Modularisierung, moderne Sensorik, das Internet der Dinge, Industrie 4.0-Konzepte und Cloud-/Edge-Computing verstärkt. Diesen Trend müssen aktuelle Entwicklungen in der Simulationstechnik spiegeln, z.B. indem Modularisierung und IT-technische Vernetzung von der physischen auf die virtuelle Ebene übertragen werden (u.a. Rosen et al. 2021). Zudem erfordern heute drängende Fragestellungen die **simulationsgestützte Betrachtung vollständiger Systeme**, um z.B. die vielfältigen Wechselwirkungen zwischen Systemkomponenten untersuchen und damit das Verhalten ganzer Systeme in unterschiedlichen Betriebsituationen analysieren, optimieren und funktional validieren zu können. Konzepte der Virtuellen Inbetriebnahme können so umfassend angewendet oder aussagekräftige Trainings- und Validierungsdaten für Maschinelles Lernen generiert werden. **Emergenz** ist hier von zentraler Bedeutung, denn aus dem Verhalten einzelner Systembestandteile kann meist nur in Grenzen auf das Verhalten des Systems geschlossen werden.

Technische Systeme müssen hierzu direkt in ihren Einsatzumgebungen in unterschiedlichen Einsatzsituationen mit einer großen Bandbreite an sinnfälligen

Eingangsstimuli umfassend simulationsgestützt untersucht werden. Dabei reicht es häufig nicht mehr aus, diese Untersuchungen nur auf höheren Skalenebenen, z.B. mittels Ereignisdiskreter Simulation, durchzuführen. Vielmehr muss auch auf Systemebene das Verhalten der beteiligten Teilsysteme detailliert betrachtet werden. Dies leisten zeitkontinuierliche Simulationsverfahren für Kinematik, Starrkörperdynamik oder Sensorik bis hin zu FEM-Methoden. Zudem müssen Datenverarbeitungs- und Kommunikationssysteme berücksichtigt werden. Grundlage einer Systemsimulation sind entsprechend **vielfach gekoppelte Multi-Skalen-Simulationsmodelle, deren Modellierung, Orchestrierung und Simulation neue Strukturen, Methoden und Prozesse erfordern**, in denen existierende Simulatoren anwendungsspezifisch miteinander gekoppelt und mit Hard-/Softwaresystemen und Menschen verbunden werden.

Wie aber können in verteilten Prozessen von unterschiedlichen Akteuren derartige Simulationen konzipiert, die notwendigen Simulationsmodelle erstellt, durch eine Kombination der am besten geeigneten Simulationsalgorithmen simuliert und das Ergebnis (mehrfach) genutzt werden? Dieser Beitrag schlägt hierfür geeignete Strukturen/Prozesse vor und untersucht hierzu die weitgehende Trennung der Modellierungsebenen durch Kombination von Model-based Systems Engineering (MBSE) mit etablierten simulationstechnischen Untersuchungen auf Komponentenebene. Ziel ist also eine Methodik, die mit bestehenden Mitteln umgesetzt werden kann.

Kapitel 2 fasst zunächst den Stand der Technik zur Systemsimulation zusammen. Kapitel 3 führt den Experimentierbaren Digitalen Zwilling (EDZ) ein und zeigt, auf welchen Ebenen mit diesem Systeme modelliert werden können. Kapitel 4 skizziert, wie EDZ System- und Simulationsbetrachtung zusammenführen. Kapitel 5 legt diesen Workflow gegen die Randbedingungen der Praxis. Dies führt zum in Kapitel 6 vorgestellten Workflow, in dem EDZ den Simulationsprozess strukturieren. Der Beitrag schließt in Kapitel 7 mit einer Zusammenfassung.

2 Stand der Technik: Modellierung auf Simulatorebene

Auch wenn Simulationen „immer komplexere und umfassendere Untersuchungen auch auf Systemebene“ erlauben, ist der Stand der Technik aktuell stark durch **„detaillierte Analysemöglichkeiten auf Komponentenebene, jedoch limitierte Fähigkeiten auf Systemebene“** geprägt (Rosen et al. 2020). Zur Untersuchung von Systemen werden **Simulationsmodelle häufig entlang der Grenzen der jeweils eingesetzten Simulatoren entwickelt** und ggfls. anschließend miteinander gekoppelt. Alternativ kombinieren anwendungsspezifische Simulatoren ausgewählte Simulationsverfahren für vorab festgelegte Simulationsaufgaben. Leistungsfähige Simulationen auf Systemebene benötigen allerdings häufig die Zusammenführung der jeweils am besten geeigneten Simulatoren. Gleiches gilt, wenn Simulationen verteilt durch unterschiedliche Akteure (z.B. Abteilungen, Komponentenanbieter, Systemintegratoren) entwickelt werden. Entsprechend rückt Co-Simulation zur Integration von Teilmodellen in eine Simulation des Gesamtsystems nach Rosen et al. (2020)/Kuhn (2017) mehr und mehr in den Mittelpunkt. Gleichzeitig führt eine verteilte Werkzeuglandschaft heute zu diversen Medienbrüchen (Blumör et al. 2017).

Ausgangspunkt der Systemsimulation ist die Modellierung auf Systemebene. Hierzu steht eine Vielzahl von Sprachen und Ansätzen u.a. aus dem MBSE-Kontext zur Verfügung. SysML hat sich zu einem Standard entwickelt, Alternativen sind z.B. CAEX, AutomationML, das Referenzmodell zur Digitalen Fabrik nach IEC 62832 sowie die

Arbeiten zu IEEE P2806 und zur Verwaltungsschale. Auf Seite der Simulatoren ist die Situation gekennzeichnet durch herstellereigenspezifische Kombinationen aus Datenmodell/Datenformat für das Simulationsmodell und dem zugehörigen Simulator. Darüber hinaus existieren wenige übergreifende Ansätze wie Modelica oder SRML. Ein neutrales Austauschformat fehlt (Weigert et al. 2017). Systeme in einem einzigen Simulationsmodell vollständig zu beschreiben erscheint daher ausgeschlossen.

Deutlich besser stellt sich die Situation hinsichtlich des Austauschs Experimentierbarer Modelle dar (Gomes et al. 2017). Der De-Facto-Standard (Schweiger et al. 2019) ist hier das Functional Mockup Interface (FMI), daneben existiert eine Vielzahl herstellereigenspezifischer Lösungen. Im Bereich der verteilten Simulation haben sich die IEEE-Standards HLA und DIS etabliert. Allerdings sind Limitierungen zu beachten. So unterstützt z.B. FMI ausschließlich gerichtete Ein-/Ausgänge. Co-Simulationsansätze wurden bereits in vielfältiger Weise zur Realisierung von Simulationen auf Systemebene eingesetzt. Beispiele sind Scheifele et al. (2019) oder Härle et al. (2020). Letztere kombinieren ein Topologiemodell (AutomationML), Teilmodelle (FMU) und einen Orchestrator (OpenModelicaSimulator). Der Orchestrator führt die Teilmodelle zusammen, steuert den Zeitfortschritt und den Datenaustausch. Als Herausforderung wurden Modelle identifiziert, die sich über mehrere Module erstrecken. Scheifele et al. 2021 fassen unterschiedliche Ansätze und Vorteile zur Co-Simulation zusammen. Ein vielversprechender neuer Standard ist SSP (System Structure and Parameterization) zur Beschreibung von Systemen aus FMU.

Es gibt diverse Überlegungen zur Zusammenführung von detaillierten Simulationen und MBSE (z.B. mit SysML) und detaillierter Simulation (z.B. Kuebler et al. 2018; Kapos et al. 2019) oder zur Ableitung von Simulationsmodellen aus Systembeschreibungen (z.B. Scheifele 2019), die weitgehende Integration liegt allerdings noch in der Zukunft (INCOSE 2014). Dabei ist die hierdurch mögliche durchgängige modellbasierte Systementwicklung zentral (Stark et al. 2020). Als ein Ansatz erscheint das Konzept des Digitalen Zwillings (DZ), der MBSE in ein Digital Twin-based Systems Engineering überführen soll (Rosen et al. 2020). Herausforderung sind hier die vielen unterschiedlichen Sichten auf Begriff, Konzept, Nutzung und technische Umsetzung.

Im Ergebnis sind die Fragen, wie Systeme mit MBSE modelliert und Komponenten im Kontext von Simulation modelliert und simuliert werden, jeweils umfassend beantwortet. Die Zusammenführung bleibt aber immer noch eine Herausforderung. Dies gilt insbesondere, wenn unterschiedliche Modellierungskonzepte und -sprachen sowie Werkzeuge zusammentreffen. Ein umfassender methodischer und anwendungsübergreifender Ansatz zur Modellierung und Orchestrierung von Co-Simulationsanwendungen ausgehend von der Modellierung von Systemen existiert bislang nicht.

3 Grundidee: EDZ beschreiben Systeme und deren Interaktion

Zur übergreifenden Umsetzung von Simulation auf Systemebene sollte sich die Modellierung an der physikalischen Architektur des betrachteten Systems ausrichten. Im Mittelpunkt sollten Komponenten und Teilsysteme, deren mechanische, elektrische, hydraulische, thermale, IT-technische Schnittstellen und deren Verbindungen und Interaktion stehen – und nicht das einzelne Simulationssystem, dessen spezifische Abstraktion der realen Welt und das hierfür mit meist großem Aufwand

von Experten anwendungsspezifisch erstellte Simulationsmodell. Eine zentrale Frage ist jetzt, wie die entstehenden Modelle, Methoden und Prozesse beschrieben und zusammengeführt werden können. Das Konzept des **Digitalen Zwilling**s liefert hier neue Ansätze. Ein Digitaler Zwilling im Sinne dieses Beitrags ist eine unter gewählten Gesichtspunkten betrachtete virtuelle digitale 1-zu-1-Repräsentation seines Realen Zwilling (RZ, insbesondere in Bezug auf Semantik, Struktur, Verhalten und Interaktion), die eine Interaktion auf Grundlage aktueller Digitaler Artefakte ermöglicht.

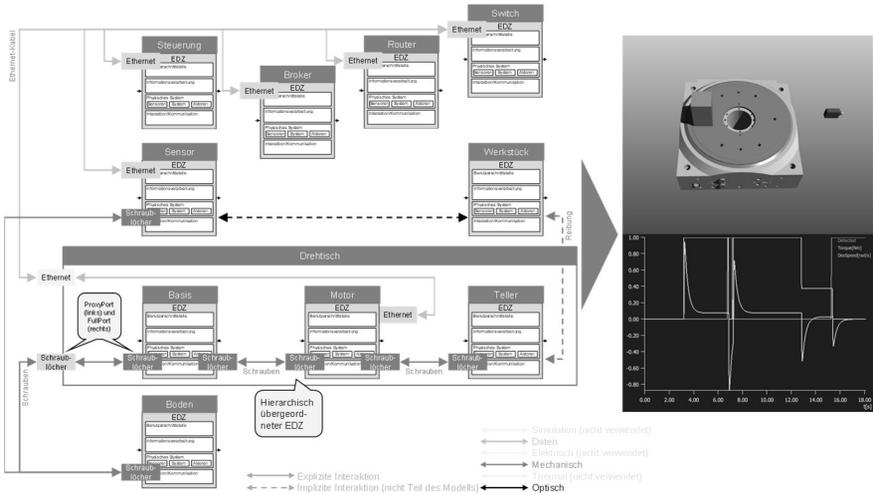


Abbildung 1: Grundlage für das Experimentierbare Systemmodell des Drehtisch-Beispiels rechts ist die Modellierung der Anwendung auf Ebene des Systems Engineerings links als Netzwerk interagierender EDZ. Die Struktur der beteiligten Komponenten orientiert sich an der physikalischen Architektur des Systems, die „Verschaltung“ der EDZ erfolgt entsprechend der Interaktion in der realen Welt.

Simulation erweckt Digitale Zwillinge zum Leben, macht sie und ihr Verhalten erfahrbar und analysierbar. Aus Digitalen Zwillingen (DZ) werden **Experimentierbare Digitale Zwillinge (EDZ)**, DZ zusammen mit aktueller Simulationstechnik bzw. DZ betrachtet aus der Perspektive der Simulationstechnik (Schluse et al. 2018). EDZ repräsentieren RZ, d.h. Physische, Informationstechnische, Mechatronische oder Cyber-Physische Systeme: **Ein EDZ ist eine virtuelle digitale 1-zu-1-Abbildung (Struktur, Verhalten, Interaktion) seines RZ in einem Experimentierbaren Modell.** EDZ/RZ interagieren miteinander in der virtuellen/realen Welt sowohl auf **physischer Ebene** durch Austausch von Energie und Materie (z.B. mechanisch, elektrisch, hydraulisch) als auch auf **informationstechnischer Ebene** durch Austausch von Nachrichten (z.B. über OPC UA, MQTT, Ethernet, CAN-Bus, serielle Schnittstelle). Aus Sicht von Modellierung und Simulation wird die Interaktion entweder **explizit** modelliert (z.B. Verschraubung oder Verkabelung) oder erfolgt **implizit** im Rahmen der Simulation (z.B. Kollision oder Detektion durch Sensor). Implizite Interaktionen sollten im Modell nicht formuliert werden; diese herzustellen ist Aufgabe der Simulation. Damit sind EDZ ideale Strukturierungselemente zur Beschreibung

von Systemen von Systemen. Die dargestellte Vorgehensweise konnte bereits für unterschiedliche Anwendungen in unterschiedlichen Domänen eingesetzt werden.

Die EDZ-Struktur wird durch Komponenten beschrieben, deren Schnittstellen durch Ports festgelegt sind und die durch Konnektoren verbunden sind (Abb. 2 links). Komponenten können selbst wieder EDZ sein. Auf oberster Hierarchieebene steht das EDZ-Szenario, welches ein konkretes Anwendungsszenario beschreibt. Die grundlegende Vorgehensweise/Darstellung orientiert sich an MBSE/SysML (Abb. 2 rechts). Die Empfehlung ist, die Modellierung an der physikalischen Architektur des Systems auszurichten. Dies senkt die Modellabstraktion, unterstützt die 1-zu-1-Beziehung zwischen EDZ und RZ und beantwortet hierdurch bereits eine Vielzahl von Fragen nach einer geeigneten Modellstruktur („Modelliere einfach so wie in der Realität!“). Dies erfolgt im Idealfall mit Ports und Konnektoren, die genauso auch in der realen Welt zu finden sind (Schraublöcher/Schrauben, Ethernet-Buchsen/Kabel etc.). Allerdings zeigt die Praxis, dass häufig auf Simulationsverfahren-spezifische Ports zurückgegriffen werden muss, z.B. zur Modellierung von Verbindungen für spezielle Kinematik-/Dynamik-Simulatoren oder bei Verwendung spezifischer Simulationssprachen wie Modelica. Schließlich benötigen Orchestratoren oft eigene Konstrukte zur Modellierung von Kommunikation (z.B. innerhalb der Komponenten auf funktionaler Ebene).

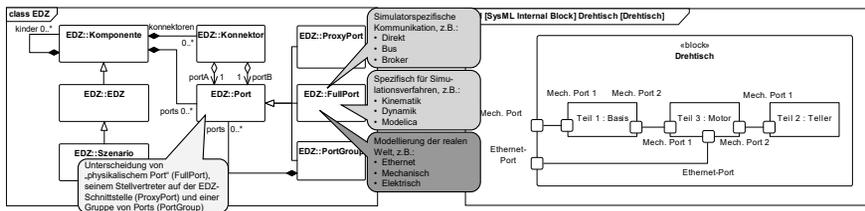


Abbildung 2: Grundelemente des abstrakten Datenmodells für EDZ (links), SysML-Modell des Drehtisch-Beispiels (rechts)

4 Konzept: EDZ führen System- und Simulationsbetrachtung zusammen

Auf dieser Grundlage führen EDZ System- und Simulationsbetrachtung zusammen (Abb. 3). Rechts unten ist der übliche Workflow der **Komponenten-/Simulator-spezifischen Modellierung und Simulation** dargestellt. Ausgehend von der (bestehenden oder neu zu entwickelnden) Realen Komponente wird ein *Konzeptuelles Komponentenmodell* aufgestellt, eine software-unabhängige Beschreibung eines Simulationsmodells, die dessen Ziele, Ein- und Ausgänge, Inhalte, Annahmen und Vereinfachungen beschreibt (Robinson 2008). Dieses wird in ein *Simulationsmodell* überführt, welches einem (meist Simulator-spezifischen) Meta-Modell folgt. Ein Simulator überführt das Simulationsmodell dann in ein *Experimentierbares Komponentenmodell*, welches ausführbar und damit Grundlage von Simulationsexperimenten ist und meist standardisiert (z.B. als FMU) zur Verfügung gestellt werden kann.

Zur Weiterverwendung dieser Ergebnisse auf einer Simulator-unabhängigen Systemebene werden die hier erzielten Ergebnisse in semantischen und insbesondere hinsichtlich ihrer Schnittstellen möglichst Simulator-unabhängigen Einheiten, den **EDZ**,

gekapselt (oben rechts). Deren Schnittstelle wird mit den Simulator-spezifischen Modellen „im Inneren“ verbunden, welche entweder als Experimentierbare Modelle oder als Simulationsmodelle zur Verfügung stehen. Aus Sicht objektorientierter Programmierung stellt der EDZ eine Fassade für die in seinem Inneren geeignet angeordneten Experimentierbaren Modelle/Simulationsmodelle dar und kapselt diese.

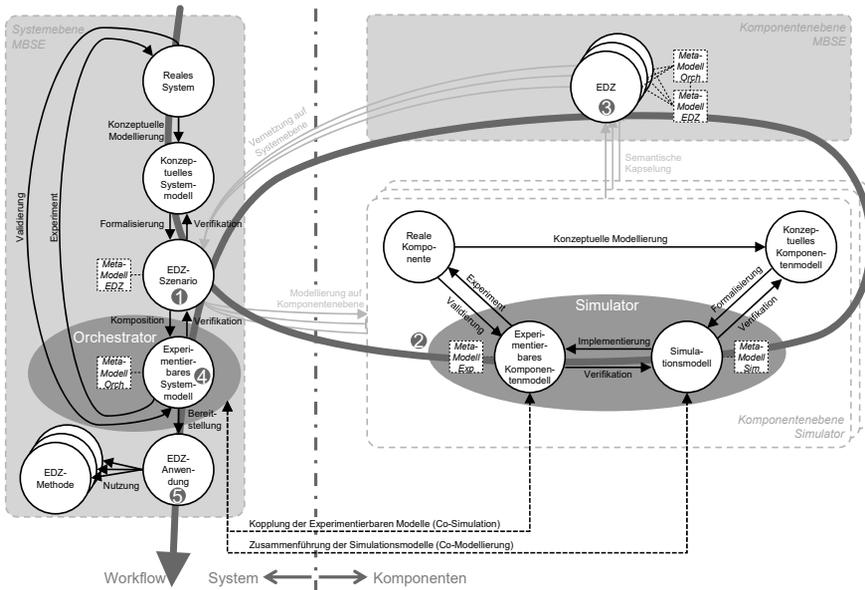


Abbildung 3: Zusammenführung von Modellierung auf Systemebene und Modellierung/Simulation auf Komponentenebene durch EDZ

Im Gegensatz hierzu sollte auf der **Systemebene** links Simulator-unabhängig gearbeitet werden. Im Idealfall werden etablierte MBSE-Methoden eingesetzt, die vollständig unabhängig von einer (später stattfindenden) Simulation und insbesondere unabhängig von einer (auf dieser Ebene möglichst zu vermeidenden) Simulationsabstraktion und einer Simulator-spezifischen Modellierung sind. MBSE und Simulation treffen sich auf Ebene der o.g. 1-zu-1-Repräsentation (hier Struktur/Interaktion). Das EDZ-Szenario legt fest, welche Systemkomponenten (d.h. welche EDZ) auf welchen Interaktionsebenen miteinander interagieren. Zur Erstellung des *Experimentierbaren Systemmodells* stellt das EDZ-Szenario den „Bauplan“ für dieses Modell zur Verfügung, entsprechend dem die EDZ bereitgestellten Experimentierbaren Modelle/Simulationsmodelle der Komponenten durch einen Orchestrator zusammengeführt und zur Verfügung gestellt werden. Die Komposition des Experimentierbaren Systemmodells erfolgt auf zwei Varianten. Einerseits können die EDZ Teil-Simulationsmodelle bereitstellen, die einem gemeinsamen Meta-Modell folgen. Diese werden zu einem Gesamt-Simulationsmodell zusammengeführt und durch einen Simulator in einem Experimentierbaren Modell implementiert (Co-Modellierung). Diese Zusammenführung erfolgt typischerweise entlang der Grenzen der Meta-Modelle, wobei theoretisch auch eine Überführung zwischen Modellen unterschiedlicher Meta-Modelle möglich wäre, wenn diese auf gleichen Meta-Meta-Modellen aufbauen

(Modelltransformation). Andererseits kann die Zusammenführung durch Kopplung der Experimentierbaren Komponentenmodelle erfolgen, die (weich (mehrere Solver) oder stark (ein Solver) miteinander gekoppelt) vernetzt werden (Co-Simulation).

Zur Umsetzung dieses Konzepts wird mindestens die **Interoperabilität** 1) der Simulatoren/Orchestratoren, 2) der Konzeptuellen Modelle (System/Komponenten), 3) der Meta-Modelle und 4) der Experimentierbaren Modelle/Simulationsmodelle der Komponenten/EDZ/EDZ-Szenarien benötigt. Zu deren Analyse haben u.a. Wang et al. (2009) im „Levels of Conceptual Interoperability Model (LCIM)“ Ebenen definiert:

0. Keine Verbindung

1. **Technisch (Physisch):** Kommunikationsverbindung besteht
2. **Syntaktisch (Daten):** Daten können auf syntaktischer Ebene interpretiert werden
3. **Semantisch (Information):** Daten können semantisch zugeordnet werden
4. **Pragmatisch (Verwendung):** Gleiches Verständnis der Verwendung der Daten
5. **Dynamisch (Verhalten):** Gleiches Verständnis des Verhaltens
6. **Konzeptuell:** U.a. Randbedingungen, Einschränkungen, Validität bekannt

LCIM ≥ 1 wird für grundsätzliche Integrierbarkeit benötigt, LCIM ≥ 4 für Interoperabilität, LCIM = 6 für Kompositionsfähigkeit. Es muss aber unterschieden werden zwischen der grundsätzlichen „Erstellbarkeit“ eines Experimentierbaren Systemmodells (Stufe 1), der Automatisierbarkeit dieses Prozesses (Stufe 2) und der Beurteilung der Validität des Ergebnisses vor dem Hintergrund des geplanten Einsatzes (Stufe 3). Für **Stufe 1** wird LCIM 3 benötigt, d.h. es müssen gemeinsame Kommunikationsprotokolle, Datenformate (z.B. Modelica, FMU) und Meta-Modelle (z.B. von Modelica oder FMU-Modellbeschreibungen) genutzt werden, anhand derer die Experimentierbaren Modelle/Simulationsmodelle in EDZ und diese dann in EDZ-Szenarien integriert werden können. Hierdurch sind die Grundlagen für die (ggfls. auch manuelle) Zusammenführung des Experimentierbaren Systemmodells gelegt. Der Verzicht auf die semantische Ebene erscheint hier nicht zeitgemäß. Für **Stufe 2** wird LCIM 4 benötigt. Hierdurch ist insbesondere eine Zuordnung zwischen den Ports der Teil-Modelle möglich (z.B. Ausgang „b“ der Regler-FMU ist die Reglerausgangsgröße und Eingang „u“ der Aktor-FMU ist der entsprechende Eingang). Für **Stufe 3** ist LCIM 6 notwendig, wobei ggfls. Anforderungen der LCIM 4 und 5 entfallen können.

5 Anwendung: Was bedeutet das in der Praxis?

LCIM 3 kann heute meist erreicht werden. Teilweise stehen Informationen zu den Experimentierbaren Modellen/Simulationsmodellen auf Ebene des Konzeptuellen Komponentenmodells zur Verfügung. LCIM 4 und 5 sind allerdings noch spärlich belegt. In der technischen Realisierung sollten und können die Meta-Modelle der EDZ/EDZ-Szenarien identisch ausgeführt werden. Die Kapselung von Experimentierbaren Komponentenmodellen in EDZ ist technisch einfach umzusetzen. Notwendig ist lediglich die Verbindung der EDZ-Ports mit den Experimentierbaren Komponentenmodellen. Hierzu stellt der Orchestrator eigene Ports/Konnektoren bereit (Abb. 4). Allerdings entsteht hierdurch eine direkte Abhängigkeit zwischen dem eingesetzten Orchestrator und der Modellierung der EDZ, die im Idealfall zu vermeiden ist. Hier besteht Bedarf einer Standardisierung.

Komplexer ist die Situation bei der Zusammenführung Simulator- und Komponentenspezifischer Simulationsmodelle. Dies ist sinnvoll, um z.B. maximal performante und

stabile Kinematik- oder Starrkörpersimulationen zu erreichen oder eine komponentenübergreifende Sensorsimulation zu ermöglichen. Hierzu müssen Modellelemente auf EDZ-Ebene (z.B. starre Verbindungen zwischen EDZ) mit Modellelementen auf Ebene der Komponenten-Simulationsmodelle (z.B. Starrkörper) zusammengeführt werden. Notwendig ist hierfür eine allgemeine Sicht z.B. auf eine mechanische Verbindung, die dann verlustfrei auf unterschiedliche Simulatoren umgesetzt werden kann, deren Sichtweise auf mechanischen Verbindungen entspricht (z.B. gerichtete vs. ungerichtete Verbindungen) und gleichzeitig alle benötigten Parameter zur Verfügung stellt. Ansätze hierfür liefert z.B. AutomationML. Ein allgemeiner Ansatz ist aktuell nicht in Sicht, auch hier besteht Bedarf einer Standardisierung – wenn sie denn überhaupt möglich ist (vgl. z.B. die Anforderungen an die Beschreibung einer mechanischen Verbindung aus Sicht von Kinematik, Starrkörperdynamik und FEM). Ein Ausweg ist, Simulator-spezifische Verknüpfungen auch auf Ebene der EDZ zu erlauben. Dies kann häufig durch „Anwendungsprofile“ für etablierte EDZ-Meta-Modelle wie z.B. SysML oder AutomationML erreicht werden, führt allerdings eine starke Abhängigkeit zwischen den eigentlich Simulator-unabhängigen EDZ und den Simulatoren ein. Die auf semantischer Ebene im EDZ-Meta-Modell notwendigen Änderungen halten sich allerdings in Grenzen und führen meist nicht zu einem benötigten tiefen Simulations-Know-how auf Ebene der Systemmodellierung (siehe Abb. 4).

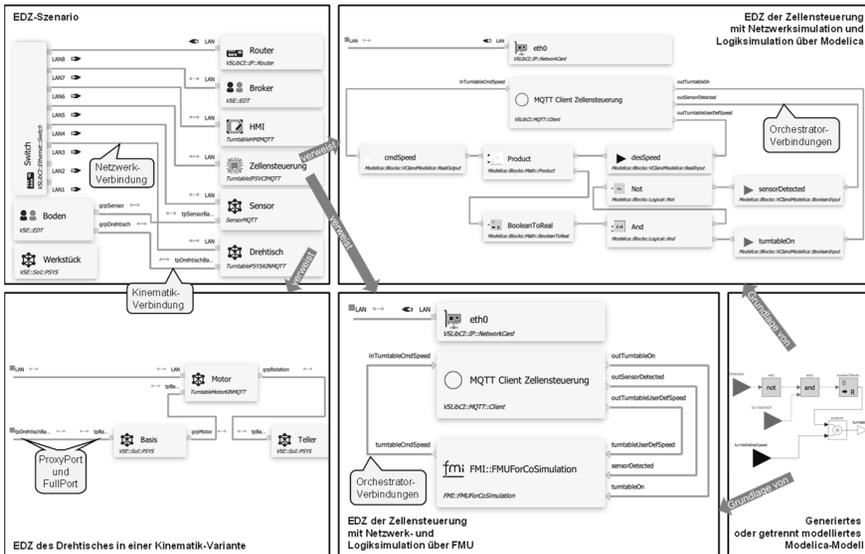


Abbildung 4: Zusammenführung von Komponentenmodellen im EDZ am Beispiel: EDZ-Szenario (l.o.), EDZ des Drehtisches in Kinematik-Variante (l.u.), EDZ der Zellensteuerung modelliert über Modelica (r.o.) oder integriert als FMU (r.u.)

6 Workflow: EDZ strukturieren den Simulationsprozess

Umgesetzt wird die skizzierte Grundidee in einem Workflow, der in fünf Stufen von der EDZ-basierten Systembeschreibung zur fertigen EDZ-basierten Simulationsanwendung führt (Blockpfeil in Abb. 3). In Schritt ① wird das EDZ-Szenario auf

übergeordneter (Gesamtsystem-) Ebene einschl. Beschreibung der Teilsysteme, deren Schnittstellen und deren Interaktion beschrieben. Diese Modellierung erfolgt im Idealfall in einem von Simulation unabhängigen Format wie z.B. SysML. Davon getrennt erfolgt in ② die Komponenten-/Simulator-spezifische Detailmodellierung. Das Ergebnis wird in ③ in EDZ zusammengefasst, die hierzu die erstellten Modelle beinhalten oder referenzieren. In ④ werden die jeweiligen Experimentierbaren Modelle/Simulationsmodelle automatisiert zusammengestellt, im Fall der Zusammenführung von Simulationsmodellen noch in Experimentierbare Modelle überführt und in standardisierten Simulationseinheiten, deren Struktur sich am FMI-Standard (Modelica 2014) orientiert, bereitgestellt. Auf dieser Grundlage wird ein Ausführungsplan abgeleitet, zur parallelen Ausführung optimiert und gekoppelt simuliert. Das Ergebnis wird dann in ⑤ in die EDZ-Anwendung integriert und so für unterschiedliche EDZ-gestützte Methoden wie Analyse, Test, Validierung, Optimierung, Inbetriebnahme, Training, Bedienung oder Steuerung genutzt.

7 Zusammenfassung

Im Ergebnis hat das dargestellte Konzept das Potenzial, die Komponenten-spezifische Modellierung auf Simulator-Ebene, die Systemmodellierung auf MBSE-Seite und den eingesetzten Orchestrator weitgehend und in Zukunft möglicherweise vollständig voneinander zu trennen und dennoch eine detaillierte Simulation auf Systemebene zu ermöglichen. Es legt weitere Grundlagen für die Überführung der Modularisierung und IT-technischen Vernetzung in die virtuelle Welt. Quasi als Nebeneffekt kann die Vernetzung über die Grenzen der virtuellen und realen Welten erfolgen, wodurch diese Welten im Sinne der Konvergenz weiter zusammenwachsen. Dies wurde z.B. für IP-basierte Kommunikation oder KNX-Bussysteme erfolgreich umgesetzt. Damit ist die Schleife vom EDZ zur Komponentensimulation und über den EDZ zurück zur Systemsimulation und zur EDZ-gestützten Anwendung konsistent geschlossen.

Der dargestellte Ansatz konnte erfolgreich in unterschiedlichen Anwendungsbereichen eingesetzt werden und hat sich darin als Grundlage für den Einsatz EDZ-gestützter Methoden bewährt. Zur praktischen Umsetzung heute muss das Konzept an den beiden beschriebenen Stellen aufgeweicht werden, durch die Abhängigkeiten zwischen Orchestrator-, Simulator- und MBSE-Meta-Modellen entstehen. Daher sind die Entscheidungen, welcher Orchestrator eingesetzt und welche Simulator-spezifischen Ports/Verbindungen eingesetzt werden sollen, elementar und zentral für den Aufbau einer EDZ-Bibliothek. Standardisierung im Kontext der Integration von Experimentierbaren Modellen/Simulationsmodellen in EDZ sowie im Bereich Simulator-spezifischer Ports/Verbindungen im EDZ-Szenario kann diese Abhängigkeit reduzieren und zu mehr Flexibilität führen, ist aber aktuell nicht in Sicht. Die Lücken bzgl. LCIM 4 und 5 verhindern eine vollautomatisierte Integration von Simulationsmodellen in EDZ, fehlende Informationen auf LCIM 6 erschweren insbesondere die Validierung des Experimentierbaren Systemmodells. Drei wichtige Empfehlungen sind, dass die Systemmodellierung sich an der physikalischen Systemarchitektur ausrichten, auf die Modellierung impliziter Interaktion verzichten und so weit wie möglich unabhängig von der späteren simulationstechnischen Umsetzung erfolgen soll.

Literatur

- Blumör, A.; Pregitzer, G.; Bothen, M.: Werkzeuge für die Entwicklung mechatronischer Systeme mit Methoden des MBSE. In: Tag des Systems Engineering, Carl Hanser Verlag, 2017.
- Gomes, C.; Thule, C.; Broman, D.; Larsen, P. G.; Vangheluwe, H.: Co-simulation: State of the art. <http://arxiv.org/abs/1702.00686>, 2017.
- Härle, C.; Barth, M.; Fay, A.: Simulationsmodellgenerierung im modularen Maschinen- und Anlagenbau - Assistenzsystem zur automatischen Komposition und Konfiguration von Co-Simulationen. *atp magazin* 09 2020.
- INCOSE: Systems Engineering Vision 2025. <http://www.incose.org/docs/default-source/aboutse/se-vision-2025.pdf?sfvrsn=4>, 2017.
- Kapos, G.-D.; Tsadimas, A.; Kotronis, C.; Dalakas, V.; Nikolaidou, M.; Anagnostopoulos, D.: A Declarative Approach for Transforming SysML Models to Executable Simulation Models. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019.
- Kübler, K., Scheifele, S., Scheifele, C., & Riedel, O.: Model-Based Systems Engineering for Machine Tools and Production Systems (Model-Based Production Engineering). *Procedia Manufacturing*, 24, 216–221, 2018.
- Kuhn, T.: Digitaler Zwilling. *Informatik-Spektrum*, 40(5), 2017.
- Modelica Association Project “FMI.”: Functional Mock-up Interface for Model Exchange and Co-Simulation - version 2.0. 2014.
- Rosen, R.; Jäkel, J.; Barth, M.; Stern, O.; Schmidt-Vollus, R.; Heinzerling, T.; ... Röhler, M.: Simulation und digitaler Zwilling im Anlagenlebenszyklus - Standpunkte und Thesen. *VDI-Statusreport* Februar 2020.
- Rosen, R.; Heinzerling, T.; Schmidt, P. P.; Jäkel, J.; Kübler, K.; Ristic, M.; ... Stern, O. (2021): Die Rolle der Simulation im Kontext des Digitalen Zwillings. *atp magazin* 04 2021.
- Scheifele, S.: Generierung des Digitalen Zwillings für den Sondermaschinenbau mit Losgröße 1. *Fraunhofer Verlag*, 2019.
- Scheifele, C.; Härle, C.; Rosen, R.; Heinzerling, T.; Röhler, M.: Co-Simulation als Realisierung digitaler Zwillinge. *atp magazin* 04 2021.
- Schluse, M.; Priggemeyer, M.; Atorf, L.; Rossmann, J.: Experimentable Digital Twins—Streamlining Simulation-Based Systems Engineering for Industry 4.0. *IEEE Transactions on Industrial Informatics*, 14(4), 2018.
- Schweiger, G.; Gomes, C.; Engel, G.; Hafner, I.; Schoeggel, J.; Posch, A.; Nouidui, T.: An empirical survey on co-simulation: Promising standards, challenges and research needs. *Simulation Modelling Practice and Theory*, 95, 2019.
- Stark, R.; Anderl, R.; Thoben, K.-D.; Wartzack, S.; Krause, F.-L.; Grässler, I., ... Göckel, N.: *WiGeP - Positionspapier: Digitaler Zwilling*, 2020.
- Weigert, D.; Aurich, P.; Reggelin, T.: Durchgehende Modellerstellung zwischen Simulations-, Visualisierungs- und Konstruktionswerkzeugen für die gesamtheitliche Planung von Produktions- und Intralogistiksystemen. In: S. Wenzel & T. Peter (Eds.), *Simulation in Produktion und Logistik* 2017.
- Robinson, S.: Conceptual modelling for simulation Part I: Definition and requirements. *Journal of the Operational Research Society*, 59(3), 278–290, 2008.
- Wang, W.; Tolk, A.; Wang, W.: The levels of conceptual interoperability model: Applying systems engineering principles to M&S. In: *Spring Simulation Multiconference* 2009



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Modulare Simulation bei der energieorientierten Planung industrieller Heiz- und Kühlsysteme

Modular Simulation in Energy-oriented Planning of Industrial Heating and Cooling Systems

Daniel Moog, Fabian Borst, Matthias Weigold, TU Darmstadt, Institut für Produktionsmanagement, Technologie und Werkzeugmaschinen (PTW), Darmstadt (Germany), d.moog@ptw.tu-darmstadt.de, f.borst@ptw.tu-darmstadt.de, m.weigold@ptw.tu-darmstadt.de

Abstract: Energy-oriented simulation-based planning of industrial heating and cooling supply systems is an ongoing research topic. The reduction of time and effort for simulation remains a central challenge. In this paper, a three-step approach for thermohydraulic system simulation is presented for this purpose. First, a modular technology model kit is developed in Modelica/Dymola. Then, the Modelica modules are transferred into a model exchange format. Finally, a (partially) automated model generation of system variants as composite models and their co-simulation in Python is performed.

1 Einleitung und Motivation

Industrielle Heiz- und Kühlsysteme wiesen im Jahr 2017 einen Anteil von nahezu 75 Prozent am Endenergieverbrauch der deutschen Industrie auf (Bundesministerium für Wirtschaft und Energie, 2019). Die betriebliche Kälte- und Wärmeversorgung ist ein zentraler Aspekt der Versorgungstechnik, welche den äußeren Rahmen, sprich die Infrastruktur der Produktion, bildet (Fraunhofer Gesellschaft, 2008). Ihre Planung ist eine Teilaufgabe der Fabrikplanung und somit wichtiger Bestandteil der Planung eines kompletten Produktions- oder Logistikstandorts. Unter anderem bedingt durch politische Rahmenbedingungen gewinnt die Berücksichtigung von Energiezielen bei dieser Planungsaufgabe zunehmend an Bedeutung (Burggräf und Schuh, 2021). Eine systemische Ausschöpfung bspw. von Energieeffizienz- oder Energieflexibilitätspotenzialen kann komplexe Systemtopologien und zugehörige Betriebsstrategien erfordern. Die Simulation als etabliertes Werkzeug im Rahmen der Digitalen Fabrik kann den Planungsprozess derartiger komplexer Systeme unterstützen (VDI 4499, 2008). Die Modellierung industrieller Heiz- und Kühlsysteme und deren Untersuchung in dynamischen Simulationen erfordert jedoch umfangreiches Know-how über die Simulationstechnologie. Weiterhin besteht eine hohe Anzahl möglicher Systemvarianten, welche nur unter hohem

Zeitaufwand modelliert, simuliert und ausgewertet werden können. Im Rahmen einer Auftragsplanung als Teil einer Ingenieur-Dienstleistung bzw. einer Energieberatungsdienstleistung kann daher bedingt durch Zeit, Aufwand und Kosten in der Regel kein dynamisches Simulationsmodell verschiedener Varianten, ausgehend von einzelnen Komponentenmodellen, aufgebaut werden. Nach Fowler und Rose (2004) stellt die Verringerung von Zeit und Aufwand im Kontext der Simulation eine fortwährende Herausforderung dar. Vor diesem Hintergrund wird in dieser Arbeit die Grundlage zur modularen simulationsgestützten Planung industrieller Heiz- und Kühlsysteme basierend auf Modelica/Dymola geschaffen. Es wird ein Technologiemodellbaukasten vorgestellt, der eine modulare und damit zeiteffiziente Modellierung komplexer Systemtopologien industrieller Heiz- und Kühlsysteme innerhalb Modelica-basierter Simulationsumgebungen ermöglicht. Anschließend erfolgt die Befähigung und der Export der Modelica-Module in ein softwareunabhängiges Modellaustauschformat. Darauf aufbauend wird ein strukturbasierter Ansatz zur (teil-)automatisierten Modellerstellung industrieller Heiz- und Kühlsysteme in Python entwickelt, in dem die Module in einem Systemmodell zur Co-Simulation (CS) verbunden werden.

2 Stand von Wissenschaft und Technik

Für Zwecke der systemischen Gebäude- und Anlagensimulation steht eine Vielzahl freier und kommerzieller Modellbibliotheken in Modelica zur Verfügung, die eine objektorientierte Modellierung ausgehend von Komponentenmodellen ermöglichen, darunter auch die freie *Modelica Standard Library (MSL)*. Eine weiterführende Übersicht verschiedener Modelica-Modellbibliotheken für die Gebäude- und Anlagensimulation ist bspw. in Müller et al. (2017) dargestellt. Die komponentengetriebene Modellierung industrieller Energiesysteme wurde in der Literatur umfassend beschrieben (siehe bspw. Wischhusen (2005), Panten (2019)). Im Kontext der Simulation gebäudeseitiger Versorgungstechnik entwickeln Flach et al. (2019) erstmalig Modelica-basierte Simulationsmodule, die wiederkehrende Verteilstrukturen in thermohydraulischen Systemen zur Wärme- und Kälteversorgung abbilden.

Trotz der umfassenden Vorarbeiten zur Modellierung gilt die Anwendung der dynamischen Simulation zur Planung industrieller Heiz- und Kühlsysteme als spezialisiert und stellt keinen flächendeckenden Standard in der Planungspraxis dar (vgl. Straßburger et al. (2006)). Aktuelle Forschungsarbeiten fokussieren diesen Ansatz jedoch, insbesondere im Kontext der Planung unter Berücksichtigung energieorientierter Zielgrößen (siehe bspw. Kurlle (2018) und Flum (2021)).

Mit dem *Functional Mock-up Interface (FMI)* steht eine standardisierte Schnittstelle zur Verfügung, gemäß derer Simulationsmodelle als Functional Mockup-Units (FMU) sprach- und umgebungsunabhängig ausgetauscht werden können (Modelica Association, 2020). Die Technologie bietet großes Potenzial, um die Kommunikation zwischen unterschiedlichen Planungswerkzeugen und letztlich auch zwischen den Planungsbeteiligten zu verbessern (Réhault et al., 2017). FMU-Modelle können zudem zur Umsetzung webfähiger Simulationswerkzeuge eingesetzt werden.

Ein Trend in simulationsgestützten Planungsprozessen ist darüber hinaus die automatisierte Modellerstellung mit dem Ziel der Aufwandsreduktion (Bergmann

und Straßburger, 2020). Nach Bergmann und Straßburger (2020) besteht eine Tendenz zur Verwendung von XML-Dokumenten als mögliche Datenquelle für eine automatisierte Modellerstellung.

Zusammenfassend lässt sich festhalten, dass umfassende Vorarbeiten im Bereich der Modelica-basierten Simulation industrieller Heiz- und Kühlsysteme bestehen. Deren modulare Modellierung zur Reduktion des Zeit- und Modellierungsaufwands ist bislang jedoch nur im Einzelfall aus Perspektive der Gebäudetechnik untersucht worden. Weiterhin beschäftigen sich aktuelle Forschungsarbeiten mit der Entwicklung simulationsgestützter Tools, die die energieorientierte Planung industrieller Heiz- und Kühlsysteme im Sinne von Assistenzsystemen unterstützen. Eine Kombination mit Ansätzen zur automatisierten Modellerstellung wurde in diesem Bereich jedoch noch nicht umgesetzt.

3 Ansatz zur modularen Simulation bei der energieorientierten Planung industrieller Heiz- und Kühlsysteme

Ausgehend vom Stand der Wissenschaft und Technik wurde ein sukzessiver Ansatz mit dem übergeordneten Ziel der Ersparnis von Zeit und Aufwand im Kontext der energieorientierten, simulationsgestützten Planung industrieller Heiz- und Kühlsysteme formuliert (Abb. 1). Die einzelnen Schritte des Ansatzes werden nachfolgend in entsprechenden Unterkapiteln ausgeführt.

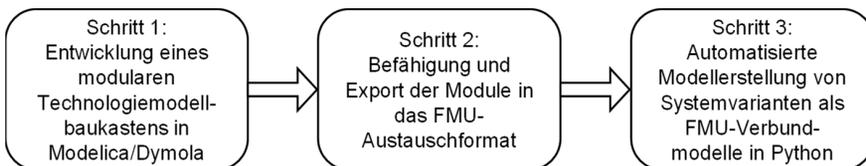


Abbildung 1: Schrittweises Vorgehen im entwickelten Ansatz

3.1 Schritt 1: Entwicklung eines modularen Technologiemodellbaukastens in Modelica/Dymola

Den Ansatz von Flach et al. (2019) aufgreifend wurde ein Technologiemodellbaukasten zur Modellierung industrieller Heiz- und Kühlsysteme entwickelt. In Anlehnung an VDI 2073-1 (2014) sowie VDI 6018 (2018) enthält er Module aus den Funktionsbereichen *Erzeugung*, *Rückkühlung*, *Entkopplung*, *Nutzenübergabe* am Verbraucher und *Wärmerückgewinnung*. Es wurde auf die Modellbasis der *MSL* (Version 3.2.3) zurückgegriffen. Komponentenmodelle versorgungstechnischer Anlagen wurden aus der *ETAFactorySimulationLibrary* (Panten, 2019) verwendet und diese um eigens entwickelte kennfeldbasierte Pumpen- und Rohrmodellklassen erweitert.

Abbildung 2 zeigt das UML-Klassendiagramm eines Systemmodells. Die kleinste strömungsführende Moduleinheit ist ein partielles *Basismodul*, das über je zwei Fluideingänge und -ausgänge, basierend auf *MSL FluidPorts*, verfügt. Die Anordnung der Fluideinlässe und -auslässe erfolgte so, dass die *Basismodule* die in

hydraulischen Verteilsystemen übliche Einteilung der Strömungsführung in Vorlauf (von Erzeugern zu Verbrauchern gerichtet) und Rücklauf (von Verbrauchern zu Erzeugern gerichtet) aufweisen. Das *Basismodul* ist eine Generalisierung des *hydraulischen Basismoduls*, welches neben Pumpen und Ventiltechnik ein Rohrmodell zur Berücksichtigung von Strömungswiderständen und Wärmeverlusten an die Umgebung enthält. Die Aktoren eines hydraulischen Basismoduls werden durch ein entsprechendes *hydraulisches Regelungsmodul* gesteuert. Hydraulische Basismodule und Regelungsmodule sind elementare Bestandteile eines *thermohydraulischen Moduls*, wobei letzteres aus mehreren dieser Klassen bestehen kann. Darüber hinaus werden darin Komponentenmodelle thermischer Wandler (Erzeuger oder Verbraucher) oder Speicher integriert. Die Regelung thermohydraulischer Module erfolgt über ein *Systemregelungsmodul*. Sofern die Erstellung des Systemmodells in einer softwareunabhängigen Umgebung erfolgen soll, erben thermohydraulische Module die Eigenschaften eines *FMI-Schnittstellenmoduls*, wodurch sie zum Export in das FMU-Modellaustauschformat befähigt werden.

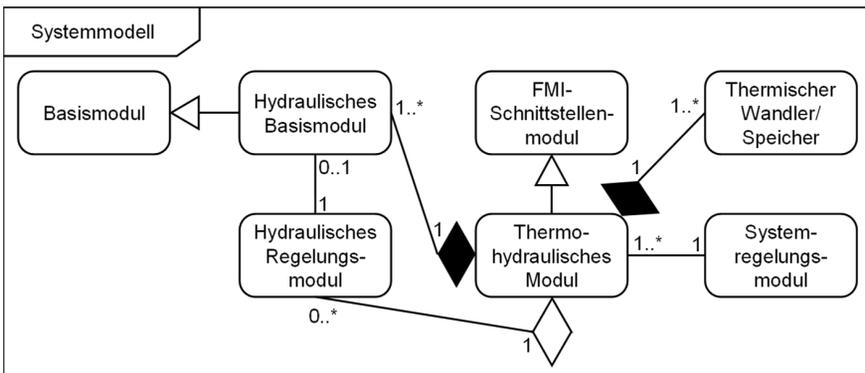


Abbildung 2: Vereinfachtes UML-Klassendiagramm des entwickelten, Modelica-basierten Technologiemodellbaukastens

Abbildung 3 zeigt beispielhaft die schematische Modellansicht eines thermohydraulischen Moduls zur Kälteerzeugung. Es hat die Funktion eine gezielte Solltemperatur des Fluids im Kaltwasservorlauf einzustellen. Im Kern beinhaltet es ein Komponentenmodell einer Kompressionskältemaschine sowie ein hydraulisches Basismodul mit integrierter Pumpe und Rohr widerständen. Die Instanzen der Kompressionskältemaschine sowie der Pumpe werden über vordefinierte Technologiedatensätze mit Herstellerdaten parametrisiert. Die hierfür eingeführten Parameter wurden auf die Modellebene des thermohydraulischen Moduls propagiert. Das Regelungsmodul übergibt Stellsignale an die Kältemaschine und an die Pumpe über ein Bussignal. Über ein weiteres Bussignal werden Sensorwerte vom Sensormodul im Kaltwasservorlauf übergeben.

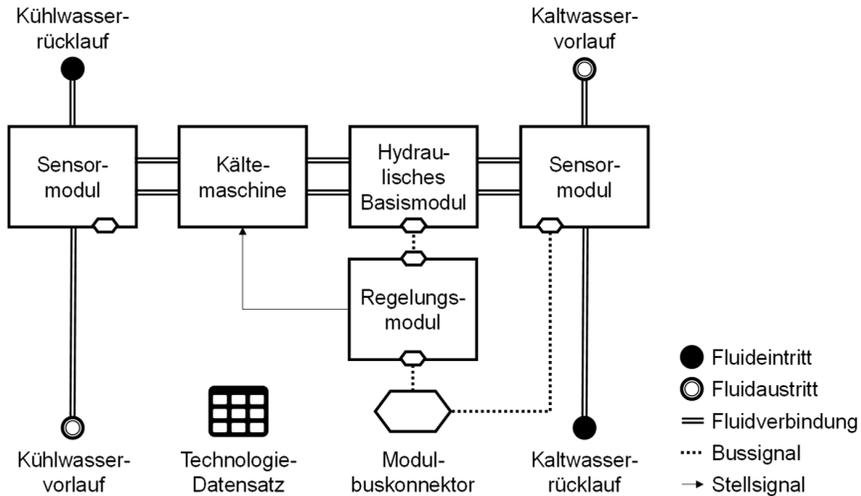


Abbildung 3: Schematische Ansicht eines thermohydraulischen Moduls zur Kälteerzeugung

3.2 Schritt 2: Befähigung und Export der Module in das FMU-Austauschformat

Im zweiten Schritt des Ansatzes erfolgt die Befähigung der Module zum optionalen Export in das FMU-Modellaustauschformat als Basis für die Python-basierte Simulation. FMU-Modelle können gemäß dem FMI-Standard 2 im Format CS zur gekoppelten Simulation oder Model Exchange (ME) zum reinen Modellaustausch exportiert werden. Die FMI-CS-Spezifikation bietet eine Modelldarstellung, bei der sowohl das Modell als auch ein Integrator innerhalb der FMI-CS gekapselt sind (Modelica Association, 2020). Im industriellen Anwendungskontext wird FMI-CS gegenüber FMI-ME eine größere Bedeutung zugeschrieben, unter anderem weil eine modellspezifische Adaptierung des Integrators bereits erfolgt ist (Ogata et al., 2014).

Auf der Modelica-Seite erfordert der Export als FMU die Einführung von Adapter-Modellen, welche akausale in kausale Signalflüsse transformieren (Hirano et al., 2015). Basierend auf den von Wetter et al. (2015) entwickelten FMI-ME-Adaptoren für Fluid-Konnektoren wurde ein FMI-Schnittstellenmodul entwickelt, das für den Export der thermohydraulischen Module als FMI-CS geeignet ist (Abb. 4). In FMI-CS fehlt grundsätzlich die Funktionalität zur direkten Ausgabe nach Veränderung von Eingangswerten ohne numerische Iteration, woraus sich Einschränkungen für die Berechnung stationärer Thermofluidströmungen in gekoppelten FMI-CS-Modellen ergeben (Wetter et al., 2015). Vereinfachend wurde daher angenommen, dass die Berechnung des Massenstroms am Ein- und Austritt eines Moduls ausschließlich von modulinternen Strömungswiderständen abhängt. Bei der Implementierung wurde in massenstromerzeugende und -empfangende Modulabschnitte unterschieden. Ein massenstromerzeugender Modulabschnitt enthält eine Strömungsmaschine und ein thermohydraulisches Rohrmodell; ein massenstromempfangendes Modul enthält ausschließlich ein thermohydraulisches Rohrmodell. Damit wird der durch die Strömungsmaschine erzeugte Massenstrom in

Abhängigkeit der Strömungswiderstände des massenstromerzeugenden Modulabschnitts berechnet werden kann, wurde eine Druckrandbedingung eingeführt.

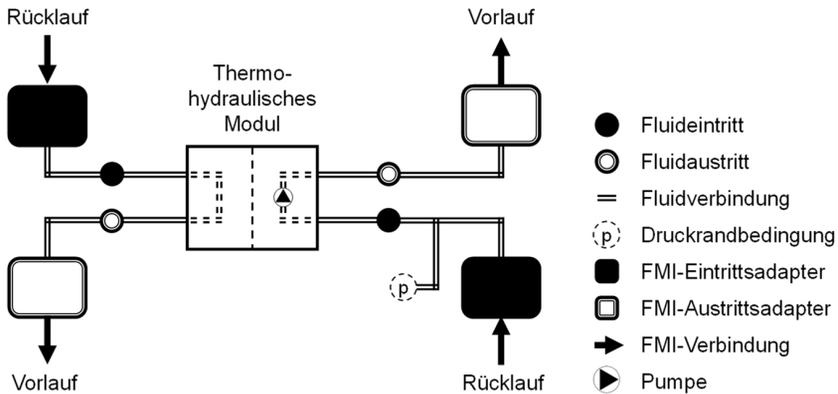


Abbildung 4: FMI-Schnittstellenmodul mit einem integrierten thermohydraulischen Modul (linke Seite: massenstromempfangend; rechte Seite: massenstromerzeugend)

3.3 Schritt 3: Automatisierte Modellerstellung von Systemvarianten als FMU-Verbundmodelle in Python

Zur (teil-)automatisierten Modellerstellung wurde ein strukturbasierter Ansatz implementiert, dessen Ausgangspunkt Strukturdaten über das abzubildende System sind (Bergmann und Straßburger, 2020). Die Systemstruktur wird dabei in Anlehnung an den Standard *System Structure and Parameterization* (SSP) 1.0 beschrieben (Modelica Association, 2019). Sie ergibt sich durch physikalische und regelungstechnische Verbindungen zwischen den Moduleingängen und -ausgängen, welche in einem XML-Datenschema hinterlegt sind. Unter Verwendung der freien Python-Bibliothek *FMPy* (Version 0.2.27) wurde eine Funktion zur automatisierten Kopplung mehrerer FMU-CS-Modelle implementiert. Die entwickelte Funktion liest die Verbindungen gemäß XML-Datenschema aus und fügt die einzelnen Module iterativ zu einem FMU-Verbundmodell zusammen. Ein ähnliches Vorgehen wurde von Noudui (2018) bei der Entwicklung einer Co-Simulationsplattform verfolgt.

4 Virtuelle Tests

Als Fallbeispiel dient ein für Betriebe der metallverarbeitenden Industrie repräsentatives zentrales Kälteversorgungssystem. Es beinhaltet ein Modul eines Kälteverbrauchers, eines thermischen Pufferspeichers zur hydraulischen Entkopplung, ein Kältemaschinenmodul zur Kälteerzeugung sowie ein Nasskühlturmmodul zu dessen Rückkühlung. In Abbildung 5 ist das entsprechende Modelica-Systemmodell schematisch dargestellt. Abgesehen von der Instanziierung thermohydraulischer Module wurden im Modell dort, wo hydraulisch geschlossene Kreise auftreten, vorgegebene Drücke aufgeprägt. Eine weitere Modellvariante

wurde unter Verwendung der in Schritt 2 vorgestellten FMI-Schnittstellenmodule erstellt.

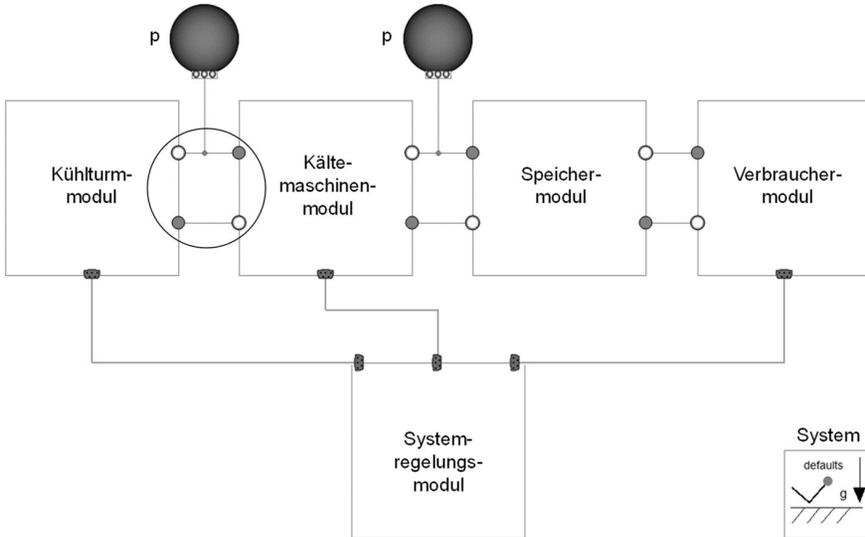


Abbildung 5: Schematische, grafische Modellansicht des manuell erstellten Modelica-Systemmodells

Zur Validierung der FMI-Schnittstellenmodule wurden Simulationsläufe mit beiden Systemmodellen (*MSL FluidPorts* bzw. FMI-Schnittstellenmodule) durchgeführt. Abbildung 6 zeigt die Ergebniswerte für Druck, Massenstrom und Temperatur an den Modulkopplungsstellen zwischen Kühlturmmodul und Kältemaschinenmodul (siehe kreisförmige Markierung in Abb. 5). Bei der Berechnung des Massenstroms im Kühlturmmodul werden ausschließlich die modulinternen Strömungswiderstände berücksichtigt. Im Vergleich zu den *FluidPorts* stellt sich daher über die FMI-Schnittstellen ein höherer Massenstrom ein (der modellierte Strömungswiderstand im Kältemaschinenmodul ist ungleich 0). Im Druckverlauf zeigt sich, dass die Aus- und Eintrittsdrücke der Module durch Verwendung der FMI-CS-Schnittstellen nicht mit den Werten an den *FluidPorts* übereinstimmen. Der Druck am Eintritt des Kältemaschinenmoduls (FMI) errechnet sich aus dem vom Kühlturmmodul empfangenen Massenstrom und dem internen Strömungswiderstand des Kältemaschinenmoduls. Die Werte des Temperaturverlaufes unter Systemanregung weichen bei Verwendung der FMI-CS-Schnittstelle geringfügig von den Werten in den entsprechenden *MSL FluidPorts* ab, wobei die relative, mittlere quadratische Abweichung 0,25 % beträgt.

Die FMI-Schnittstellenmodule zeigen das erwartete Verhalten. Im Vergleich zu den *MSL FluidPorts* ergibt sich ein bleibender Fehler im Druckverlauf. Um diesen Fehler möglichst klein zu halten, müssen die dominierenden hydraulischen Widerstände in den massenstromerzeugenden Modulen berücksichtigt werden. Der Fehler verschwindet für den Fall, dass in einem massenstromempfangenden Modul kein Strömungswiderstand vorherrscht.

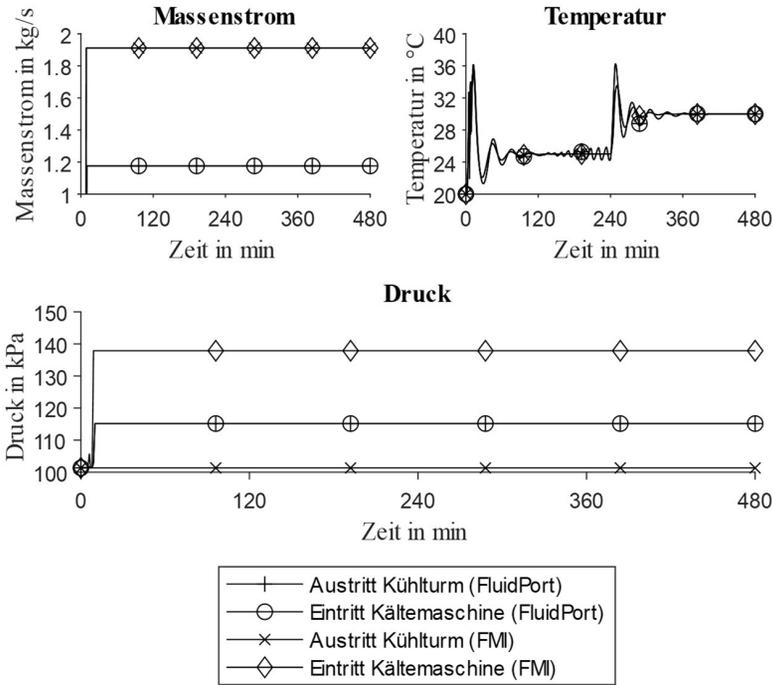


Abbildung 6: Bilanzierung zwischen Eintritt in das Modul zur Kälteerzeugung (Kältemaschine) und Austritt aus dem Modul zur Rückkühlung (Kühlturm)

5 Bewertung des Ansatzes

Der Hauptvorteil des im ersten Schritt entwickelten Technologiemoellbaukastens ergibt sich durch die Zeitersparnis für die Modellierung komplexer Systeme im Rahmen simulationsgestützter Planungsprozesse. Durch die Verwendung von Modulen müssen im Vergleich zur Modellierung auf Komponentenbasis deutlich weniger Verbindungen manuell gesetzt werden. Damit sinkt zugleich die Fehleranfälligkeit bei der Modellierung. Im Rahmen einer Vorplanung können somit schneller unterschiedliche Systemvarianten für weitere Analysen aufgebaut werden. Nachteilig ist, dass insbesondere die Modellierung der Regelung noch weitestgehend vom Anwender selbst implementiert werden muss.

Mit dem Export der Module in das softwareunabhängige FMU-Modellaustauschformat im zweiten Schritt wurde die Grundlage für eine (teil-)automatisierte Modellbildung geschaffen. Einschränkend wurde gezeigt, dass ein Export der entwickelten Modelica-Module in das FMU-Modellaustauschformat nicht in allen Fällen ohne Weiteres möglich ist und aus der CS bekannte numerische Herausforderungen mit sich bringen kann (siehe auch Andersson (2016)). In angrenzenden Untersuchungen werden alternative Möglichkeiten der (teil-)automatisierten Modellierung und Simulation ohne Verwendung der FMI-Technologie, z. B. unter Verwendung des Dymola-Python-Interface, betrachtet.

Weiterführend wurde im dritten Schritt des Ansatzes die (teil-)automatisierte Modellbildung von Systemvarianten mit einem strukturbasierten Ansatz implementiert. Die Simulation des erzeugten FMU-Verbundmodells erfolgt in einer Python-basierten CS. Damit besteht die Perspektive für die anwendungsfreundliche Gestaltung softwareunabhängiger Assistenzsysteme zur energieorientierten Planung industrieller Heiz- und Kühlsysteme.

6 Fazit und Ausblick

Im Beitrag wurde ein Ansatz zur Verringerung von Zeit und Aufwand für die Simulation in der Planung industrieller Heiz- und Kühlsysteme vorgestellt. Zum einen wurde dies durch einen Technologiemoellbaukasten zur modularen Systemmodellbildung erreicht. Zum anderen wurde ein *Proof of Concept* zur (teil-)automatisierten Modellerstellung industrieller Heiz- und Kühlsysteme in einer Python-basierten CS umgesetzt. Für künftige Arbeiten ist eine Integration der Funktionalitäten in ein Web-fähiges Planungstool sowie dessen Anwendung auf reale Planungsbeispiele vorgesehen. Damit sollen die Hemmnisse zur Nutzung der Simulationstechnologie im Kontext der energieorientierten Planung industrieller Heiz- und Kühlsysteme weiter reduziert werden.

Danksagung

Der vorgestellte Ansatz entstand im Kopernikus-Projekt „SynErgie“ (Förderkennzeichen 03SFK3A0-2) unter Förderung durch das Bundesministerium für Bildung und Forschung (BMBF). Modelle hydraulischer Basiskomponenten wurden im Projekt „ETA im Bestand“ (Förderkennzeichen 03EN2048A-I) unter Förderung durch das Bundesministerium für Wirtschaft und Energie (BMWi) entwickelt. Die Autoren bedanken sich für die finanzielle Unterstützung.

Literatur

- Andersson, C.: *Methods and Tools for Co-Simulation of Dynamic Systems with the Functional Mock-up Interface*. Lund, Lund University, Academic Thesis, 2016.
- Bergmann, S.; Straßburger, S.: Automatische Modellgenerierung – Stand, Klassifizierung und ein Anwendungsbeispiel. In: Mayer, G.; Pöge, C.; Spieckermann, S.; Wenzel, S. (Hrsg.): *Ablaufsimulation in der Automobilindustrie*. Berlin, Heidelberg: Springer 2020, S. 333–347.
- Bundesministerium für Wirtschaft und Energie (Hrsg.): *Energieeffizienz in Zahlen. Entwicklungen und Trends in Deutschland 2019, 2019*.
- Flum, D.: *Eine Methode zur modellbasierten Planung energieoptimierter Kälte- und Wärmeversorgungssysteme in Produktionsbetrieben*. Darmstadt, TU Darmstadt, Institut für Produktionsmanagement, Technologie und Werkzeugmaschinen (PTW), Dissertation, 2021.
- Fowler, J.W.; Rose, O.: Grand Challenges in Modeling and Simulation of Complex Manufacturing Systems. *SIMULATION* 80 (2004) 9, S. 469–476.
- Fraunhofer Gesellschaft (Hrsg.): *Energieeffizienz in der Produktion: Untersuchung zum Handlungs- und Forschungsbedarf*. Unter Mitarbeit von Reimund

- Neugebauer, E. Westkämper, F. Klocke, A. Kuhn, M. Schenk, A. Michaelis et al., 2008, München.
- Hirano, Y.; Shimada, S.; Teraoka, Y.; Seya, O.; Ohsumi, Y.; Murakami, S.; Hirono, T.; Sekisue, T.: Initiatives for acausal model connection using FMI in JSAE. In: Proceedings of the 11th International Modelica Conference, Versailles, France, September 21-23, 2015, S. 795–801.
- Kurle, D.: Integrated planning of heat flows in production systems: Sustainable Production, Life Cycle Engineering and Management. Braunschweig, TU Braunschweig, Dissertation, 2018.
- Matthes, P.; Flach, M.; Schmied, S.; Teichmann, J.; Müller, D.; Glück, F.; Oppermann, J.; Mathis, P.; Knapp, T.; Kümpel, A.; Baranski, M.: Zukünftige hydraulische Systeme für die Verteilung von Wärme und Kälte in modernen neuen oder sanierten Gebäuden mit unterschiedlichen Abnehmern und Erzeugern. Gemeinschaftlicher Abschlussbericht über alle Teilprojekte. 2019.
- Modelica Association: System Structure and Parameterization 1.0, 2019.
- Modelica Association: Functional Mock-up Interface for Model Exchange and Co-Simulation 2.0.2, 2020.
- Müller, D.; Both, P.; Maile, T.; Wimmer, R.; Treeck, C.; Nytsch-Geusen, C.: EnEff-BIM: Abschlussbericht, 2017.
- Nouidui, T.S.; Coignard, J.; Gehbauer, C.; Wetter, M.; Joo, J.-Y.; Vrettos, E.: CyDER - An FMI-based Co-Simulation Platform for Distributed Energy Resources, 2018.
- Ogata, Y.; Loyer, B.; Viel, A.: New trends and methods for the co-simulation of strongly coupled systems using the Functional Mock-up Interface 2.0. In: JSAE Annual Congress, Yokohama, May 23, 2014.
- Panten, N.: Deep Reinforcement Learning zur Betriebsoptimierung hybrider industrieller Energienetze. Darmstadt, TU Darmstadt, Institut für Produktionsmanagement, Technologie und Werkzeugmaschinen (PTW), Dissertation, 2019.
- Réhault, N.; Dang, M.; Mitterhofer, M.; Stratbücker, S.; Benndorf, G.: Planung, Auslegung und Betriebsoptimierung von energieeffizienten Neu- und Bestandsbauten durch Modellierung und Simulation auf Basis von Bauwerkinformationsmodellen (EnEff-BIM), 2017.
- Straßburger, S.; Seidel, H.; Schady, R.; Masik, S.: Werkzeuge und Trends der digitalen Fabrikplanung: Analyse der Ergebnisse einer Onlinebefragung. In: Wenzel, S. (Hrsg.): Simulation in Produktion und Logistik 2006. Erlangen: SCS Publ. House 2006.
- VDI 2073-1: Hydraulik in Anlagen der Technischen Gebäudeausrüstung. Hydraulische Schaltungen. Berlin: Beuth 2014.
- VDI 4499: Digitale Fabrik. Blatt 1 - Grundlagen. Berlin: Beuth 2008.
- VDI 6018: Kälteversorgung in der technischen Gebäudeausrüstung in der technischen Gebäudeausrüstung Planung, Bau, Betrieb. Berlin: Beuth 2018.
- Wetter, M.; Fuchs, M.; Nouidui, T.S.: Design Choices for Thermofluid Flow Components and Systems that are Exported as Functional Mockup Units. In: Proceedings of the 11th International Modelica Conference, Versailles, France, September 21-23, 2015, S. 31–41.
- Wischhusen, S.: Dynamische Simulation zur Bewertung von komplexen Energiesystemen. Hamburg, TU Hamburg-Harburg, Dissertation, 2005.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Simulationsbasierte Optimierung zur Konfiguration von ökoeffizienten Supply Chains

Simulation-based optimization for eco-efficient supply chain configuration

Lucas Schreiber, Fraunhofer Institut für Materialfluss und Logistik Dortmund,
Dortmund (Germany), lucas.schreiber@iml.fraunhofer.de

Christian Niehus, TU Dortmund, Dortmund (Germany), christian.niehus@tu-
dortmund.de

Nikolas Moroff, Fraunhofer Institut für Materialfluss und Logistik Dortmund,
Dortmund (Germany), nikolas.moroff@iml.fraunhofer.de

Abstract: The efficient design of supply chains incorporating ecological objectives is a strategic task that is increasingly attracting the attention of companies. This paper introduces a simulation-based optimization approach to eco-efficiently orchestrate a supply chain with a target system consisting of three sub-targets: Costs, energy-efficiency and service level. Regarding a use case from the steel processing industry, an event-discrete simulation model of the corresponding supply chain was configured. By interfacing the simulation model with a Nondominated Sorting Genetic Algorithm, new configurations of decision variables are generated after a set of simulation runs. The evaluation of the experiments and the resulting pareto sets led to the identification of promising eco-efficient configurations and the derivation of corresponding decision variable assignments for the use case which consist of material allocation, reorder point and replenishment level.

1 Einleitung

Eine weltweit steigende Nachfrage an Energie und Rohstoffen stellt die produzierende Wirtschaft vor große Herausforderungen. Durch die fortlaufende Verknappung der Verfügbarkeit von fossilen Energieträgern kommt der Nachhaltigkeit im industriellen Sektor eine steigende Bedeutung zu. Gleichwohl ist parallel zu dieser Entwicklung laut dem deutschen Bundesministerium für Wirtschaft und Energie ein deutlich steigender Energieverbrauch im industriellen Sektor zu erkennen (Bundesministerium für Wirtschaft und Energie (BMWi), 2019). Um Unternehmen befähigen zu können, diese konträren Entwicklungen zu adressieren, sind Anstrengungen in der Schnittstelle zwischen Forschung und Anwendung notwendig.

Im Rahmen des Supply Chain Managements werden zunehmend ökologische Zielgrößen in Unternehmensentscheidungen integriert. Zur Konfiguration von nachhaltigen Supply Chains stellen ökoeffektive Herangehensweisen eine entscheidende Strategie dar (Braungart et al., 2007). Dabei sollen ökologische Zielgrößen simultan zu einer Beibehaltung oder Steigerung der Wirtschaftlichkeit in das Zielsystem des Wertschöpfungsnetzwerks eingebunden werden. Supply Chains sind jedoch hochgradig dynamische und komplexe Systeme mit mannigfaltigen dynamischen Wirkungszusammenhängen. Es bedarf demnach effizienten methodischen Werkzeugen, welche die Zusammenhänge der Supply Chain abbilden und bewerten sowie eine Verbesserung von Supply-Chain-Parameterkonfigurationen erzielen können. Für das digitale Rekonstruieren und bewerten unterschiedlicher Steuerungslogiken und Parametersetzungen einzelner Entitäten ist die ereignisdiskrete Simulation ein etabliertes Instrumentarium. Für die Optimierung von Wertschöpfungsnetzwerken bieten sich unter anderem lineare Optimierungsmodelle oder auch Metaheuristiken an (Schreiber, 2019). Aufgrund der beschriebenen Komplexität und einem damit einhergehend großem Suchraum an potentiellen Gestaltungsoptionen für eine Wertschöpfungskette, ist es schwierig, exakte Lösungen für eine gegebene modellierte Problemstellung zu generieren. Zudem gestaltet sich eine rein mathematische Formulierung des Modells aufgrund der Vielzahl an Objekten, Wirkungszusammenhängen und stochastischen Unsicherheiten oftmals als Herausforderung.

Eine Kombination aus Metaheuristik und Simulation bietet sich an, um die Vorteile beider Werkzeuge vollends auszuschöpfen (Baudach et al., 2013). Diese sogenannte simulationsbasierte Optimierung kann laut VDI auf vier unterschiedliche Arten implementiert werden (VDI 3633 Blatt 12, 2020). In dieser Ausarbeitung soll eine integrative Kopplung als Entscheidungsunterstützung implementiert werden, bei welcher die (multikritiellen) Simulationsergebnisse als Zielfunktionswerte für einen genetischen Algorithmus dienen. Dies entspricht gemäß VDI einem Ansatz der „Kategorie D“. Das dreidimensionale Zielgrößensystem besteht aus den Kriterien Kosten, Energieverbrauch und Servicegrad. Mithilfe des genetischen Algorithmus werden nach einem Satz an Simulationsläufen neue Konfigurationen der Supply-Chain-Input-Parameter für das Simulationsmodell generiert. Für diese Ausarbeitung werden als genetische Algorithmen sogenannte „Nondominated Sorting Genetic Algorithms (NSGA)“ in Form des NSGA-II und des NSGA-III Algorithmus verwendet. Durch eine implementierte Schnittstelle werden die neu generierten Konfigurationen automatisiert und iterativ im darauffolgenden Simulationslauf getestet. Aufbauend auf dem Stand der Technik und der Abgrenzung des vorgestellten Ansatzes zu bisherigen Forschungsarbeiten, wird ein spezifischer demonstrativer Anwendungsfall aus der Stahlbearbeitungsbranche vorgestellt. Basierend auf einer Sensitivitätsanalyse zur Eingrenzung des Untersuchungsraums zum initialen Ausschluss von defizitären Lösungen, wird die entwickelte Methodik experimentell erprobt und die Ergebnisse ausgewertet.

2 Simulationsbasierte Optimierung für Supply Chains

Die Nutzung der simulationsbasierten Optimierung für die Orchestrierung und Konfiguration von Supply Chains ist ein breit gefächertes Forschungsfeld. Dieses Kapitel widmet sich der Betrachtung von aktuellen Lösungsansätzen der

simulationsbasierten Optimierung, insbesondere für die strategische und taktische Ausgestaltung von Prozessen in Wertschöpfungsnetzwerken.

Eine Methodik für eine simulationsbasierte Optimierung mit einem NSGA-II Algorithmus der Supply Chain eines Stahlhandelsunternehmens wurde von Rabe et al. vorgestellt, in welchem benutzerdefinierte Maßnahmenpläne konzipierbar sind, welche unter anderem Bestandsparameter und Materialumschichtungen adressieren (Rabe et al., 2017). Das Zielsystem beinhaltet die beiden Zielgrößen Kosten und Servicegrad, integriert jedoch keine ökologischen Kennzahlen und nutzt neben dem NSGA-II als alternativen Algorithmus einen Ansatz des Deep Reinforcement Learning (DRL) anstatt des NSGA-III. Ein weiterer Ansatz der „Kategorie D“ von Benyoucef und Xie fokussiert sich auf einen Anwendungsfall in der Automobilindustrie, in welchem ebenfalls ein zweidimensionales Zielsystem, bestehend aus Kosten und Servicegrad betrachtet wurde. Auch hier wurde der Untersuchungsraum eingeschränkt, indem konfigurierbare Parameter wie die Bestellmenge und der Meldebestand mit Ober- und Untergrenzen in der potentiellen Belegung begrenzt wurden (Benyoucef und Xie, 2011). Es existieren weitere Ansätze, welche in zweidimensionalen Zielsystemen Emissionswerte einbeziehen (Keramydas et al., 2017) oder Ausarbeitungen, welche andere Metaheuristiken, wie eine Particle Swarm Optimization (PSO) implementieren (Salem und Haouari, 2017).

Eine aktuelle und ausführliche Zusammenfassung der aktuellen Forschungslandschaft der simulationsbasierten Optimierung für Wertschöpfungsnetzwerke kann in den Ausarbeitungen von Tordecilla et al. nachvollzogen werden (Tordecilla et al., 2021). Daraus wird insbesondere deutlich, dass das Zielgrößensystem des in diesem Beitrag betrachteten Anwendungsfalls mit den drei Dimensionen Kosten, Servicegrad und Energieeffizienz eine Abgrenzung zu bisherigen Ansätzen darstellt.

3 Beschreibung des Anwendungsfalls

Der Anwendungsfall für diese Ausarbeitung sei eine Supply Chain eines Stahlbearbeitungsdienstleisters mit mehreren Distributions- bzw. Bearbeitungszentren. Der Dienstleister bietet Serviceleistungen an, bei welchen Standardgrößen von Stahlprodukten auf kundenspezifische Abmaße und Formen durch unterschiedliche Bearbeitungsmaschinen transformiert werden. Dazu steht ein Pool an Bearbeitungsmaschinen an den verschiedenen Standorten zur Verfügung. Die Lieferkette besteht aus 5 Distributionszentren sowie 25 Kundenregionen und ist in Abbildung 1 vereinfacht dargestellt.

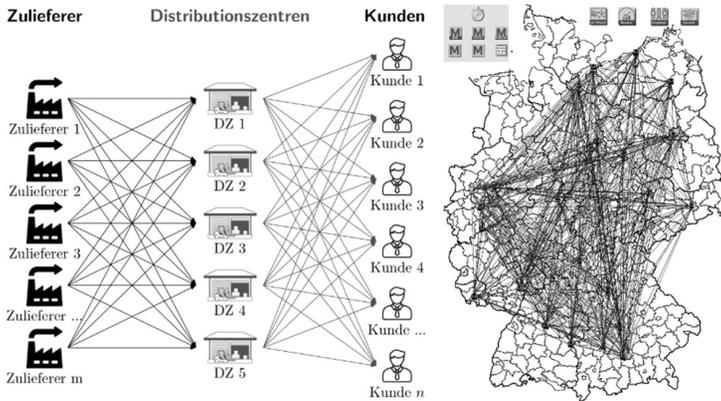


Abbildung 1: Darstellung der Lieferkette des Dienstleisters

3.1 Beschaffung, Produktion und Distribution im Anwendungsfall

Folgend werden die grundlegenden Prozesse der Beschaffung, Produktion und Distribution innerhalb der Supply Chain des Dienstleisters aufgezeigt, welche in die Mechanismen des Simulationsmodells integriert sind.

Eingehende Kundenaufträge werden zunächst in Auftragspositionen aufgeteilt. Diese sind basierend auf den Lagerbeständen in den Distributionszentren dahingehend zu prüfen, ob eine Annahme des Auftrages erfolgen kann. Aufgrund einer individuellen Materialallokation, können Szenarien existieren, in denen nicht alle Standardgrößen in allen Distributionszentren bevorratet werden. Für den Fall, dass mehrere Distributionszentren das Material eines Auftrags in ausreichender Quantität lagern, wird eine entfernungs-basierte Zuweisung getroffen, sodass jenes mengenmäßig befähigte Distributionszentrum den Auftrag annimmt, welches geographisch die geringste Entfernung zur Kundenregion des Auftraggebers aufweist. Für den Fall, dass nicht ausreichend Material in der Lieferkette vorhanden ist, wird der Auftrag abgelehnt, was entsprechend im Servicegrad zu beachten ist. Die Überwachung der Lagerbestände der Distributionszentren stellt demnach einen integralen Bestandteil dar. Der Dienstleister operiert mit einer Bestellpunkt-Lagerniveau-Strategie, genannt (s,S)-Bestellpolitik (Tempelmeier, 2012). Unterschreitet im Rahmen dieser Bestellpolitik ein Lagerbestand einen definierten Meldebestand s , stößt der Dienstleister einen Bestellauftrag beim Lieferanten an, welcher neue Standardgrößen auf einen individuell pro Produkt definierten maximalen Richtbestand S auffüllt.

Der Produktionsprozess stellt die kundenindividuelle Anarbeitung von vorhandenen Standardgrößen oder Restgrößen aus bisherigen Bearbeitungsprozessen dar. Dabei wird unterschieden zwischen dem Sägen von Profilen und Trägern sowie dem Schneiden von Blechen (vgl. Abbildung 2). Die Zeiten der jeweiligen Säge- und Schneidprozesse sind abhängig von dem zu durchtrennenden Querschnitt des Materials als auch der Säge- bzw. Schneidgeschwindigkeit der Maschine. Zudem werden die Vorschub-Geschwindigkeiten sowie Be- und Entladezeiten der Maschinen in die Gesamtbearbeitungszeit eingebunden. Jedem Distributionszentrum steht ein individueller Pool an Maschinen mit unterschiedlichen Attributen bzgl.

Bearbeitungsgeschwindigkeit, Befähigungen zur Bearbeitung spezifischer Materialien und Energieverbräuchen zur Verfügung.

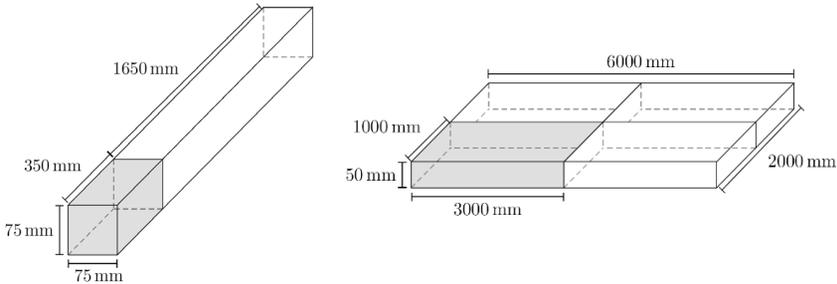


Abbildung 2: Beispielhafte Bearbeitungsvorgänge des Dienstleisters

Distributionsseitig erfolgt die Auftragskommissionierung und die Verladung basierend auf dem Gewicht und Volumen der fertig bearbeiteten Produkte und der maximalen Mengen- und Volumenkapazität der Transport-LKW. Diese versorgen je nach Dringlichkeit des Kundenwunschtermins eines Auftrags sowie basierend auf einer Routenminimierung. Neben betriebseigenen LKW steht zudem die Möglichkeit des Abrufs eines Spediteurs zur Verfügung.

3.2 Zielsystem des Anwendungsfalls

Ein wichtiger Aspekt ist die Befähigung der Quantifizierung der einzelnen Komponenten des Zielsystems bestehend aus Kosten, Energieverbrauch und Servicegrad.

3.2.1 Kosten

Als erste Zielfunktion des Zielsystems werden die Gesamtkosten berechnet, welche aus der Summe der Transportkosten, der Bestellkosten, der Bestandskosten sowie der Fehlmengenkosten bestehen. Die Zielfunktion ergibt sich damit zu:

$$ZF_1 = C_t + C_o + C_{inv} + C_{bo} \quad (1)$$

mit:

C_t : Gesamte Transportkosten [in €]

C_o : Gesamte Bestellkosten [in €]

C_{inv} : Gesamte Bestandskosten [in €]

C_{bo} : Fehlmengenkosten [in €]

Die Kostenarten Transportkosten und Bestellkosten werden darüber hinaus in einen fixen und einen variablen Anteil aufgeteilt. Fixe Transportkosten fallen einmalig für einen Transport an. Dabei wird zwischen den Fixkosten für einen Spediteur und den Fixkosten für den Transport eines betriebseigenen LKW unterschieden. Dies gilt in Analogie für die variablen Transportkosten. Für den Spediteur wird ein abweichender

Kostensatz im Vergleich zu den betriebseigenen LKWs verwendet. Die variablen Kosten berechnen sich auf Stundenbasis in Abhängigkeit der Dauer einer Tour. Diese resultiert aus der Gesamtstrecke der Tour dividiert durch die Durchschnittsgeschwindigkeit eines LKW. Bestellkosten fallen immer an, sofern ein Lagerbestand einen vorgegebenen Meldebestand unterschreitet und daraufhin eine Bestellung des jeweiligen Materials beim Lieferanten ausgelöst wird. Die variablen Bestellkosten einer Bestellung ergeben sich aus der Multiplikation der Bestellmenge mit einem variablen Bestellkostensatz. Zusätzlich wird ein Fixkostensatz je Bestellung aufaddiert. Bestandskosten sind stets variable Kosten. Für jede Lagereinheit stellt die individuelle Lagerdauer die Grundlage zur Berechnung der Bestandskosten dar. Diese wird zur Ermittlung der Bestandskosten mit dem täglichen Lagerhaltungskostensatz multipliziert. Letztlich ergeben sich die Fehlmengenkosten aus der Multiplikation der Mengeneinheiten der Auftragspositionen, die nicht erfüllt werden konnten mit dem entsprechenden Fehlmengenkostensatz. Die Gesamtkosten sind in diesem Optimierungsproblem zu minimieren.

3.2.2 Energieverbrauch

Im Rahmen der zweiten Zielfunktion wird die Minimierung des Leistungsverbrauch der Säge- und Schneidemaschinen in der Einheit Megajoule [MJ] angestrebt. Die Bearbeitungszeiten, Vorschübe und technischen Gegebenheiten der einzelnen Maschinen spielen bei der Berechnung eine essentielle Rolle. Um den energetischen Leistungsverbrauch einer Säge zu ermitteln, wird angenommen, dass Energie nur durch eine Säge verbraucht wird, wenn sich ein Werkstück auf dieser im Vorschub befindet oder die Säge mit der Bearbeitung eines Werkstücks beschäftigt ist. Für die betrachteten Maschinen sind die Antriebsleistungen dieser beiden Aktoren in der Einheit Kilowatt [kW] bekannt. Anhand der Prozesszeiten für den Vorschub sowie der Zerspanung einer Auftragsposition und den Antriebsleistungen beider Aktoren einer spezifischen Säge lässt sich ermitteln, wie viel Kilowatt für die Bearbeitung des jeweiligen Auftrags verbraucht wird. Durch die Multiplikation mit den jeweiligen Bearbeitungszeiten kann demnach der Energieverbrauch in Kilowattstunden ermittelt werden. Für eine Bearbeitung eines Produkts einer Auftragsposition ergibt sich je nach genutzter Bearbeitungsmaschine die folgende Berechnungsvorschrift für den Energieverbrauch:

$$P_{AP} = \frac{t_{GesVor}}{3600 \frac{s}{h}} \times P_{Vor} + \frac{t_{GesSäg}}{3600 \frac{s}{h}} \times P_{Säg} \quad (2)$$

mit:

- P_{AP} : Leistungsverbrauch einer Säge der aus der Bearbeitung der Auftragsposition resultiert [in kWh]
- $t_{GesSäg}$: Gesamtsägezeit einer Auftragsposition [in s]
- t_{GesVor} : Zeit für den Vorschub aller Werkstücke einer Auftragsposition durch die Sägemaschine [in s]
- P_{Vor} : Antriebsleistung des Aktors zur Steuerung des Vorschubs der Werkstücke [in kW]

$P_{Säg}$: Antriebsleistung des Aktors für den Sägebetrieb [in kW]

Durch die Multiplikation mit dem Faktor 3,6 MJ/kWh erfolgt die Umrechnung von Kilowattstunde in Megajoule. Da ein Auftrag in die einzelnen Positionen aufgeteilt wird und diese eventuell in verschiedenen Distributionszentren bearbeitet werden, muss zur Ermittlung des Energieverbrauchs für einen Auftrag, die Summe der Leistungsverbräuche über Positionen ($i=1, \dots, m$) erfolgen:

$$ZF_2 = \sum_{i=1}^m P_{AP_i} \times 3,6 \text{ MJ/kWh} \quad (3)$$

3.2.3 Servicegrad

Die dritte Dimension des Zielsystems stellt der erreichte Servicegrad bzw. Lieferbereitschaftsgrad dar. Dieser ist eine logistische Kennzahl, die Auskunft über die durchschnittliche Lieferfähigkeit eines Unternehmens innerhalb eines betrachteten Zeitraums gibt. Für die Berechnung des Lieferbereitschaftsgrades existieren prinzipiell verschiedene Berechnungsgrundlagen. Im Rahmen des Anwendungsfalls wird der sogenannte β -Servicegrad zur Quantifizierung der Lieferfähigkeit herangezogen. Aus praktischer Sicht wird dieser als bevorzugtes Leistungskriterium herangezogen, weil die Höhe einer Fehlmenge in der Berechnungsgrundlage inkludiert ist (Tempelmeier, 2012). Mathematisch ist dieser wie folgt definiert:

$$\beta = \frac{\text{Erfüllte Periodennachfragemenge}}{\text{Periodennachfragemenge}} \quad (4)$$

Die Berechnung des β -Servicegrad impliziert, dass die beste Ausprägung bei einem Wert von $\beta=1$ erzielt wird, da in diesem Fall jegliche Periodennachfrage erfüllt werden konnte. Im Sinne einer Optimierung ist der Wert entsprechend grundsätzlich zu maximieren. Da die anderen beiden Zielfunktionen zu minimieren sind und eine Kombination aus Minimierung und Maximierung mit einer multikriteriellen Metaheuristik schwieriger realisierbar ist, wird im Anwendungsfall die Gegenwahrscheinlichkeit des Lieferbereitschaftsgrades minimiert. Mathematisch ergibt sich die dritte Zielfunktion somit zu:

$$ZF_3 = 1 - \beta \quad (5)$$

4 Entscheidungsvariablen und Untersuchungsraum

Das Zielsystem des Unternehmens besteht wie im vorherigen Kapitel beschrieben aus den drei Komponenten Kosten, Energieverbrauch und Servicegrad. Um dieses multikriterielle Zielsystem zu optimieren, werden Variationen von logistischen Parametern untersucht, welche die Entscheidungsvariablen des Systems darstellen. Die zu untersuchenden Parameter sind der Meldebestand s je Produkt p in Distributionszentrum j , der Richtbestand S je Produkt p in Distributionszentrum j sowie die Allokation von Produkten auf bestimmte Distributionszentren in Form von Materialumschichtungsstrategien.

S_j^p : Richtbestand für Produkt p in Distributionszentrum j in Tagen $\in \mathbb{Z}^+$

s_j^p : Meldebestand für Produkt p in Distributionszentrum j in Tagen $\in \mathbb{Z}^+$

l_j^p : Binäre Materialallokation für Produkt p in Distributionszentrum $j \in [0,1]$

Um das initial aufgezeigte Problem des großen Lösungsraumes zu adressieren, wurde vor der Kopplung der Simulation mit der Metaheuristik eine Sensitivitätsanalyse der Supply Chain auf verschiedene Bestandsparameter durchgeführt, um den Suchraum auf vielversprechende Kombinationsmöglichkeiten aus Meldebestand und Richtbestand zu beschränken. Mit dieser Analysetechnik wird die Empfindlichkeit der Zielfunktionen in Abhängigkeit der Eingabeparameter ermittelt (Rabe et al., 2008). Dazu wurden die durchschnittlichen Nachfragemengen pro Produkt in Bestandsreichweiten überführt und alle potentiellen Kombinationsmöglichkeiten bis zu einer definierten Obergrenze iterativ in automatisch angestoßenen Simulationsläufen gleichverteilt auf die Distributionszentren getestet. Abbildung 3 zeigt die Gesamtkosten des Netzwerks in Abhängigkeit der Parametersetzungen.

Es ist ersichtlich, dass die Kostenfunktion zumindest für die getesteten Parameter eine konvexe Form annimmt, was eine gute Implikation dafür ist, dass ein gefundenes lokales Minimum auch gleichzeitig das globale Minimum darstellt.

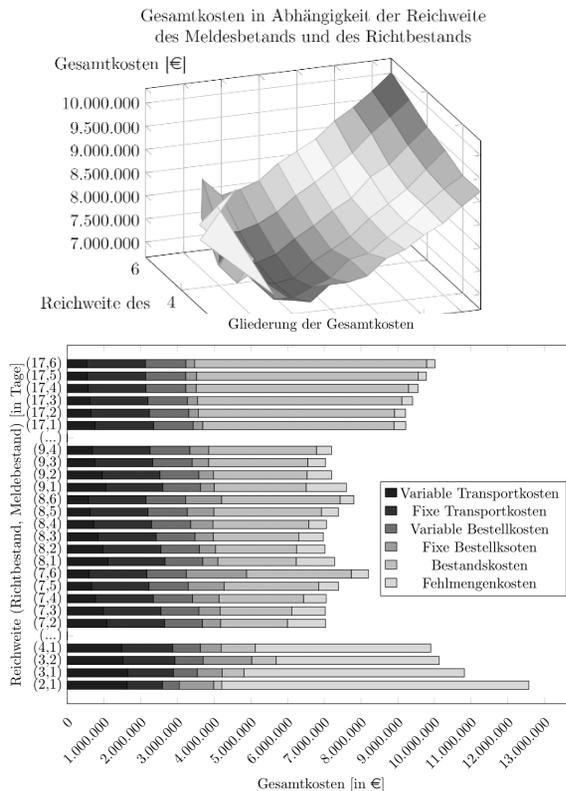


Abbildung 3: Sensitivitätsanalyse der Gesamtkosten des Netzwerks in Abhängigkeit der Bestandsparameter

Um eine bessere Interpretation der Ergebnisse hinsichtlich der erreichten Kosten zu erzielen, sind in der in Abbildung 3 zudem die subsummierten Gesamtkosten aufgeschlüsselt. Dadurch kann aufgezeigt werden, dass die signifikantesten Unterschiede bei den Fehlmengenkosten und den Bestandskosten auftreten. Diese Erkenntnis lässt sich auf einen klassischen Trade-Off Effekt in der Lagerhaltung zurückführen, in welchem hohe Bestände zu hohen Kapitalbindungskosten führen, wohingegen geringe Bestände zu Fehlmengenkosten führen. Bei geringen Melde- und Richtbeständen ist das Risiko entsprechend höher, eine hohe Summe an Fehlmengenkosten zu verbuchen. Umgekehrt führt eine Minimierung dieses Risikos zu hohen Bestandskosten.

Diese Sensitivitätsanalysen wurden zudem für die Zielgrößen Energieverbrauch und Servicegrad durchgeführt. Aus diesen Untersuchungen wurde ersichtlich, dass sich der Servicegrad schon bei recht geringen Reichweiten bei einem Wert nahe von 100% einpendelt, was einhergehend mit geringen Fehlmengenkosten ist. Die Güte der Werte für die Energieverbräuche korrelieren stark mit denen für den Servicegrad. Dieser Effekt lässt sich dadurch erklären, dass für die Sensitivitätsanalyse vorerst nur die ersten beiden Entscheidungsvariablen rudimentär untersucht wurden und keine Materialumschichtungen integriert sind. Es können entsprechend nur Produkte bearbeitet werden und entsprechend Energie verbraucht werden, wenn Produkte im Bestand verfügbar sind. Durch diese Sensitivitätsanalyse der ersten beiden Entscheidungsvariablen konnte bereits ein entsprechender Metaheuristik-Suchraum für die Bestandsparameter definiert werden, welcher defizitäre Lösungskandidaten einschränkt. Die Variation der Materialallokation auf die Distributionszentren ist im weiteren Verlauf näher zu untersuchen, da darin ein großer Hebel für die Minimierung der Energieverbräuche vermutet wird. Die Materialumschichtungen wurden ebenfalls mit einer Logik zur Reduzierung der Rechenintensität versehen. Zur Untersuchung von verschiedenen Zuordnungen von Produkt und Distributionszentrum kann vor dem Start eines Optimierungslaufes ein Anteil an durchzuführenden Materialumschichtungen definiert werden. In der Initialpopulation und mit jeder Mutation des genetischen Algorithmus wird entsprechend dieses Anteils ein Material zufällig auf eines oder mehrere Distributionszentren umverteilt.

5 Beschreibung der Methodik und der Werkzeuge

Anknüpfend an die Definition der Entscheidungsvariablen und der Zielfunktion ist die methodische Konfiguration auszugestalten. Als Metaheuristiken werden zwei sogenannte „Nondominated Sorting Genetic Algorithms“ (NSGA-II (Deb et al., 2002) und NSGA-III (Deb und Jain, 2014)) in verschiedenen Experimenten als Metaheuristiken genutzt, welche auf dem Prinzip der genetischen Algorithmen basieren. Kernbausteine sind danach die Selektion, die Rekombination sowie die Mutation. Die beiden Algorithmen eignen sich im Gegensatz zu klassischen evolutionären Verfahren insbesondere für die Anwendung bei diskreten multikritiellen Optimierungsproblemen und sind entsprechend mit Mechanismen ausgestattet, welche zur Ermittlung einer Pareto-Rangfolge befähigen. Obwohl der NSGA-III im Gegensatz zum NSGA-II mit einer Lotnormalisierung zur Distanzberechnung und einem Niching-Mechanismus ausgestattet ist, liefert dieser nicht für jeden Anwendungsfall bessere Ergebnisse. Für eine genauere Erläuterung der Algorithmen sei auf die entsprechende Literatur verwiesen.

Die Metaheuristiken werden in der Programmiersprache Python implementiert und mithilfe der *Component Object Model (COM)*-Schnittstelle an die Software *Tecnomatix Plant Simulation* von *Siemens* gekoppelt, in welcher das Simulationsmodell erstellt wurde. Die Schnittstelle ermöglicht es, *Tecnomatix Plant Simulation* von anderen Programmen aus zu steuern, sodass diese bspw. Simulationsläufe starten, Modellparameter verändern und Ergebnisse erfassen können. Nach der Generierung von Lösungskandidaten in der Metaheuristik in Python, werden diese an das Simulationsmodell übergeben und daraufhin aus Python ein Simulationslauf gestartet. Die Terminierung der Simulation wird über ein Ereignis in der COM-Schnittstelle mitgeteilt, worauf die Ergebnisse des Simulationsmodell in Python übernommen werden. Um die Rechengeschwindigkeit zu erhöhen, wurde zusätzlich eine Parallelisierung von Simulationsläufen auf mehrere Kerne der CPU realisiert. Abbildung 4 stellt den Prozessablauf der Simulationsoptimierung grob anhand eines Prozessdiagramms dar.

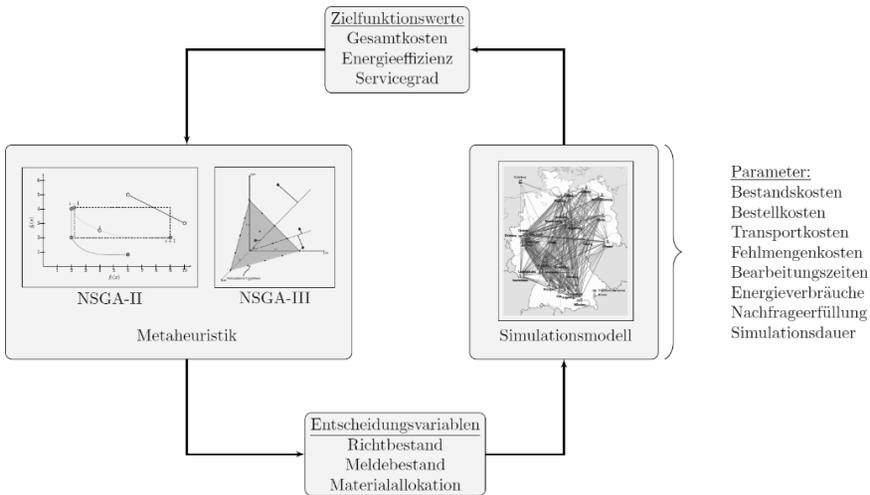


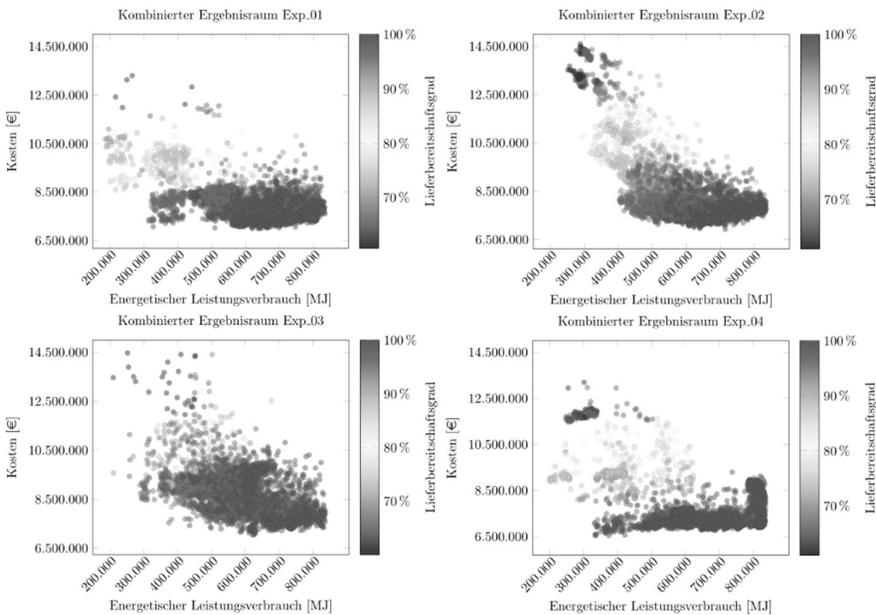
Abbildung 4: Prozessschaubild der Simulationsoptimierung

6 Ergebnisse der Experimente

Für den Vergleich der verschiedenen Algorithmen mit variierenden Konfigurationen wurde ein Experimentierplan konzipiert. Dieser besteht aus vier Experimenten, welche die Güte der verschiedenen Algorithmen sowie die unterschiedlichen Anteile an Materialumschichtungen (MU) bei gegebener Populationsgröße (PG) untersuchen. Der Experimentierplan ist Tabelle 1 zu entnehmen. Eine Metaheuristik terminiert in dieser Simulationsstudie nach 50 Generationen. Daraus resultieren 2.500 Lösungskandidaten aus dem Durchlauf einer Metaheuristik. Um stochastische Signifikanz zu berücksichtigen werden je Experiment zwei Replikationen durchgeführt. Somit stehen 5.000 Lösungskandidaten für die Auswertung eines Experiments zur Verfügung. Der Simulationszeitraum beträgt je Simulationslauf ein Jahr im Anwendungsfall. Die Ergebnisse als Streudiagramm können der Abbildung 5 entnommen werden.

Tabelle 1: Experimentierplan

Experiment	ID	Algorithmus	MU	PG	Generationen
Experiment 1	EXP 01	NSGA-II	33 %	50	50
Experiment 2	EXP 02	NSGA-II	0 %	50	50
Experiment 3	EXP 03	NSGA-III	33 %	50	50
Experiment 4	EXP 04	NSGA-II	66 %	50	50


Abbildung 5: Streudiagramme der Zielfunktionswerte

Es zeigt sich, dass vor allem in den Experimenten mit geringem Anteil an Materialumschichtungen (Experimente 1-3) annähernd lineare Abhängigkeiten zwischen den Zielgrößen zu beobachten sind. Diese Erkenntnis lässt sich darauf zurückführen, dass ein geringerer Servicegrad damit einhergeht, dass die gewünschte Produktionsmenge nicht erreicht wurde und dementsprechend die Maschinen weniger Energie durch weniger Produktionsgesamtmenge verbraucht haben. Da in die Kostenfunktion neben den Transport-, Bestell- und Bestandskosten auch die Fehlmengenkosten integriert sind, entstehen bei solchen Lösungskandidaten entsprechend hohe Kosten durch schlechten Lieferservice bei geringem Energieverbrauch. Bei einem höheren Anteil an Materialumschichtungen können vielversprechende Lösungskandidaten identifiziert werden, welche geringe Kosten und Energieverbräuche bei einem guten Servicelevel erzielen. Dies ist möglich, da die Anlagen der Distributionszentren unterschiedlich energieeffizient sind. Es werden demnach Materialallokationen mit effizienten Bestandsparametern identifiziert,

sodass die Materialien einerseits kundennah als auch an jeweils energieeffizienten Standorten hergestellt werden.

Weiterhin ist ersichtlich, dass im Cluster für das Experiment 3 eine große Streuung auftritt. Diese lässt sich durch einen abweichenden Selektionsoperator des NSGA-III im Vergleich zum NSGA-II erklären. Eine noch dezidiertere Möglichkeit zur Interpretation der Ergebnisse liefern die Pareto-Mengen der Experimente in Abbildung 6.

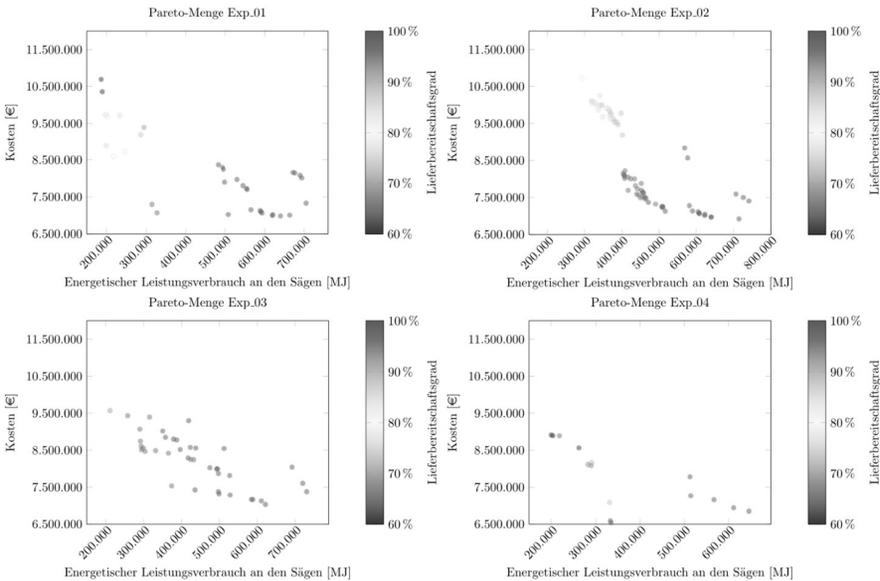


Abbildung 6: Pareto-Mengen der Experimente

Die tendenziell vielversprechendste Pareto-Menge konnte im Rahmen von Experiment 4 generiert werden. Im Vergleich zu den Pareto-Mengen aus den übrigen Experimenten weisen mehrere Lösungskandidaten in Experiment 4 sowohl bessere Lieferbereitschaftsgrade, Energieverbräuche als auch bessere Gesamtkosten auf. Eine Vielzahl an Lösungskandidaten aus der Pareto-Menge aus Experiment 4 würden große Teile der restlichen Pareto-Mengen dominieren.

Eine beste Lösung, welche in allen drei Zielfunktionsdimensionen dominiert, wurde erwartungsgemäß aufgrund der diversen Trade-Off Effekte innerhalb einer Supply Chain nicht gefunden. Dennoch können aus den generierten Pareto-Mengen je nach individueller Gewichtung des Zielsystems im Anwendungsfall Handlungsempfehlungen abgeleitet werden. Experiment 4 lieferte Lösungskandidaten mit Gesamtkosten von ca. 6,5 Mio. Euro, einem Leistungsverbrauch von ca. 330.000 MJ und einem Lieferbereitschaftsgrade von ca. 99%. Es existieren zwar Lösungen mit besseren Energieverbräuchen, jedoch verschlechtern sich bei der Auswahl dieser Lösungen der Lieferbereitschaftsgrad als auch die Gesamtkosten signifikant. Sofern die individuellen Präferenzen im Anwendungsfall den Energieverbrauch und eine perfekte Lieferbereitschaft nicht

übermäßig stark priorisieren, ist diese zugrundeliegende Kombination aus Entscheidungsvariablen eine äußerst vielversprechende ökoeffiziente Konfiguration für den vorliegenden Anwendungsfall.

7 Zusammenfassung und Ausblick

Die in diesem Beitrag vorgestellte simulationsbasierte Optimierung zur Konfiguration von ökoeffizienten Supply Chains stellt ein performantes Werkzeug zur Generierung von zielsystemspezifischen Pareto-Mengen dar. Die Einbindung der Kennzahl Energieeffizienz in ein dreidimensionales Zielsystem angereichert mit den Dimensionen Kosten und Servicegrad stellt zudem in Kombination mit den gewählten Algorithmen und Simulationswerkzeugen ein innovatives Vorgehen dar, welches sich vom bisherigen Stand der Technik abgrenzt. Im Rahmen der Auswertung der Experimente konnten vielversprechende ökoeffiziente Konfigurationen ermittelt werden und daraus entsprechende Belegungen der Entscheidungsvariablen für den Anwendungsfall hinsichtlich Materialallokation, Meldebestand und Richtbestand abgeleitet werden. Die Experimente verdeutlichten zudem, dass in den Betrachtungen für diesen Anwendungsfall, der NSGA-II bessere Lösungskandidaten ermitteln konnte, als der NSGA-III.

Als weiterer Forschungsbedarf kann abgeleitet werden, dass weitere Methoden wie das Reinforcement Learning anstelle der in dieser Ausarbeitung gewählten Metaheuristiken zu untersuchen sind. Zudem könnte die Wahl eines anderen Simulationsinstrumentariums (bspw. SimPy), welches komplett auf Animationen verzichtet, eine performantere simulationsbasierte Optimierung ermöglichen, was ein zügigeres Generieren von potentiellen Lösungskandidaten ermöglicht. Weiterhin ist es möglich zusätzliche Komponenten an Energieverbräuchen, bspw. verursacht durch Transporte, in das Zielsystem zu integrieren.

Danksagung

Dieser Beitrag entstand im Rahmen des Forschungsprojekts „E²-Design“, gefördert durch das Bundesministerium für Wirtschaft und Energie (FKZ 03ET1558A).

Literaturverzeichnis

- Baudach, J.; Voll, R.; Eufinger, L.; Meier, F.; Sender, J.; Goedicke, I.; Thaller, C.: Modellentwicklung. In: Clausen, U.; Gegier, C. (Hrsg.): Verkehrs- und Transportlogistik. Berlin, Heidelberg: Springer Berlin Heidelberg 2013, S. 327–403.
- Benyoucef, L.; Xie, X.: Supply chain design using simulation-based NSGA-II approach. In: Wang, L.; Ng, A.H.C.; Deb, K. (Hrsg.): Multi-objective Evolutionary Optimisation for Product Design and Manufacturing. London: Springer-Verlag London 2011, S. 455–491.
- Braungart, M.; McDonough, W.; Bollinger, A.: Cradle-to-cradle design: creating healthy emissions – a strategy for eco-effective product and system design. *Journal of Cleaner Production* (2007) 15, S. 1337–1348.
- Bundesministerium für Wirtschaft und Energie: Energieeffizienz in Zahlen - Entwicklungen und Trends in Deutschland 2019 2019.

- Deb, K.; Jain, H.: An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach. *IEEE Transactions on Evolutionary Computation* (2014) 18, S. 577–601.
- Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* (2002) 6, S. 182–197.
- Keramydas, C.; Mallidis, I.; Dekker, R.; Vlachos, D.: Cost and environmental trade-offs in supply chain network design and planning: the merit of a simulation-based approach. *Journal of Simulation* (2017) 11, S. 20–29.
- Rabe, M.; Dross, F.; Schmitt, D.; Ammouriova, M.: Decision Support for Logistics Networks in Materials Trading Using a Simheuristic Framework and User-generated Action Types. In: Wenzel, S.; Peter, T. (Hrsg.): *Simulation in Produktion und Logistik 2017*. Kassel: Kassel University Press 2017, S. 109–118.
- Rabe, M.; Spiekermann, S.; Wenzel, S.: *Verifikation und Validierung für die Simulation in Produktion und Logistik: Vorgehensmodelle und Techniken*. Springer-Verlag Berlin Heidelberg 2008.
- Salem, R.W.; Haouari, M.: A simulation-optimisation approach for supply chain network design under supply and demand uncertainties. *International Journal of Production Research* (2017) 55, S. 1845–1861.
- Schreiber, L.: Optimization and simulation for sustainable supply chain design. In: *Hamburg International Conference of Logistics (HICL) 2019, Hamburg, 2019*, S. 271–298.
- Tempelmeier, H.: *Bestandsmanagement in Supply Chains 2012*.
- Tordecilla, R.D.; Juan, A.A.; Montoya-Torres, J.R.; Quintero-Araujo, C.L.; Panadero, J.: Simulation-optimization methods for designing and assessing resilient supply chain networks under uncertainty scenarios: A review. *Simulation Modelling Practice and Theory* (2021) 106, S. 102–166.
- VDI 3633 Blatt 12: *Simulation von Logistik-, Materialfluss- und Produktionssystemen - Simulation und Optimierung 2020*.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Analyse von konventionellen Prioritätsregeln zur Reduktion von CO₂-Emissionen durch den Einsatz von Photovoltaikanlagen

Analysis of commonly used priority rules to reduce CO₂ emissions through the use of photovoltaic systems

Hajo Terbrack, Thorsten Claus, Matthias Götz, TU Dresden, Dresden (Germany),
hajo.terbrack@mailbox.tu-dresden.de, thorsten.claus@tu-dresden.de,
matthias.goetz@tu-dresden.de

Frank Herrmann, OTH Regensburg, Regensburg (Germany),
frank.herrmann@oth-regensburg.de

Maximilian Selmair, BMW Group, Regensburg (Germany),
maximilian.selmair@bmw.de

Abstract: This article discusses job shop scheduling through commonly used priority rules in the context of energy efficiency and emission pollution. While including an onsite photovoltaic system as a volatile energy source, the link between production-related energy consumption and emission pollution is analysed. By simulating production scheduling and onsite energy generation in different weather scenarios, we identify a nonlinear relation between total energy consumption and emissions caused by energy procurement from the grid. This optimization potential is addressed by a combination of two priority rules that considers generation output of the photovoltaic system in production scheduling. Our simulation results show an improvement in total energy consumption, emission pollution and throughput time.

1 Einführung

Die Reduktion von Treibhausgasemissionen spielt weltweit eine immer größere Rolle. So wurde in Deutschland zu Beginn des Jahres 2021 eine CO₂-Steuer eingeführt, um einen Anreiz zur Emissionsreduzierung zu schaffen. Mittelfristig wird in Deutschland das Ziel verfolgt, die Treibhausgasemissionen bis zu dem Jahr 2030 um 65% in Bezug auf das Jahr 1990 zu senken. Bis 2045 soll in Deutschland Klimaneutralität erreicht werden. Damit orientiert sich die Bundesregierung am Ziel des Pariser Abkommens, dass in der zweiten Hälfte dieses Jahrhunderts weltweit Treibhausgasneutralität erreicht werden soll (BEHG 2020; Bundestag 2021).

Einige wissenschaftliche Arbeiten beschäftigen sich bereits damit, CO₂-Emissionen durch entsprechende Produktionsplanung, etwa auf Ebene des Scheduling, zu verringern. Eine detaillierte Analyse dieser Artikel ist in der Literaturdatenbank zu Nachhaltiger Produktionsplanung von Terbrack et al. (2020) zu finden. Eine Argumentation in diesen Arbeiten ist, dass ein deutlicher Anteil elektrischer Energie durch fossile Brennstoffe erzeugt wird und eine rationelle Nutzung von Energie erhebliche Potentiale zur Reduzierung von energiebezogenen Treibhausgasen bietet (Ding et al. 2016). Im Artikel von Liu et al. (2017) etwa wird ein Job Shop Scheduling-Ansatz vorgestellt, der neben der Minimierung der Zykluszeit auf die Minimierung energiebezogener Emissionen abzielt. Wang et al. (2019) präsentieren ein Optimierungsmodell für ein Ein-Maschinen-Problem zur Minimierung von Emissionen und Zykluszeit. In Guo et al. (2020) werden für ein Fließfertigungsproblem die Zykluszeit, die Maschinenlautstärke sowie die energiebezogenen Emissionen minimiert. Die Minimierung des Stromverbrauchs resultiert dabei in der Regel in einer Verringerung der Emissionen.

Die Reduktion der CO₂-Emissionen korreliert aber nicht zwangsweise mit der Reduktion des Energieverbrauchs. So ist ein weiterer Ansatz zur Verringerung der CO₂-Emissionen von Produktionsunternehmen die Stromproduktion durch erneuerbare Energiequellen wie Photovoltaikanlagen (PV-Anlagen) (Liu 2016). Durch den Einsatz solcher Anlagen kann bei der Produktion ein gewisser Anteil des gesamten Stromverbrauchs durch emissionsfreien Strom abgedeckt werden. Jedoch hängt die generierte Menge des Solarstroms stark von unterschiedlichen Faktoren wie der Tageszeit, den Sonnenstunden und der Saison ab. Deshalb stehen Unternehmen mit Photovoltaikanlagen täglich unterschiedlich hohe Mengen an Solarstrom zur Verfügung. Aufgrund dieses Zusammenhangs zwischen Zeit und Menge an Solarstromerzeugung gilt es für Unternehmen zu berücksichtigen, dass die Produktion von emissionsfreiem Strom im Tagesverlauf schwanken kann.

Wie bereits in zahlreichen Studien zur Analyse energetischer Einflussfaktoren in Produktion und Logistik Simulationsmodelle zum Einsatz kommen (Wenzel et al. 2017), wird der Zusammenhang zwischen produktionsbezogenem Energieverbrauch und energiespezifischen Emissionen, kombiniert mit emissionsfreier Energieeigenerzeugung, im vorliegenden Artikel simulativ untersucht.

So wird in dieser Arbeit durch eine Simulationsstudie aufgezeigt, dass es Produktionsunternehmen gibt,

- a. bei denen eine Reduktion von CO₂-Emissionen durch Photovoltaikanlagen signifikant von dem konkreten Belegungsplan der Anlagen abhängt.
- b. bei denen das nach (a) vorhandene Potential durch konventionelle Prioritätsregeln – bedingt durch einen nichtlinearen Zusammenhang zwischen Energieverbrauch und Emissionsausstoß – entweder nicht oder nur zufällig genutzt und teilweise erschlossen wird.
- c. bei denen sich das nach (a) vorhandene Potential durch eine neu aufgestellte Kombination aus Prioritätsregeln in Verbindung mit der Berücksichtigung der Eigenenergieerzeugung besser nutzen lässt als bei einer Belegungsplanung durch konventionelle Prioritätsregeln.

Kapitel 2 stellt im Folgenden das Vorhaben der Simulationsuntersuchungen vor. Kapitel 3 erläutert den Aufbau und die Ergebnisse der Simulationsläufe. In Kapitel 4 werden die Erkenntnisse der Arbeit zusammengefasst sowie ein Ausblick präsentiert.

2 Untersuchungsgegenstand

In der Maschinenbelegungsplanung wird bestimmt, zu welchem Zeitpunkt und in welcher Reihenfolge die Produktionsaufträge an den verfügbaren Ressourcen, i.e. Produktionsmaschinen, bearbeitet werden. Dafür werden Optimierungsmodelle oder Heuristiken verwendet. Unter industriellen Rahmenbedingungen werden insbesondere einfache Heuristiken, wie Prioritätsregeln, eingesetzt. Im Kern werden dazu die zur Bearbeitung freigegebenen Aufträge als Operationen in Warteschlangen vor die einzelnen Stationen gereiht. Sobald eine Maschine frei ist, wird aus der entsprechenden Warteschlange der Auftrag mit der höchsten Priorität der Maschine zur Bearbeitung zugeteilt (Claus et al. 2015; Herrmann 2011).

Die Simulationsstudie verwendet als Heuristiken die folgenden sechs, in der Unternehmenspraxis häufig verwendeten Prioritätsregeln.

- FIFO (First-In-First-Out-Regel) und LIFO (Last-In-First-Out-Regel): Die Aufträge werden entsprechend der Länge der Wartezeit in absteigender (FIFO) oder aufsteigender (LIFO) Reihenfolge priorisiert.
- KOZ (Kürzeste Operationszeit-Regel) und LOZ (Längste Operationszeit-Regel): Die Priorität eines Auftrags bestimmt sich durch die Länge der Operationszeit auf der Maschine. Höchste Priorität erhält der Auftrag mit der kürzesten Operationszeit (KOZ) oder der längsten Operationszeit (LOZ).
- GRB (Größte Restbearbeitungszeit-Regel) und KRB (Kürzeste Restbearbeitungszeit-Regel): Abhängig von der Restbearbeitungszeit der noch ausstehenden Operationen wird die Priorität eines Auftrags bestimmt. Nach der GRB-Regel erhält der Auftrag mit der größten Restbearbeitungszeit die höchste Priorität, nach der KRB-Regel der Auftrag mit der kürzesten Restbearbeitungszeit.

Die resultierenden Belegungspläne werden hinsichtlich der Gesamtdurchlaufzeit aller Aufträge, des Gesamtenergieverbrauchs im Shopfloor sowie der energiebezogenen Emissionen bewertet. Weiterhin werden fünf Wetterszenarien betrachtet, in denen sich die Stromerzeugung von PV-Anlagen, bedingt durch variierende Sonneneinstrahlung, unterscheidet.

Ziel dieser Untersuchung ist es, gemäß (a) den Zusammenhang zwischen der Maschinenbelegungsplanung durch konventionelle Prioritätsregeln, dem resultierenden Energieverbrauch und dem Emissionsausstoß unter Berücksichtigung von Eigenenergieerzeugung zu analysieren. Dazu soll – bezogen auf (b) – Verbesserungspotential aufgezeigt werden. Im Anschluss wird, Punkt (c) entsprechend, dargelegt, dass durch eine Kombination aus zwei Prioritätsregeln dieses Potential besser adressiert werden kann als durch konventionelle Prioritätsregeln.

3 Simulationsuntersuchung

Mit Hilfe der Simulationssoftware Plant Simulation Version 16 wird in diesem Beitrag die Ressourcenbelegungsplanung in einer Werkstattfertigung simuliert. Zunächst wird untersucht, wie die unterschiedlichen Belegungspläne abhängig von der gewählten Prioritätsregel und dem Wetterszenario Unterschiede im Gesamtenergieverbrauch, in der Nutzung von eigenerzeugter Solarenergie und in den resultierenden Emissionen aufweisen. Im Anschluss wird eine Kombination aus zwei Prioritätsregeln vorgestellt, die in jeder der betrachteten Szenarien zu einem

niedrigeren Energieverbrauch sowie zu niedrigeren energiebezogenen Emissionen führt. Dazu wird simulativ für jedes Wetterszenario ein Schwellenwert bestimmt, mit dem die aufgestellte Kombination aus Prioritätsregeln zu besseren Ergebnissen führt.

3.1 Aufbau der Simulationsuntersuchung

Der Simulationsaufbau ist in Abbildung 1 dargestellt. Das modellierte Produktionssystem basiert auf der Arbeit von Selmair et al. (2015) und entspricht einer Werkstattfertigung. Es besteht aus fünf Maschinen mit jeweils einer Warteschlange. Über den Baustein *Energy Analyzer* wird der gesamte produktionsbezogene Stromverbrauch ermittelt, welcher bei den unterschiedlichen Belegungsplänen entsteht. Für die fünf Maschinen werden die Zustände „Arbeitend“, „Rüstend“, „Betriebsbereit“, „Standby“ und „Aus“ mit unterschiedlichen energetischen Leistungswerten zwischen 0 und 7 kW angenommen. Diese Leistungswerte werden in der vorliegenden Arbeit als konstant angenommen und für jede Maschine gilt: $P_{\text{Arbeitend}} > P_{\text{Rüstend}} > P_{\text{Betriebsbereit}} > P_{\text{Standby}} > P_{\text{Aus}}$.

Die Zeiten für die Zustandswechsel „Aus → Betriebsbereit“, „Betriebsbereit → Aus“, „Betriebsbereit → Standby“, „Standby → Betriebsbereit“ und „Standby → Aus“ sind für die fünf Maschinen gleich und entsprechen Werten zwischen 2 und 30 Minuten. Dabei gilt:

$$t_{\text{Aus} \rightarrow \text{Betriebsbereit}} > t_{\text{Standby} \rightarrow \text{Betriebsbereit}} > t_{\text{Betriebsbereit} \rightarrow \text{Aus}} > t_{\text{Betriebsbereit} \rightarrow \text{Standby}} = t_{\text{Standby} \rightarrow \text{Aus}}$$

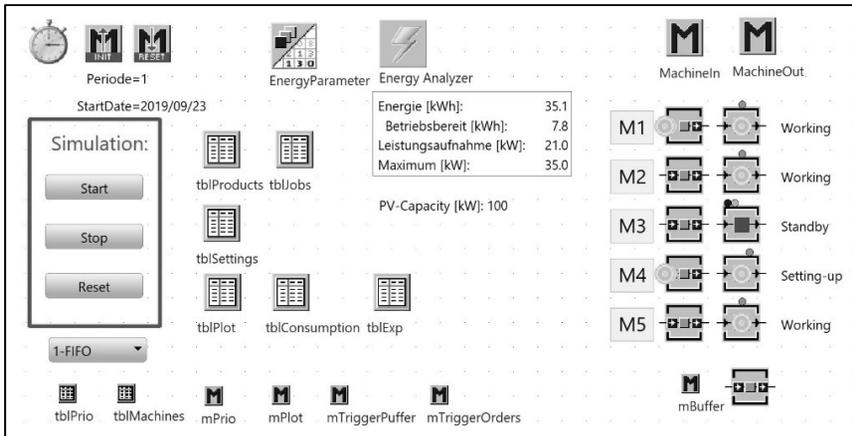


Abbildung 1: Aufbau des Simulationsmodells

In der Fallstudie werden drei Produkte (P1, P2, P3) gefertigt. Diese durchlaufen die Maschinen in unterschiedlicher Reihenfolge, wobei zur Fertigung von P1 und P3 jeweils 7 Operationen, zur Fertigung von P2 8 Bearbeitungsschritte durchgeführt werden. Hierzu werden produktspezifische und maschinenspezifische Rüst- und Bearbeitungszeiten zwischen 5 und 17 Minuten angenommen. Der Auftragsvorrat, dessen Bearbeitung auf den Produktionsmaschinen simuliert wird, besteht aus 9 Produktionsaufträgen mit jeweiligen Mengen zwischen 1 und 4 Stück. Der Auftragsvorrat ist zu Simulationsstart bekannt. Je nach Prioritätsregel werden die

Aufträge unterschiedlich in der Warteschlange vor der jeweiligen Produktionsmaschine angeordnet und in der resultierenden Reihenfolge auf der Maschine abgearbeitet. Sobald eine Produktionsmaschine eine Operation beendet hat und der soeben bearbeitete Auftrag vor die nächste Station gelegt wird, werden die Aufträge in den Warteschlangen erneut, entsprechend der jeweiligen Prioritätsregel, sortiert.

Des Weiteren wird eine Photovoltaikanlage mit Nennleistung 100 kW und 30° Südausrichtung berücksichtigt, die die Produktionsmaschinen mit Solarstrom versorgt. Die Dimensionierung der PV-Anlage basiert auf Expertenbefragungen unseres Praxispartners. Die Höhe der Eigenerzeugung von Solarenergie durch diese Photovoltaikanlage wird in Form von Leistungswerten je 15 Minuten angegeben. Wie Abbildung 2 zeigt, variiert das Erzeugungsprofil hierbei in fünf Wetterszenarien. Die Wetterszenarien entsprechen typischen Septembertagen in Regensburg, Deutschland, und sind der Simulationssoftware für Photovoltaikanlagen PV*SOL® entnommen.

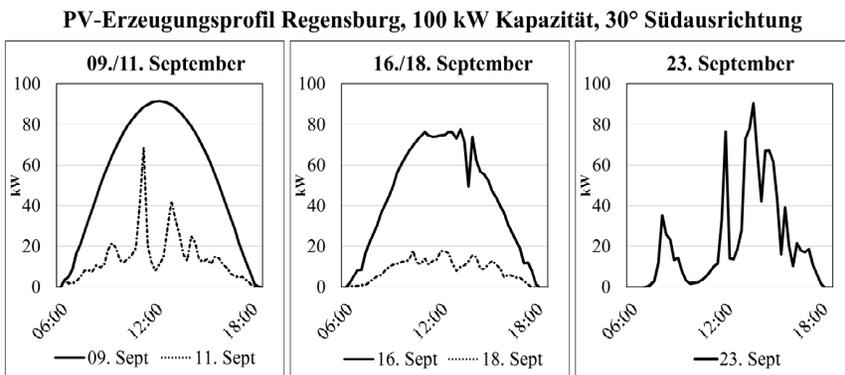


Abbildung 2: Erzeugungsprofile für fünf Septembertage in Regensburg.

Sofern der durch die Produktionsmaschinen ausgehende Energiebedarf nicht durch die PV-Anlage bedient werden kann, wird elektrische Energie aus dem Stromnetz bezogen. Dazu wird angenommen, dass der aus dem Stromnetz zugekaufte Strom Emissionen in Höhe von 401 g CO₂ pro kWh verursacht. Dieser Wert ist nach Icha und Kuhs (2020) eine Hochrechnung der Emissionen des deutschen Energiemixes für das Jahr 2019. An dieser Stelle sei erwähnt, dass, obwohl in der Realität von einem zeitvariablen Emissionsfaktor ausgegangen werden kann, ein weiter Teil der Literatur zu energieorientierter Produktionsplanung einen konstanten Umrechnungsfaktor annimmt (wie beispielsweise Guo et al. 2020, Liu et al. 2017, Wang et al. 2019). Daher wird die Annahme eines konstanten Faktors von 401 gCO₂/kWh für die Experimente als hinreichend geeignet eingeschätzt.

Für die einzelnen Wetterszenarien wird jeweils die Ressourcenbelegungsplanung nach den sechs genannten Prioritätsregeln simuliert. Arbeitsbeginn ist 06.00 Uhr morgens am entsprechenden Septembertag.

3.2 Ergebnisse der konventionellen Prioritätsregeln

Die Ergebnisse der Simulationsläufe mit Ressourcenbelegung durch konventionelle Prioritätsregeln sind in Tabelle 1 aufgeführt. Diese zeigt die Gesamtdurchlaufzeit aller Aufträge (DLZ), den Gesamtenergieverbrauch im Shopfloor (EV_{Gesamt}) und den Energiebezug aus dem Stromnetz (EV_{Netz}) sowie die resultierenden Emissionen (EM) für die durchgeführten Experimente.

Tabelle 1: Ergebnisse der Ressourcenbelegung durch konventionelle Prioritätsregeln in fünf Wetterszenarien.

		FIFO	LIFO	KOZ	LOZ	GRB	KRB	Minimum
DLZ	[h]	13.48	12.95	13.02	12.78	12.87	14.15	LOZ 12.78
EV_{Gesamt}	[kWh]	233.50	233.75	229.75	231.00	226.50	240.50	GRB 226.50
09. Sept	EV_{Netz} [kWh]	20.63	19.75	12.00	14.25	18.13	26.50	KOZ 12.00
	EM [kg]	8.27	7.92	4.81	5.71	7.27	10.63	KOZ 4.81
11. Sept	EV_{Netz} [kWh]	91.58	87.16	83.33	84.66	84.25	97.10	KOZ 83.33
	EM [kg]	36.72	34.95	33.42	33.95	33.79	38.94	KOZ 33.42
16. Sept	EV_{Netz} [kWh]	27.31	25.56	17.55	20.05	25.06	32.06	KOZ 17.55
	EM [kg]	10.95	10.25	7.04	8.04	10.05	12.86	KOZ 7.04
18. Sept	EV_{Netz} [kWh]	128.46	128.71	124.71	125.96	121.46	135.46	GRB 121.46
	EM [kg]	51.51	51.61	50.01	50.51	48.71	54.32	GRB 48.71
23. Sept	EV_{Netz} [kWh]	82.24	85.77	78.52	80.49	81.07	92.27	KOZ 78.52
	EM [kg]	32.98	34.39	31.49	32.28	32.51	37.00	KOZ 31.49

Die LOZ-Regel erzielt in dieser Fallstudie die geringste Durchlaufzeit von 12.78 h. Die Anordnung nach der LOZ-Regel resultiert dabei nicht im geringsten Energieverbrauch, sondern die Priorisierung nach der GRB-Regel erreicht diesen mit 226.50 kWh. Die Unterschiede im Energieverbrauch lassen sich ableiten durch die Energiebedarfe bei Rüstvorgängen, Hoch- und Herunterfahren von Maschinen und in Standby-Zeiten, deren Anzahl und Länge je nach Prioritätsregel variiert. Wie bereits in Selmaier et al. (2015) analysiert, kann aus einem solchen nichtlinearen Zusammenhang zwischen Durchlaufzeit und Energieverbrauch im Shopfloor Optimierungspotential hinsichtlich der Energieeffizienz abgeleitet werden.

Begründet durch die unterschiedliche Sonneneinstrahlung und die damit einhergehende Erzeugungsleistung der Photovoltaikanlage variiert in den unterschiedlichen Wetterszenarien der Bezug elektrischer Energie aus dem Stromnetz sowie die daraus folgenden Emissionen. Auch hierbei zeigen sich deutliche Unterschiede zwischen den Ressourcenbelegungsplänen der einzelnen Prioritätsregeln, wodurch Aussage (a) bestätigt werden kann. In vier der fünf Wetterszenarien ist die Maschinenbelegung nach der KOZ-Regel diejenige mit dem geringsten Energiebezug aus dem Stromnetz und folglich auch diejenige mit den niedrigsten Emissionen. So werden in vier von fünf Wetterszenarien durch die Ressourcenbelegung nach der KOZ-Regel die Aufträge geeigneter nach der PV-Erzeugung angeordnet, sodass der Energiebezug aus dem Stromnetz und die

Emissionen niedriger sind als bei den anderen Belegungsplänen. Nur in einem der fünf Szenarien, am 18. September, entsprechen die Emissionen bei der energieärmsten GRB-Regel auch dem Minimum unter den konventionellen Prioritätsregeln. Somit erzielt in einem Großteil der Wetterszenarien eine Regel, mit der nicht der niedrigste Gesamtenergieverbrauch erreicht wird, eine Ressourcenbelegungsplanung mit den niedrigsten Emissionen. Daraus kann abgeleitet werden, dass es Szenarien gibt, in denen unter Berücksichtigung von emissionsfreier Energieeigenerzeugung und einem konstanten Emissionsfaktor [CO₂-Äquivalente g pro kWh Strombezug] zwischen dem Energieverbrauch im Shopfloor und den energiebezogenen Emissionen keine lineare Korrelation besteht. Weiterhin wird in vier von fünf Wetterszenarien selbst bei der emissionsniedrigsten Prioritätsregel nicht die gesamte Solarenergie (Ertrag_{PV}) verwendet, wie Tabelle 2 jeweils für den emissionsärmsten Belegungsplan eines jeden Wetterszenarios zeigt. Dazu wird der PV-Reinertrag eines jeden Szenarios zwischen Arbeitsbeginn und Produktionsende mit dem Energieverbrauch verglichen. Basierend auf diesen Ergebnissen kann gefolgert werden, dass Produktionsunternehmen existieren, bei denen die in (b) beschriebenen Kriterien auftreten.

Tabelle 2: Energieverbrauch und -quellen der emissionsärmsten Belegungspläne.

	EV _{Gesamt} [kWh]	EV _{PV} [kWh]	Ertrag _{PV} [kWh]	EV _{Netz} [kWh]	EM [kg]
09. Sept: KOZ-Regel	229.75	217.75	693.21	12.00	4.81
11. Sept: KOZ-Regel	229.75	146.42	177.93	83.33	33.42
16. Sept: KOZ-Regel	229.75	212.20	556.57	17.55	7.04
18. Sept: GRB-Regel	226.50	105.04	105.04	121.46	48.71
23. Sept: KOZ-Regel	229.75	151.23	283.86	78.52	31.49

Zur Erhöhung des Nutzungsgrads von emissionsfreier Eigenenergieerzeugung und zur Vermeidung von Energieverlusten können Energiespeicher oder Einspeiseoptionen in das Stromnetz genutzt werden. Ein weiterer Lösungsweg besteht in der Ausrichtung der Maschinenbelegungsplanung auf die Erzeugungsleistung. Hierzu wird im Folgenden eine Möglichkeit aufgezeigt.

3.3 Kombination aus zwei Prioritätsregeln zur Erhöhung der Eigenenergienutzung und Reduzierung von Emissionen

Die Ergebnisse aus Kapitel 3.2 zeigen, dass die Prioritätsregeln LOZ, GRB und KOZ für die Durchlaufzeit, den gesamten Energieverbrauch sowie die Emissionen die besten Ergebnisse liefern. Dabei wird das Potential der Eigenerzeugung von Solarenergie zur Emissionsreduktion jedoch zumeist nicht vollständig durch diese Prioritätsregeln ausgeschöpft.

Zur Verbesserung der energie- und emissionsbezogenen Kennzahlen sollte die Maschinenbelegungsplanung die Eigenenergieerzeugung berücksichtigen. Dazu werden in weiteren Simulationsläufen für jedes Wetterszenario die Ressourcenbelegungsplanung durch eine Kombination aus LOZ- und KOZ-Regel durchgeführt. Die Wahl der Prioritätsregel basiert auf der als bekannt angenommen mittleren Erzeugungsleistung der PV-Anlage für die nächste Stunde ab dem

Planungszeitpunkt. Ist diese mittlere Erzeugungsleistung größer gleich einem Schwellenwert s , der relativ zur Nennleistung der PV-Anlage von 100 kW ausgedrückt wird, werden die Aufträge in der Warteschlange nach der LOZ-Regel angeordnet. Bedingt durch konstante energetische Leistung während des Maschinenzustands „Arbeitend“ stellen die Aufträge mit längerer Operationszeit die energieintensiven Aufträge in der Warteschlange dar. Somit werden bei hoher PV-Erzeugung durch die LOZ-Regel die energieintensiven Aufträge priorisiert bearbeitet. Ansonsten, wenn die mittlere Erzeugungsleistung für die nächste Stunde kleiner als der Wert $s \cdot \text{Nennleistung}$ ist, werden die Aufträge nach der KOZ-Regel priorisiert.

Durch eine Wiederholung der Simulationsläufe, bei denen iterativ der Schwellenwert s um 0.01 im Intervall $[0, 1]$ erhöht wird, wird für jedes Wetterszenario ein Schwellenwert ermittelt, bei dem diese Kombination aus LOZ- und KOZ-Regel zu besseren Ergebnissen hinsichtlich Energieverbrauch und Emissionen führt. Die Ergebnisse dieser Simulationsläufe zeigt Tabelle 3.

Tabelle 3: Ergebnisse der Simulationsläufe mit einer Kombination aus LOZ- und KOZ-Regel abhängig von Schwellenwert s .

LOZ-KOZ-s-Regel	Schwellenwert s	DLZ [h]	EV _{Gesamt} [kWh]	EV _{PV} [kWh]	EV _{Netz} [kWh]	EM [kg]
09. Sept	0.65	12.42	225.50	220.33	5.17	2.07
11. Sept	0.16	12.20	222.00	149.92	72.08	28.90
16. Sept	0.50	12.42	225.50	214.95	10.55	4.23
18. Sept	0.13	12.73	223.75	105.04	118.71	47.60
23. Sept	0.33	12.73	225.75	153.98	71.77	28.78

Für die fünf betrachteten Wetterszenarien resultiert diese Kombination aus LOZ- und KOZ-Regel bei geeigneter Wahl des Schwellenwerts s jeweils in einem Belegungsplan mit niedrigeren Emissionen. Indem die Erzeugungsleistung der Photovoltaikanlage Berücksichtigung findet, erfolgt ein geringerer Energiebezug aus dem Stromnetz in jedem der fünf Wetterszenarien. Damit kann (c) bestätigt werden.

Weiterhin kann in den betrachteten Szenarien durch die Kombination aus LOZ- und KOZ-Regel die Gesamtdurchlaufzeit aller Aufträge sowie der Energieverbrauch im Shopfloor reduziert werden – respektive einem geeigneten Schwellenwert s . Für diesen ist zu beobachten, dass in den Szenarien mit insgesamt höherer Sonneneinstrahlung der geeignetste Schwellenwert s auch einen höheren Wert annimmt und somit die Kombination aus LOZ- und KOZ-Regel auch in diesen Szenarien angestrebt wird. So reicht es für die aufgestellte LOZ-KOZ-s-Regel nicht aus, einen Schwellenwert für alle Wetterszenarien zu nutzen. Stattdessen ist die Wahl des Schwellenwerts entscheidend.

Am Beispiel des zweiten Wetterszenarios zeigt Abbildung 3 den Leistungsbedarf des emissionsärmsten Belegungsplan unter den konventionellen Prioritätsregeln (KOZ), den Leistungsbedarf des Belegungsplans nach der entworfenen LOZ-KOZ-s-Regel sowie Ausschnitte des Erzeugungsprofils der PV-Anlage. Der zumeist höhere Leistungsbedarf der Kombination (LOZ-KOZ-s) als der KOZ-Regel in Zeiträumen mit hoher PV-Erzeugung sowie in umgekehrter Weise für Zeiträume mit niedriger

PV-Erzeugung begründet die niedrigeren Emissionen für den Belegungsplan nach LOZ-KOZ-s-Kombination. Dennoch zeigt die Abbildung ebenso das weiterhin verbleibende Optimierungspotential, indem der Leistungsbedarf der Maschinenbelegung noch mehr an das PV-Erzeugungsprofil ausgerichtet werden kann.

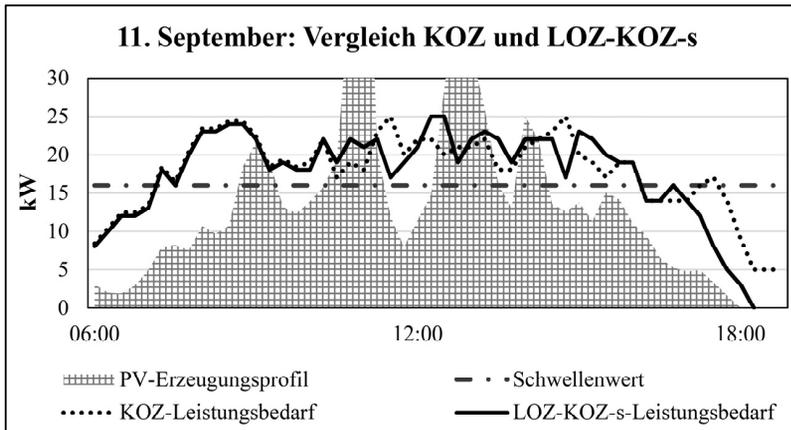


Abbildung 3: Vergleich der KOZ-Regel mit der aufgestellten Kombination aus LOZ- und KOZ-Regel für den 11. September.

4 Erkenntnisse und Ausblick

In dieser Arbeit wurden Belegungspläne einer Werkstattfertigung nach konventionellen Prioritätsregeln in Verbindung mit Eigenstromerzeugung einer PV-Anlage untersucht und hinsichtlich energiebezogener Emissionen analysiert. Durch einen nichtlinearen Zusammenhang aus Energieverbrauch und Emissionen in vier der fünf Wetterszenarien konnte Optimierungspotential festgestellt werden. Zur Emissionsreduzierung durch die Maschinenbelegungsplanung ist somit nicht ausschließlich der Umfang des Energieverbrauchs entscheidend, sondern auch der Zeitraum des Verbrauchs, da letzterer in Verbindung mit einer volatilen PV-Energieerzeugung die produktionsspezifischen, energiebezogenen Emissionen beeinflussen kann. Eine wie in dieser Arbeit erfolgte Kombination konventioneller Prioritätsregeln zur Ausrichtung auf eine Eigenenergieerzeugung kann bereits bessere Ergebnisse hinsichtlich der Nutzung volatiler Energiequellen erzielen und zu einer Reduktion von energiebezogenen Emissionen führen, ohne notwendigerweise ökonomische Kriterien wie die Gesamtdurchlaufzeit zu verschlechtern.

Zur Validierung der vorgestellten LOZ-KOZ-s-Regel sind Simulationsläufe zu weiteren Szenarien und Fallstudien geplant. Möglicherweise führt diesbezüglich ebenso eine Anpassung des Vorhersagehorizonts zur PV-Erzeugung zu besseren Ergebnissen. Auch soll in weiteren Experimenten die Abbildung von Unsicherheit bezüglich des Wetters bei der Bestimmung eines geeigneten Schwellenwerts s berücksichtigt werden. Bezogen auf den in dieser Arbeit angenommen konstanten Umrechnungsfaktor zwischen Emissionen und elektrischer Energie soll in weiteren Simulationsläufen davon abgewichen werden und die Annahme eines zeitvariablen Faktors untersucht werden. Weiterhin könnte die Erhöhung der Flexibilität der

Produktion, etwa durch die Einführung variabler Maschinengeschwindigkeiten oder heterogener paralleler Maschinen mit Unterschieden in den Leistungsbedarfen, das Potential der Eigenenergieerzeugung stärker ausschöpfen. Ein weiterer für zukünftige Forschung geplanter Ansatz besteht in der Modellierung und Lösung eines entsprechenden Optimierungsmodells, das Energieverbrauch, Stromerzeugung und Emissionen simultan betrachtet.

Literatur

- BEHG (2020): Brennstoffemissionshandelsgesetz vom 12. Dezember 2019 (BGBl. I S. 2728), das durch Artikel 1 des Gesetzes vom 3. November 2020 (BGBl. I S. 2291) geändert worden ist.
- Bundestag (2021): Drucksache des Deutschen Bundestages 19/30230 vom 02.06.2021: Entwurf eines Ersten Gesetzes zur Änderung des Bundes-Klimaschutzgesetzes.
- Claus, T.; Herrmann, F.; Manitz, M.: Produktionsplanung und -steuerung: Forschungsansätze, Methoden und deren Anwendungen. Springer-Verlag, 2015.
- Ding, J.; Song, S.; Wu, C.: Carbon-efficient scheduling of flow shops by multi-objective optimization. *European Journal of Operational Research*, 248 (2016) 3, pp. 758-771.
- Guo, H.; Li, J.; Yang, B.; Mao, X.; Zhou, Q.: Green scheduling optimization of ship plane block flow line considering carbon emission and noise. *Computers & Industrial Engineering*, 148 (2020), pp. 106680.
- Herrmann, F.: Operative Planung in IT-Systemen für die Produktionsplanung und -steuerung: Wirkung, Auswahl und Einstellhinweise von Verfahren und Parametern. Springer-Verlag, 2011.
- Icha, P.; Kuhs, G.: Entwicklung der spezifischen Kohlendioxid-Emissionen des deutschen Strommix in den Jahren 1990 - 2019. Dessau: Umweltbundesamt, 2020.
- Liu, C.: Mathematical programming formulations for single-machine scheduling problems while considering renewable energy uncertainty. *International Journal of Production Research*, 54 (2016) 4, pp. 1122-1133.
- Liu, Q.; Zhan, M.; Chekem, F.; Shao, X.; Ying, B.; Sutherland, J.: A hybrid fruit fly algorithm for solving flexible job-shop scheduling to reduce manufacturing carbon footprint. *Journal of Cleaner Production*, 168 (2017), pp. 668-678.
- Selmair, M.; Herrmann, F.; Claus, T.; Teich, E.: Potentiale in der Reduzierung des Gesamtenergieverbrauchs einer Werkstattfertigung in der Maschinenbelegungsplanung. In: Rabe, M.; Clausen, U. (Eds): *Simulation in Produktion und Logistik 2015*. Fraunhofer IRB Verlag, Stuttgart, 2015, S. 177-186.
- Terbrack, H.; Frank, I.; Herrmann, F.; Claus, T.; Trost, M.; Götz, M.: A literature database on ecological sustainability in hierarchical production planning, Anwendungen und Konzepte der Wirtschaftsinformatik 12 (2020), pp. 36-40.
- Wang, J.; Yao, S.; Sheng, J.; Yang, H.: Minimizing total carbon emissions in an integrated machine scheduling and vehicle routing problem. *Journal of Cleaner Production*, 229 (2019), pp. 1004-1017.
- Wenzel, S.; Peter, T.; Stoldt, J.; Schlegel, A.; Groß, G.; Pitsch, H.; Rabe, M.; Seewaldt, M.: Betrachtungen energetischer Einflussfaktoren in der Simulation in Produktion und Logistik: Eine Literaturanalyse. In: Wenzel, S.; Peter, T. (Eds): *Simulation in Produktion und Logistik*, kassel university press, Kassel, 2017, S. 9-18.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Material- und Energieflusssimulation zur prädiktiven Bestimmung von Ofenreinigungsintervallen

***Material and energy flow simulation for predictive determination of
furnace cleaning intervals***

Johannes Dettelbacher, Wolfgang Schlüter, HS Ansbach, Ansbach (Germany),
johannes.dettelbacher@hs-ansbach.de, wolfgang.schlüter@hs-ansbach.de

Abstract: Within the scope of this work, an aluminium casting operation is modelled with a material and energy flow simulation. The used simulation model focusses on the melting operation, where grey box models describe the melting furnaces. The essential physical model is based on a differential equation system, which displays the material and energy flows in the furnace. In order to determine optimal cleaning intervals for the melting furnaces, the simulation is coupled with an optimization. As an optimization method, a genetic algorithm is compared with a brute force method. Within various specific operating situations the quality of the simulation-based optimization is proved. As a result, the model can reduce operational disturbances and increase energy efficiency.

1 Einleitung

Durch die zunehmende Globalisierung und steigende Energiepreise erhöht sich der Wettbewerbsdruck auf die energieintensive Gussindustrie in Deutschland. In diesen Betrieben nimmt zwar das Bewusstsein für die Energieeffizienz zu, jedoch liegt der Fokus auf der Produktion der Bauteile und Effizienzmaßnahmen werden nur angewendet, wenn diese die Produktionskennzahlen nicht negativ beeinflussen. Während Verbesserungen durch technische Maßnahmen in der Branche bereits umgesetzt werden, bieten Optimierungen im Betriebsablauf noch weitere Potentiale (Herrmann et al. 2013; LFU 2005). Insbesondere die Werkzeuge der mathematischen Modellierung mit Simulationen und Optimierungen können hierbei zielführend für die Beurteilung von Effizienzmaßnahmen eingesetzt werden (Gutenschwager et al. 2017; Römer und Strassburger 2016). In vorherigen Forschungsarbeiten konnten auf diesem Weg Einsparpotentiale bei der Verteilung des festen und flüssigen Aluminiums und der Betriebsweise der Schmelzöfen identifiziert werden (Schlüter et al. 2017). Im Rahmen dieser Untersuchung wird das Simulationsmodell für die Prognostizierung von geeigneten Reinigungsintervallen der Schmelzöfen verwendet.

Ziel hierbei ist, kritische Betriebsituationen durch Reinigungsphasen zu vermeiden und einen störungsfreien und energieeffizienten Betrieb der Produktionsmaschinen zu ermöglichen. In der vorliegenden Arbeit werden der betrachtete Gussbetrieb, das Simulationsmodell sowie das Steuerungsmodul beschrieben und verschiedene Optimierungsoptionen betrachtet. Die Optimierungsergebnisse werden anhand spezifischer Betriebsituationen auf ihre Güte geprüft.

2 Aluminium-Gussbetriebe und Ofenreinigungen

Anhand der Abbildung 1 wird der Betriebsablauf in einem typischen Aluminium-Gussbetrieb beschrieben. Dabei wird besonders auf die Reinigung der Schmelzöfen, deren zeitlicher Ablauf optimiert werden soll, eingegangen.

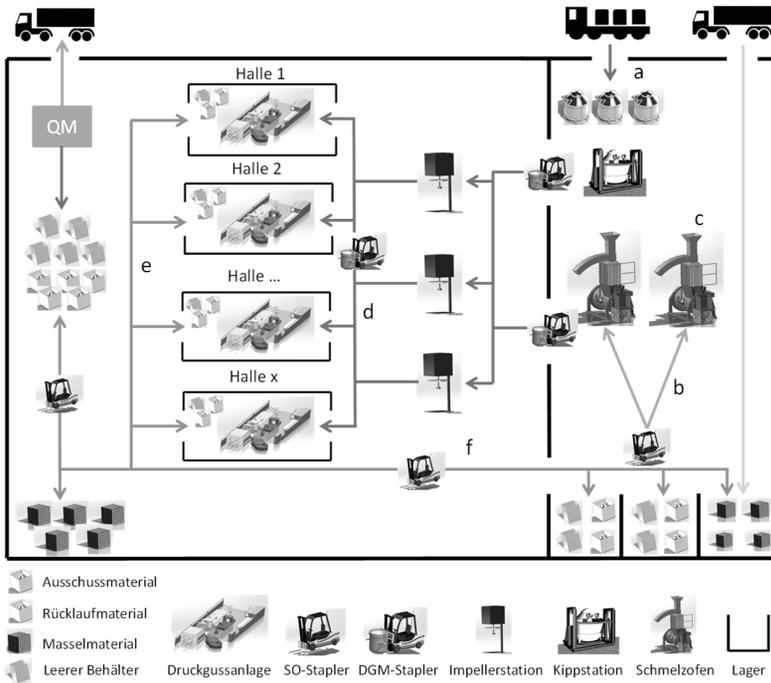


Abbildung 1: Schema eines Aluminium-Gussbetriebs mit Prozessschritten

Die zugrundeliegenden Prozesse sind:

- Anlieferung von flüssigem Aluminium (Abb. 1, a),
- Beschickung der gasbetriebenen Schachtschmelzöfen über Stapler mit Masseln (Metallbarren), Rücklauf- oder Ausschussmaterial (Abb. 1, b)
- Erwärmen, Schmelzen und Überhitzen bzw. Warmhalten des Metalls (Abb. 1, c)
- Verteilung des flüssigen Aluminiums mit Staplern auf die Dosieröfen der Druckgussmaschinen (Abb. 1, d)

- Produktion von Gussteilen in den Druckgussanlagen und Qualitätsprüfung (Abb. 1, e)
- Transport von Materialbehältern aus dem Druckgussbetrieb oder von Masselpaketen aus dem Lager zum Schmelzbetrieb (Abb. 1, f)

Im Fokus dieser Arbeit stehen die Reinigungsintervalle der Schmelzöfen (Freischmelzen). Da beim Schmelzprozess Ablagerungen an den Ofenwänden entstehen, welche die Effizienz und Lebensdauer der Anlagen reduzieren, müssen die Schmelzöfen entsprechend gereinigt werden. Nach Herstellerangaben soll jeder Schmelzofen pro Betriebsschicht einmal gereinigt werden, um Ablagerungen an den Ofenwänden zu vermeiden, was in der betrieblichen Praxis auch so umgesetzt wird. Beim Freischmelzen werden die Brenner genutzt, um die Ablagerungen abzuschmelzen und es kann kein neues Material beschickt werden. Die Druckgussmaschinen sind für die Produktion von Bauteilen auf die Flüssigaluminiumversorgung der Schmelzöfen angewiesen. Reinigungsphasen in kritischen Betriebssituationen (Flüssigaluminiummangel) können zu Störungen der Produktion führen. Durch die Reinigung erfolgt eine Unterbrechung des Schmelzvorganges, was zu einer Reduzierung des warmzuhaltenden Flüssigaluminiums führt. Dadurch beeinflusst der Reinigungszeitpunkt auch den Energieverbrauch der Öfen.

3 Simulationsmodell

Ausgehend von einem bestehenden Simulationsmodell wird das System in ein Teilmodell Schmelzbetrieb und ein Teilmodell Gussbetrieb aufgegliedert. Mit dem Teilmodell Gussbetrieb können die Materialverbräuche der einzelnen Produktionsmaschinen über die Bruttoproduktgewichte und den Taktzeiten bestimmt werden. Somit können der Aluminiumbedarf und die optimalen Belieferungszeiten der Maschinen über eine vorangehende eigenständige Optimierung bestimmt werden (Dettelbacher und Schlüter 2020).

Dies ermöglicht, die Systemgrenzen beim weiteren Simulationsmodell um den Schmelzbetrieb zu legen. Ziel des Schmelzbetriebes ist es, den Bedarf an Flüssigaluminium mit den Schmelzöfen zu decken. Hierfür sind im Teilmodell Schmelzbetrieb die Schmelzöfen abgebildet, welche sich jeweils in einen Schmelzschacht und die Ofenwanne unterteilen lassen. Für die Schmelzöfen wird ein Greyboxmodell verwendet, welches ein physikalisches Grundmodell mit, an Betriebsdaten und CFD-Simulationen kalibrierten, Parametern kombiniert. Das physikalische Grundmodell basiert als dynamisches System auf ein Differentialgleichungssystem, welches ausgehend von einer Verbrennungsrechnung die Wärmeübertragung vom Rauchgas auf das Aluminium und den Schmelz- und Warmhalteprozess im Schmelzofen abbildet. Der Material- und Energiefluss im Ofen wird im Simulationsmodell bidirektional gekoppelt. Ein Schema des Schmelzofenmodells ist in Abbildung 2 dargestellt. Im Modell wird neben der Änderung der Aluminiumoberfläche, welche einen Einfluss auf die Wärmeübertragung hat, auch der Wärmeverlust über die Ofenwand berücksichtigt.

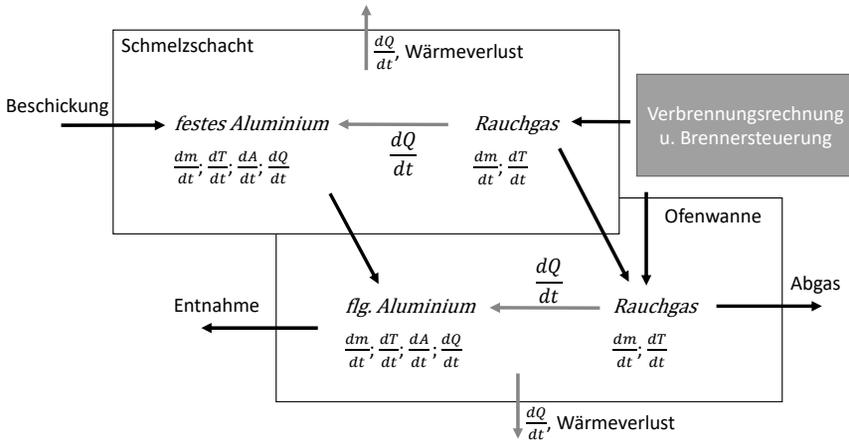


Abbildung 2: Schema des Schmelzofenmodells

Das Modell der Schmelzöfen ist mit einer Steuerung verbunden, welche die Staplertransporte und die Ofenreinigungen veranlasst. Über die Stapleraufträge kann die Beschickung der Schmelzöfen mit festem Aluminium oder die Entnahme des Flüssigaluminiums erfolgen. Die Vorgabe der Reinigungsphasen kann in den Simulationsdurchläufen variiert werden, um geeignete Freischmelzintervalle zu bestimmen. Validiert wurde das Material- und Energieflussmodell anhand der Betriebsdaten von zwei realen Betrieben.

Die Modelle werden in der Umgebung Matlab/Simulink von MathWorks erstellt. Eine Herausforderung stellt hierbei der modellprädiktive Einsatz der Modelle im Rahmen einer simulationsbasierten Optimierung dar. Hierfür müssen neben der Anpassung der Systemgrenze auch softwaretechnische Änderungen in den Simulink-Modellen vorgenommen werden, um die Rechenzeit zu reduzieren. Hierfür wurde das Simulationsmodell anhand des Simulink-Profilers auf die Simulationsdauer optimiert und stellenweise in C-Code kompiliert.

4 Optimierung der Ofenreinigungsintervalle

Während die Energieeffizienz stark an Bedeutung gewinnt, liegt die Priorität stets auf der Sicherstellung der Produktion. Im Fall eines Gusstriebes bedeutet das, dass der störungsfreie Betrieb der Druckgussmaschinen gewährleistet ist. Die Reinigungsphasen der Öfen stellen hierbei unvermeidbare Unterbrechungen des Schmelzbetriebes dar. Es gilt, die Reinigungsphasen so zu optimieren, dass ein störungsfreier und effizienter Betrieb der Produktionsmaschinen weiterhin möglich ist. Hierfür kann eine Methode für die Verteilung der Reinigungsintervalle über Kennzahlen eingesetzt werden (Dentel et al. 2018).

Im Rahmen dieser Arbeit wird der Ansatz gewählt, die Simulation des Schmelzbetriebes mit einem Optimierungsmodul zu verknüpfen. Dieses Modul übergibt mögliche Freischmelzintervalle an die Simulation. Da die Öfen nur einmal pro Schicht gereinigt werden, sind die Variablen des Optimierungsproblems auf die Anzahl der Öfen begrenzt. Anhand von zeitlichen Restriktionen sind

Nebenbedingungen für mögliche Freischmelzintervalle vorgegeben. Ofenreinigungen können bis zu zwei Stunden vorgezogen oder aufgeschoben werden. Somit ergibt sich ein Zeitintervall von vier Stunden, welches für die Optimierung in Zeitabschnitte diskretisiert wird. Als weitere Nebenbedingung wird festgelegt, dass aufgrund des Personalbedarfs jeweils nur ein Ofen gleichzeitig gereinigt werden kann.

Beurteilt wird die Simulation an den Größen Betriebsstörungen und spezifischen Energieverbrauch. Die Teilziele werden über Gewichtungsfaktoren zu einer Zielfunktion f verbunden, welche im Rahmen der Optimierung minimiert wird (s. Gl. 1). Während der spezifische Energieverbrauch über den Gasverbrauch der Schmelzöfen und der geschmolzenen Masse bestimmt wird, werden die Betriebsstörungen anhand der Abdeckung des Flüssigaluminiumbedarfs bemessen.

$$f(x) = w_1 * EV(x) + w_2 * BS(x) \quad (1)$$

x : *Reinigungszeiten der Schmelzöfen*

w_1 und w_2 : *Gewichtungsfaktoren der Teilziele*

$EV(x)$: *spezifischer Energieverbrauch der Schmelzöfen*

$BS(x)$: *Betriebsstörung aufgrund von Materialmangel*

Die Gewichte sind auf reale Betriebsziele so angepasst, dass vorrangig Betriebsstörungen vermieden werden und anschließend der Energieverbrauch optimiert wird. Als Optimierer wird ein genetischer Algorithmus verwendet, welcher sich für multikriterielle simulationsbasierte Optimierungsprobleme eignet (Sobottka 2017). Hierfür wurde der Algorithmus aus der Optimization Toolbox von Matlab auf das vorliegende Problem angepasst. Neben der Mutation und dem Crossover beinhaltet dieser Algorithmus auch Elitismus, welcher die besten Lösungen einer Generation für weitere Generationen abspeichert.

Der Einsatz der Optimierungsmethode wird mit der einfachen Brute-Force-Methode verglichen, bei der unter Berücksichtigung der Nebenbedingungen alle möglichen Varianten durchsimuliert werden. Zum einen kann dieser Ansatz bei niedriger Anzahl der Variationen als Optimierungsmethode eingesetzt werden, zum anderen dient er zur Ermittlung des globalen Optimums und somit zur Beurteilung des Ergebnisses des genetischen Algorithmus. Für eine feinere Diskretisierung des möglichen Reinigungsfensters oder für deutlich mehr Öfen ist der Brute-Force-Ansatz aufgrund der dann exponentiell steigenden Rechendauer jedoch nicht mehr praktikabel einsetzbar. Verglichen werden die optimierten Ofenreinigungsintervalle mit einer statischen Reinigungsstrategie, welche auf einer Reinigung nach Reihenfolge basiert und auf Grund ihrer Einfachheit auch in den Betrieben praktiziert wird.

Bei der Implementierung der simulationsbasierten Optimierung wird neben der Optimization Toolbox auch die Parallel Computing Toolbox von Matlab verwendet. Für die Berechnungen wird ein Rechnercluster mit 20 CPU-Kernen eingesetzt, um die Simulationen parallel auszuführen.

5 Ergebnisse

Für die Untersuchung wird ein Aluminium-Gussbetrieb mit vier Schmelzöfen verwendet. In der Simulation wird ein Zeitraum von 12 Stunden (1 ½ Betriebsschichten) betrachtet, um mögliche Auswirkungen der Ofenreinigungen in

einer Schicht auf die nachfolgende Schicht mit abzubilden. Es werden zwei Szenarien betrachtet. Im Szenario 1 wird der Betrieb in einer Schicht betrachtet, in welcher zu jedem Zeitpunkt die Versorgung mit Flüssigaluminium gewährleistet ist. Im Szenario 2 wird eine kritische Betriebsituation mit hohem Flüssigaluminiumbedarf betrachtet, bei welcher ungeeignete Ofenreinigungsintervalle zu Betriebsstörungen hinsichtlich der Flüssigaluminiumentnahme führen.

Im Szenario 1 kürzt sich das Teilziel Betriebsstörung aus der Zielfunktion, da die Flüssigaluminiumversorgung unabhängig von der Wahl der Reinigungsintervalle stets gewährleistet ist. Somit wird lediglich der Energieverbrauch der Öfen optimiert. In Abbildung 3 ist der beispielhafte Verlauf der jeweils besten Lösung einer Generation des genetischen Algorithmus gezeigt. Die Werte sind auf dem Fitnesswert, welcher sich aus der statischen Reinigungsstrategie ergibt, normiert. Zudem wird das globale Optimum, welches über den Brute-Force-Ansatz ermittelt wird, angegeben. In diesem Fall wird das globale Optimum vom genetischen Algorithmus nach 32 Generationen erreicht. Da der genetische Algorithmus stochastische Bestandteile enthält, kann sich der Verlauf der Fitnesswerte je nach Durchlauf unterscheiden und das Erreichen des globalen Optimums ist nicht gewährleistet. Die Rechenzeit des genetischen Algorithmus beläuft sich in diesem Beispiel auf unter 1/2 Minute pro Generation. Mit der Optimierung konnte eine Verbesserung des spezifischen Energieverbrauches von 1,4 % erreicht werden.

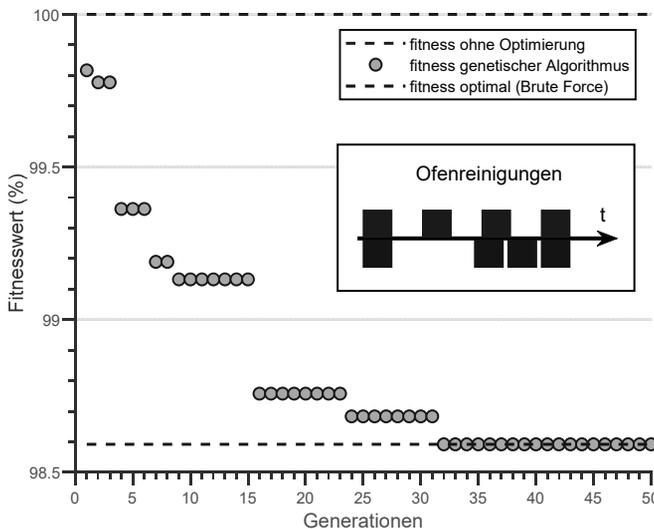


Abbildung 3: Optimierung für Szenario 1

Beim Szenario 2 treten bei der statischen Reinigungsstrategie zwei Betriebsstörungen aufgrund von Materialmangel auf. Somit beeinflussen in diesem Optimierungsproblem beide Teilziele die Zielfunktion. Der Verlauf der Optimierung ist in Abbildung 4 abgebildet. Aufgrund der möglichen Betriebsstörungen ist im Vergleich zum Szenario 1 eine größere Differenz zwischen dem Fitnesswert ohne Optimierung und dem optimalen Fitnesswert erkennbar. Dieses Szenario führt dazu,

dass in den ersten Generationen des genetischen Algorithmus vorrangig eine Lösung gesucht wird, welche zu keiner Betriebsstörung führt. Anschließend wird in den darauffolgenden Generationen bei diesen Variationen der Energieverbrauch optimiert. In diesem Fall wird lediglich ein lokales Optimum erreicht, welches keine Betriebsstörungen aufweist und knapp über dem globalen Optimum liegt. Der Fitnesswert reduziert sich durch die Vermeidung der Betriebsstörungen um etwa 10 % und der spezifische Energieverbrauch wird um 2 % verringert.

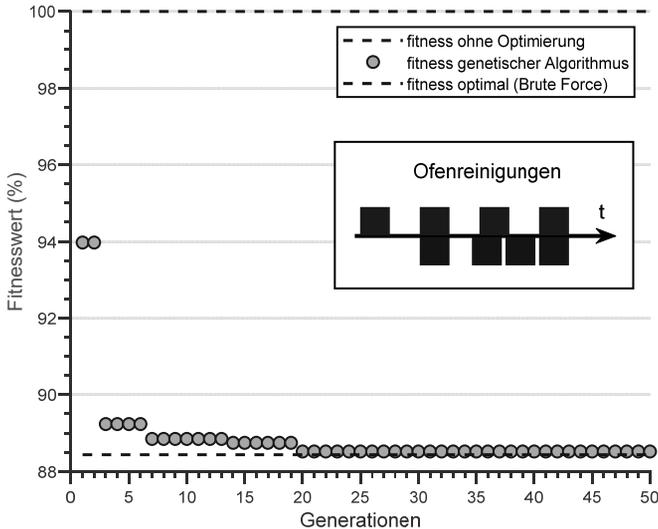


Abbildung 4: Optimierung für Szenario 2

Es zeigt sich, dass der genetische Algorithmus für dieses Optimierungsproblem geeignet ist. Die Optimierung der Ofenreinigungen bietet insbesondere in kritischen Betriebsituation ein großes Potential: Zum einen können Betriebsstörungen vermieden werden, zum anderen haben die Ofenreinigungszeitpunkte bei hoher Materialnachfrage auch einen erhöhten Einfluss auf die Energieeffizienz.

6 Ausblick

Um die Optimierung durch den genetischen Algorithmus zu verbessern, können die Optimierungsparameter noch über eine Parameterstudie mit Gittersuche angepasst werden. So können Parameter wie Populationsgröße, der Anteil des Crossovers und der Elite-Parameter weiter auf das vorliegende Optimierungsproblem abgestimmt werden. Des Weiteren kann der Einfluss der Diskretisierung des Zeitintervalls der möglichen Ofenreinigung untersucht werden. Während der Brute Force Ansatz hier bei feinerer Diskretisierung nicht mehr praktikabel eingesetzt werden kann, gilt es, insbesondere den genetischen Algorithmus oder weitere Algorithmen zu testen.

Der Einsatz einer simulationsbasierten Optimierung für Gussbetriebe konnte anhand dieses Beispiels gezeigt werden, welches aufgrund der begrenzten Ofenreinigungen jedoch nur einen kleinen Lösungsraum bietet. Ein wesentlich größeres

Einsparpotential bietet die Optimierung der Beschickung der Schmelzöfen, welche im Rahmen weiterer Untersuchungen zur simulationsbasierten Optimierung betrachtet werden soll. Im Rahmen dieses Vorhabens soll die Rechenzeit des Simulationsmodells weiter reduziert werden, indem das Modell in eine eigenständige ausführbare Datei kompiliert wird. Zudem sollen weitere Optimierungsmethoden in Betracht gezogen werden.

Literatur

- Bayerisches Landesamt für Umweltschutz (LfU): Effiziente Energieverwendung in der Industrie – Teilprojekt „Metallschmelzbetriebe“ Effiziente Energienutzung in Nicht-Eisen-Metall-Schmelzbetrieben. Augsburg, 2005.
- Dentel, A.; Schlüter, W.; Schmidt J.: Vorrichtung und Verfahren zur Verteilung von Reinigungsintervallen bei n Schmelzöfen in einem Schmelz- und Druckgussbetrieb. Hochschule für angewandte Wissenschaften Ansbach (Ansbach, DE), Technische Hochschule Nürnberg Georg Simon Ohm (Nürnberg, DE), 2018. DE102016119702
- Dettelbacher, J.; Schlüter, W.: Simulationsgestützte Optimierung des Materialflusses in einem Aluminium-Gussbetrieb. In: Deatcu, C; Lückcrath, D; Ullrich, O.; Durak, U. (Hrsg.): Proceedings ASIM SST 2020, S. 343–348.
- Gutenschwager, K.; Rabe, M.; Spieckermann, S; Wenzel, S.: Simulation in Produktion und Logistik. Berlin: Springer 2017.
- Herrmann, C.; Pries, H.; Hartmann, G.: Energie- und ressourceneffiziente Produktion von Aluminiumdruckguss. Berlin: Springer 2013.
- Roemer, A. C.; Strassburger, S.: A review of literature on simulation-based optimization of the energy efficiency in production. In: Winter Simulation Conference, 2016. S. 1416–1427.
- Schlüter, W; Henninger, M.; Buswell A.; Schmidt, J.: Schwachstellenanalyse und Prozessverbesserung in Nichteisen-Schmelz- und Druckgussbetrieben durch bidirektionale Kopplung eines Materialflussmodells mit einem Energiemodell. In: Wenzel S.; Peter, T. (Hrsg.): Simulation in Produktion und Logistik, Kassel, 2017, S. 19–28.
- Sobottka, T.: Eine anwendungsorientierte simulationsbasierte Methode, unter Berücksichtigung von Energieeffizienz, in der optimierenden Planung von Produktion und Logistik. Dissertation TU Wien. 2017.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

A Criteria-based Database for Research and Applications of Energy-oriented Simulation in Production and Logistics

Eine Datenbank für klassifizierte Forschungs- und Anwendungsberichte zu energieorientierter Simulation in Produktion und Logistik

Johannes Stoldt, ABB Stotz-Kontakt GmbH, Heidelberg (Germany),
johannes.stoldt@de.abb.com

Bastian Prell, TH Wildau, Wildau (Germany), bastian.prell@th-wildau.de

Markus Rabe, TU Dortmund, Dortmund (Germany), markus.rabe@tu-dortmund.de

Sigrid Wenzel, Universität Kassel, Kassel (Germany), s.wenzel@uni-kassel.de

Sebastian Thiede, University of Twente (Netherlands), s.thiede@utwente.nl

Abstract: Within the Section “Simulation in Production and Logistics” of the association for simulation in the German-speaking area (ASIM), the relevance of energy-related considerations led to the foundation of a dedicated scientific workgroup. One of their primary results is a comprehensive literature overview of related approaches from research and industrial perspectives. The experts cooperated to identify, evaluate, and classify relevant literature. Interim findings have been published with special emphasis on logistics and manufacturing. Recently, work was focused on converting the findings into a structured database, to foster the involvement of other experts by providing consistent access to all results in the public domain and to provide means to submit new contributions proactively to the group. Thus, the workgroup intends to raise awareness and motivate other researchers to join the effort. This paper summarises the rationale and methodology of the activities as well as the provision as a public database.

1 Introduction

Simulation has been established as a tool that is of great value in both planning new and revising existing processes in production and logistics. The method and the respective software suites have evolved to cover a wide area of applications within different industrial branches and planning scopes. When the societal awareness for climate change grew, it was natural that this would be reflected in research and application of simulation. A primary reason for this shift is the environmental impact of the industrial and the transport sectors. As of 2016, the operations of these two

sectors caused approx. 24 % and 16 % of all yearly greenhouse gas emissions worldwide (Ritchie and Roser 2020).

Initial works that integrated energy-related considerations into simulation studies in the production and logistics sectors were primarily focusing on the economic effects of energy utilization. This came at a time (from approx. 2005 onward) when electricity and oil prices were rising to the extent that energy costs became a notable issue especially for energy-intensive manufacturers as well as logistics operators. In parallel, the awareness for the overall environmental impact grew in some industries, which also prompted actions to increase transparency on and reduce the footprint of economic operations. In addition to the afore-described energy efficiency and eco efficiency, more recent works focus on energy flexibility. The latter focusses on aligning industrial energy demands with volatile, decentralised energy supply (e.g. from renewable energy sources) (Uhlig et al. 2018).

The Section “Simulation in Production and Logistics (SPL)” of the association for simulation in the German-speaking area (Germany, Switzerland, and Austria) (ASIM), responded to the importance of these developments by founding the “Workgroup on the Investigation of Energy-related Influences in SPL” in 2014. It has gathered an extensive and structured collection of relevant works to shed light on findings of various groups or organisations as well as on knowledge gaps (Wenzel et al. 2017; Uhlig et al. 2018; Poeting et al. 2019). To make these results even more widely available and attract further contributions, the workgroup has recently decided to publish a database of its findings on research and applications of energy-oriented simulation in production and logistics, which is structured by a set of classification criteria. The database will be made accessible through a specific website that allows for filtering, according to specific criteria in terms of content keywords, year of publication, authors or institutions, and others. This paper is intended to provide an overview of the workgroup’s results, the methodology of the literature review conducted, and the realization of the new online database. The structure of the paper will follow these main topics and conclude with a brief summary.

2 Overview on Workgroup’s Results

The initial impetus for founding the “Workgroup on the Investigation of Energy-related Influences in SPL” within the ASIM Section SPL came in the wake of the section’s bi-yearly conference in 2013. It consisted and still consists of a group of simulation software vendors, industrial users, and researchers. During its constituting meeting one year later and in the following assemblies, the participating parties shared insights into their respective previous and current case studies that related to the theme of the workgroup. Moving onward from there, discussions took place to decide on the scope of future work. It was agreed that standardization of vocabulary and procedures was desirable but would require an in-depth understanding of existing and documented works in the field. This prompted the group to compile an exhaustive overview on research and applications of simulation based on published articles.

The remainder of this section will review the primary findings of the workgroup’s previous publications (Wenzel et al. 2017; Uhlig et al. 2018; Poeting et al. 2019). Additionally, statistics will summarise the distribution of identified works considering the criteria defined as part of the methodology that was applied in the literature review (cf. Section 3).

Previous publications have shown the diversity of knowledge that is comprised in such literature overviews: By plotting the number of literature entries that were determined to be relevant by the workgroup categorised by publishing year, Uhlig et al. (2018) showed an increased interest in this topic's exploration, starting from 2011. Apart from the quantitative analysis, causal interpretations have been made possible. Mutual dependencies between applied key performance indicators and the level of detail for modelling could be discerned as well as their effect on additional requirements for simulation input data. Furthermore, particularities resulting from either a logistics or manufacturing scope of the case studies were investigated by Poeting et al. (2019). These showed, e.g., that fewer documented simulation applications could be found for logistics than for manufacturing. Reflecting the workgroup's aim to identify actors in the field from academia or industry, distinct inferences were made, e.g. by mapping the origin of the included literature's authors by country and region within Central Europe (cf. Wenzel et al. 2017). The initial Germany-centred focus of the workgroup's findings came naturally from the members' ASIM affiliation and their self-conception as a German-speaking institution, which made German publications more available. However, the workgroup diligently worked to meet expectations of the internationally linked research community by incorporating English publications wherever possible. A recurring topic that has been discussed by the members was the ambivalence of usually discrete simulation of production or logistics processes and the usually continuous flow of resources or energy. Considering this aspect, Uhlig et al. (2018) referred to Thiede (2012) concerning typical paradigms for the simulation of energy flows in manufacturing systems. Poeting et al. (2019) built on this classification for simulations with a focus on logistics and underpinned it with empirical findings.

As extension to the previous analyses, Fig. 1 **Fehler! Verweisquelle konnte nicht gefunden werden.** gives a combined overview of those paradigms and the related architectures to the simulation tools that were used in publications over the years. Quite clearly, discrete event simulation (DES) with integrated consideration of energy turns out to be the most dominantly used architecture. These approaches were mainly implemented with tools like Plant Simulation and Arena. The strong increase of publications with this approach especially after 2012 can certainly also be explained through the introduction of standard energy features with Plant Simulation version 11. However, when it comes to other architectures, a diversity of further tools comes into play. This includes, for example, multilevel approaches that bring together different elements into overall system simulations (e.g. a factory with different subsystems). This can either take place within one tool, e.g. in hybrid tools that can incorporate different modelling paradigms like AnyLogic (e.g. Thiede 2012) or Matlab (e.g. Pawletta et al. 2017). The alternative is to couple several tools and enable communication through a middleware interface or directly (e.g. Peter and Wenzel 2017).

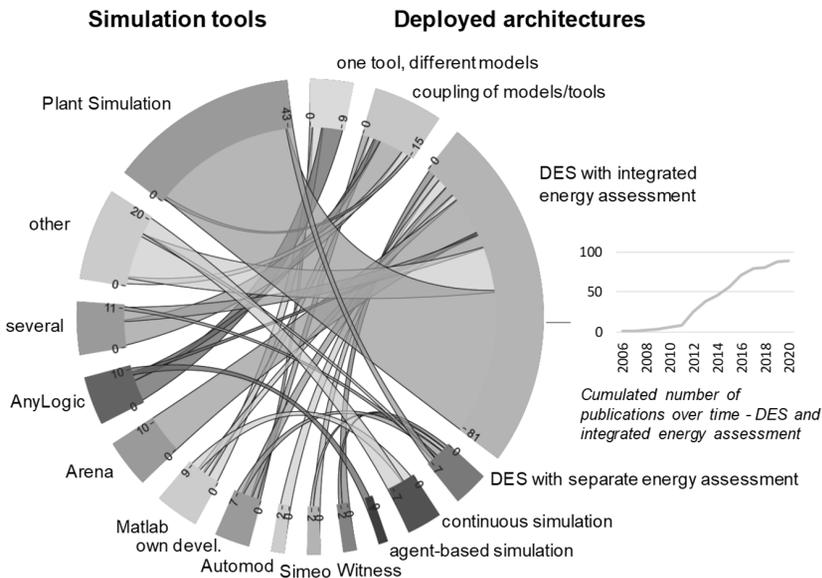


Figure 1: Deployed architectures in relation to simulation used in documented publications (Note: “own devel.” refers to self-developed simulation software)

3 Methodology of the Literature Review

A primary goal of the research method *literature review* is to generate new insights based on other primary works. Various general approaches can be distinguished that differ primarily in the way they try to incorporate quantitative data from identified articles. Meta-analyses aim to mathematically integrate the results of different studies. Systematic reviews focus on summarising quantitative studies regarding the effects of certain measures. Scoping reviews are intended to cover not just scientific literature but all relevant publications, for instance, to identify knowledge gaps or the extent of existing research (Sturma et al. 2016). The workgroup applied the latter method for its purposes.

The general process applied follows the methodology for scoping reviews proposed for use by Colquhoun et al. (2014). It has been described in Uhlig et al. (2018) and is summarised in Fig. 2. The depicted list of keywords is not applied automatically nor strictly but rather manually and for approximate matches in order to filter for relevant works. As of May 2021, a total of 207 publications from 29 journals and 47 conference proceedings as well as other publication channels (e.g. thesis works) have been identified and added to the scoping review’s body of literature through the application of this methodology.

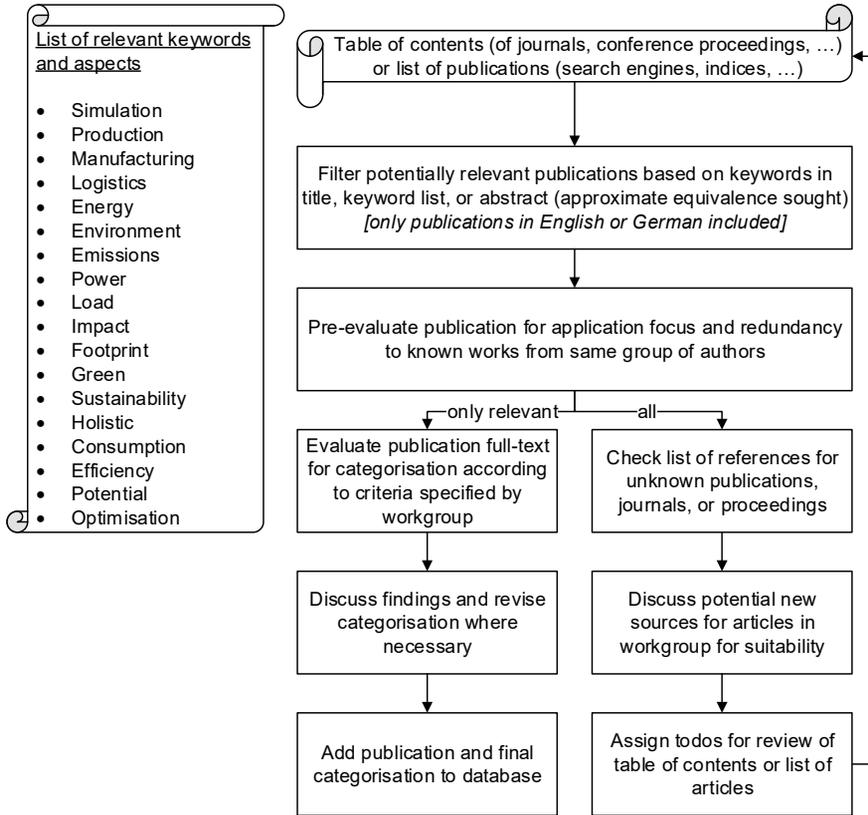


Figure 2: Overview of process for scope review

Three general options to facilitate the involvement of third-party actors exist. Firstly, the workgroup can be notified of journals, journal issues, conference proceedings, and further compendium works that might include potentially relevant publications. Secondly, possibly relevant publications can be brought to the attention of the workgroup, directly. Lastly, fully evaluated publications (i.e. a suggestion for a publication and its classification according to the specified criteria) can be forwarded to the workgroup. They will be discussed in regular meetings before they are added to the database. It is foreseen that the public frontend of the online database will feature ways to support these options. Contributions are presently encouraged to be provided directly to members of the workgroup or through the website, once completed. The mandatory step “discussion in the workgroup” will serve to maintain consistency and suitability of added publications throughout the future extension of the database, following the scoping review methodology.

4 Database and Website

In order to enable an efficient maintenance and exploitation of the knowledge that has been acquired by the workgroup, a web-based service is under development. This

section describes the envisaged general structure of the database, the planned access mechanisms, as well as the major technical characteristics.

4.1 Underlying Data Structure

The data are structured in a relational database model with the following basic entities:

- PAPER: title, language, source, year, abstract, administrative information (e.g. date generated, date changed).
- AUTHOR: first name and surname.
- INSTITUTION: name of institution (and department), city, country.
- CONTRIBUTOR: sequence number in author list.
- KEY: criteria to classify publications (cf. Wenzel et al. 2017) regarding energy-specific key performance indicators (KPI), process scope, production structures, production type, industrial sector, level of detail, integration architecture, simulation tool, and added input data.
- KEYWORD: key class (i.e. characteristics for each criterion for the classification of case studies; e.g. energy consumption per system or per system element over time for criterion energy KPI).

In addition to these basic entities, there are a number of relational entities, which comprise PAPER–CONTRIBUTOR, CONTRIBUTOR–AUTHOR, CONTRIBUTOR–INSTITUTION, PAPER–KEYWORD; and KEYWORD–KEY (Fig. 3). The list of keys is a fixed catalogue, and a change requires access to the underlying database. All keywords are related to exactly one of these keys, and each paper may be associated with several keywords, also allowing more than one keyword of the same key category.

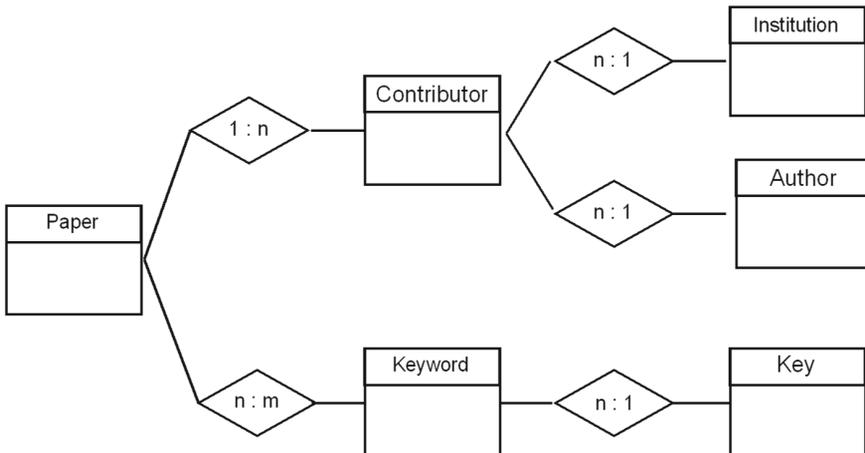


Figure 3: Simplified Entity-Relationship Model of the database (cardinality in Chen-notation)

4.2 Maintenance and Query Functionalities

The functionalities that are provided through the web interface can be differentiated into administration, maintenance of the data, and query. Administrative functions, like user right assignment or backup, are not discussed here. Maintenance functions enable privileged users to load any of the basic entity entries based on a unique identification (ID) code, which is automatically assigned to all entries. The entry (except the ID) can be changed in the interface and sent back to the database for permanent storage. Furthermore, relations can be added or edited. This function is supported in the user interface showing the content (e.g. the name of the key characteristic) instead of the ID (cf. Fig. 4). Furthermore, text-element-based filter functions will ease the search in large tables.

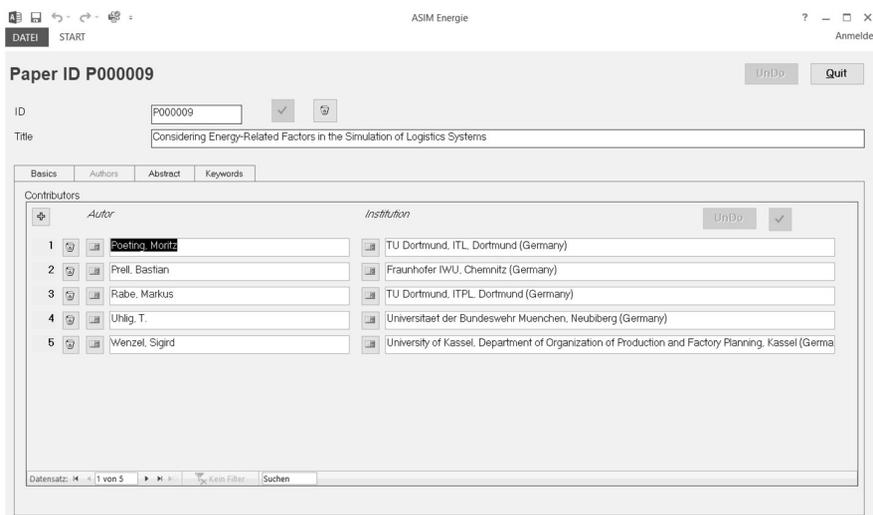


Figure 4: Screenshot of the prototypical maintenance graphical user interface (GUI)

In order to keep the development effort reasonable, a Microsoft Access interface is used for the maintenance functions. A later full implementation of these functions on the web server is possible (then, both mechanisms would work in parallel), if suitable resources are provided. For performance reasons, the data are cached in the local Access database and only cross-checked during editing operations and updated when changes are conducted in the local interface.

Query functions in the public interface work (i) through text elements of the entry or (ii) through keywords. In the first case, the user may enter any piece of text or a regular expression (RegEx), either in general or associated with one of the basic entities. Thus, searches can be limited to, e.g. titles or institutions. A search in the keywords or keys is not possible with this function. In the second case, the user will select a number of keywords. Multiple keywords for the same key are allowed, as well as a combination of keywords from different keys. In the first approach, all items will be found that

match any of the specified keywords. Depending on the usage, the workgroup might later consider more sophisticated mechanisms for the selection of items.

The result of a query is in the first step the number of items found. Depending on the user's rights, functions for downloading an overview or the resulting references, a bibliographic list in bibTex format, or the full available information can be generated.

4.3 Technical Characteristics and Roles

In order to enable a low-barrier access, the complete system will be web-based and deployed on the web server of the ASIM Simulation in Production and Logistics Conference (ASIM 2021). The data are held in a mySQL database, with the interface coded in PHP7, applying standard HTTP POST mechanisms for the data transfer (Fig. 5).

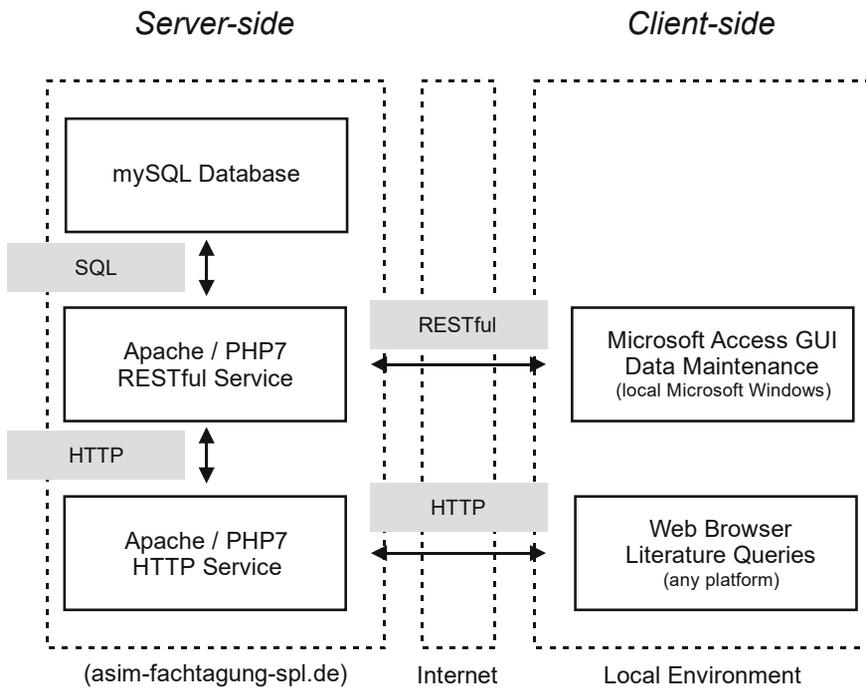


Figure 5: General architecture and deployment of the application

In detail, there is a RESTful service implemented as a shell for the database. The RESTful service comprises standard functions for adding, replacing, or deleting basic entries and further functions to relate such entries, e.g. to relate an author to a paper or to an institution. These services are conceptualised as integrity-guaranteeing, i.e. changes are only permitted if the database's integrity is not violated by the operation. The same RESTful service will also conduct the authorization approval activities, checking the user's permission for specific activities on the database.

In the first preliminary version, the access to the complete page will be restricted by a suitable username and password for authorised workgroup members as beta users. In a second stage, specific differentiated user rights will be assigned. The envisaged user groups are:

- EDITOR, authorised to change the content and add new information, as well as to accept contributed information.
- CONTRIBUTOR, authorised to propose changes to the database that are visible to editors, but not to the public.
- REQUESTER; authorised to query the database by text items or keywords, and to access the full information available, including full references and abstracts.
- GENERAL, authorised to query some overview information, e.g. the institutions that can be associated with specific keywords. The categories of the open information are not yet decided.
- MANAGER, authorised to add or delete users and user rights.
- ADMIN, comprising all rights described above plus the right to conduct changes in the database without consistency approval, thus allowing for corrective actions in the case of technical failures.

Obviously, there will be a user access control for most of these roles. While roles like EDITOR or MANAGER will require the permission of the workgroup, the rules for some other rules will have to be decided. For practical reasons, at least CONTRIBUTORS might be requested to register with their institution and email address. Also, REQUESTERS could be asked to make themselves known before full access is granted, in order to give the workgroup a chance to get in touch with them and to monitor the interest in their work. This seems to be a fair approach given that the workgroup publishes its knowledge in such an open way.

It should be noted that – through the implementation as a RESTful service – there is an application programming interface (API) that could also be used by any evaluation software, independently of the user interfaces that are described in this paper.

5 Summary

This paper gives a short overview on the main work and the results of the "Workgroup on the Investigation of Energy-related Influences in SPL". The literature overview, which has been created in the workgroup since 2014, was structured via keywords, so that all contributions could be classified systematically albeit manually and analysed laboriously so far. In order to make the extensive results more usable and internationally visible, to be able to easily add further relevant third-party literature, and to provide efficient features for data analyses, a database has been designed. This database is currently in the realisation phase. The implemented database schema was briefly presented herein. The backend functions have already been realised using Microsoft Access.

Future work will focus on the implementation of data query and maintenance functions in a web frontend as well as the establishment of an organisational evaluation process for new third-party contributions within the workgroup. Furthermore, it is planned to carry out and publish more analyses of the recorded publications using the database. The long-term goal is an international visibility of the database including a sustainable and simple organizational concept. The access to the database will be

available on the website of the ASIM Section "Simulation in Production and Logistics" under the URL <http://asim-fachtagung-spl/asimenergy>.

Acknowledgement

We would like to thank the members of the ASIM Section's Simulation in Production and Logistics (SPL) workgroup on the Investigation of Energy-related Influences in SPL for their support over many years in researching and evaluating relevant publications as well as in making the database usable and applicable.

References

- ASIM: Conference on Simulation in Production and Logistics. <http://www.asim-fachtagung-spl.de>, accessed May 3rd, 2021.
- Colquhoun, H.L.; Levac, D.; O'Brien, K.K.; Straus, S.; Tricco, A.C.; Perrier, L.; Kastner, M.; Moher, D.: Scoping reviews: time for clarity in definition, methods, and reporting. *Journal of Clinical Epidemiology* 67 (2014) 12, S. 1291–1294.
- Pawletta, T.; Schmidt, A.; Junglas, P.: A Multimodeling Approach for the Simulation of Energy Consumption in Manufacturing. *Simulation Notes Europe* 27 (2017) 2, S. 115–124.
- Peter, T.; Wenzel, S.: Coupled Simulation of Energy and Material Flow using Plant Simulation and MATLAB Simulink. *Simulation Notes Europe* 27 (2017) 2, S. 105–113.
- Poeting, M.; Prell, B.; Rabe, M.; Uhlig, T.; Wenzel, S.: Considering Energy-Related Factors in the Simulation of Logistics Systems. In: 2019 Winter Simulation Conference (WSC). National Harbor, MD, USA: IEEE 2019, pp. 1849–1858.
- Ritchie, H.; Roser, M.: Sector by sector: where do global greenhouse gas emissions come from? In: *Our World in Data*. <https://ourworldindata.org/ghg-emissions-by-sector>, accessed February 6th, 2021.
- Sturma, A.; Ritschl, V.; Dennhardt, S.; Stamm, T.: Reviews. In: Ritschl, V.; Weigl, R.; Stamm, T. (eds.): *Wissenschaftliches Arbeiten und Schreiben*. Berlin, Heidelberg: Springer Berlin Heidelberg 2016, pp. 207–221.
- Thiede, S.: *Energy Efficiency in Manufacturing Systems*. Berlin, Heidelberg: Springer 2012.
- Uhlig, T.; Wenzel, S.; Peter, T.; Stoldt, J.; Schlegel, A.; Jósваи, J.: Considering Energy in the Simulation of Manufacturing Systems. In: Rabe, M.; Juan, A.A.; Mustafee, N.; Skoogh, A.; Jain, S.; Johansson, B. (eds.): *Proceedings of the 2018 Winter Simulation Conference*. Gothenburg: IEEE 2018, pp. 3275–3286.
- Wenzel, S.; Peter, T.; Stoldt, J.; Schlegel, A.; Groß, G.; Pitsch, H.; Rabe, M.; Seewaldt, M.: Betrachtungen energetischer Einflussfaktoren in der Simulation in Produktion und Logistik: Eine Literaturanalyse. In: Wenzel, S.; Peter, T. (eds.): *Simulation in Produktion und Logistik 2017*. Kassel: kassel university press 2017, pp. 9–18.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Simulation-based Assessment of Energy Demand and Costs Associated with Production Scrap in the Battery Production

Simulationsbasierte Bewertung des Energiebedarfs und der Kosten unter Berücksichtigung des Produktionsausschusses in der Batterieproduktion

Gabriela Ventura Silva, Matthias Thomitzek, Tim Abraham, Christoph Herrmann, TU Braunschweig, Battery LabFactory Braunschweig, Braunschweig (Germany), g.ventura-silva@tu-braunschweig.de, matthias.thomitzek@tu-braunschweig.de, tim.abraham@tu-braunschweig.de, c.herrmann@tu-braunschweig.de

Abstract: The shift in the mobility sector towards electric vehicles is responsible for a growth in the market demand for lithium-ion batteries. To follow this trend, the current 200 GWh global production capacity of lithium-ion batteries will present an annual increase up to 300 GWh in the next years. Characterized by an energy-intensive process chain and high material costs, battery production is sensitive to scrap. Current works on energy and cost assessment in the battery production consider scrap rate based on static values derived from historical production data. Thus, there is a lack of works describing the influence of different scrap rates on the process chain dynamics, e.g. machine states and utilisation capacity. To tackle this challenge and contribute to a more cost competitive and environmentally sustainable battery production, this work presents a simulation-based methodology to assess the indirect and direct energy demand and costs associated with production scrap. Based on a combined discrete event and agent-based simulation, scenarios with different scrap rates are simulated. The results show that the effects associated with production scrap varies for each process and are influenced by various factors, e.g. process characteristics, position in the process chain, material costs and energy demand.

1 Introduction

Lithium-ion batteries offer a wide range of applications, with the mobility sector accounting for more than 60% of the 200 GWh global demand in 2019. To follow the electromobility growth, studies predict that the global capacity of production of lithium-ion batteries will present an annual increase up to 300 GWh in the next years (VDMA Battery Production 2020). Due to its energy-intensive process chain, raw material extraction and manufacturing are responsible for up to 45% of the battery

cradle-to-gate environmental impacts (Drachenfels et al. 2021). Besides the environmental impact, production is also the main cost driver. Here material is a decisive aspect, accounting for up to 70 % of the costs of a single battery (Kwade et al. 2018). Therefore, a more environmentally sustainable and cost competitive battery cell production depends on a material and energy efficient production. The reduction of production scrap, i.e. material waste intrinsic to the process or resultant from material flaws, increases the material efficiency and reduces the production costs. For large scale production, production scrap rate varies from 5 to 10% (Drachenfels et al. 2021). Different works in the battery production context with focus on energy efficiency (Thomitzek et al. 2019a; Weeber et al. 2020; Erakca et al. 2021) and cost estimation (Nelson et al. 2019; Schünemann 2015; Mauler et al. 2021) consider production scrap in their models and calculations. Nevertheless, the influence of production scrap on the process chain dynamics (e.g. machine states and utilisation capacity) has not yet been discussed. Simulation-based approaches represent a well-established tool for understanding complex relationships and dynamics of process chains, and have already been applied in the analysis of material and energy flows as well as production improvements (Schönemann et al. 2019; Weeber et al. 2020). Against this background, this work proposes a combined discrete event and agent-based simulation approach to (i) dynamically study the effect of different scrap rates on a process chain level and (ii) provide an identification of critical processes from energetic and economic perspectives.

2 Theoretical Background

2.1 Lithium-ion Battery Production

The battery cell production is characterized by a rigidly interlinked process chain with numerous heterogeneous process steps. In general, the process chain can be divided into electrode production, cell production and cell conditioning. However, slight variations might occur in the battery process chain depending on the respective process technology and the battery cell design, e.g. pouch, cylindrical or prismatic. In the electrode production, anodes and cathodes are produced in batch and continuous processes, located in separate production lines to avoid contamination (Schünemann 2015). After a dry and wet mixing process, the respective material suspension is coated and subsequently dried to produce a composite structure. Afterwards, anode and cathode coils are calendered to reduce their porosity, and slit to width and length before they enter the dry room for the cell production, characterized by discrete processes. First, the coils are further cut to single electrode sheets. For pouch cells, the individual electrode sheets are stacked together with a separator. The electrode-separator assembly is contacted internally and afterwards inserted into a pouch bag housing. The housing is then filled with electrolyte and subsequently sealed. In the cell conditioning, the formation and aging of the battery cells are conducted (Kwade et al. 2018). Scrap rate information in the literature is diverse and limited, usually derived from input-output rates and historical data. Based on previous publications, Drachenfels et al. (2021) present variations in scrap rates according to production scales, e.g. 5 - 20 % for small and 5 - 10 % for large factories. Nelson et al. (2019) present process-specific scrap rates, varying from 1 to 8 % according to the process characteristics. Schünemann (2015) proposes even lower rates, e.g. 1 % for mixing

process and 0.2 % for stacking. Production scrap rate has also a major influence on production energy demand and costs.

2.1.1 Energetic Perspective

The battery cell production requires a significant amount of electrical energy, especially caused by its energy-intensive processes, e.g. coating/drying, calendaring and formation (Thomitzek et al. 2019a). In addition, the technical building services (TBS), which provides the necessary environmental conditions, also contributes to a significant share of the total energy demand (Wessel et al. 2021). The literature reports large variations in energy demand per energy storage capacity at industrial scale, ranging from 47 to 162 Wh per Wh (Erakca et al. 2021). These variations can be explained by the production scale, the complex and dynamic combination of continuous and discrete processes as well as the selected process parameters and boundary conditions (Drachenfels et al. 2021; Thomitzek et al. 2019b). The assessment of energy considering scrap rates has been shown in different works. Thomitzek et al. (2019a) present a material and energy flow analysis based on input-output ratios and the measured energy demand. Weeber et al. (2020) propose a simulation on process chain and process levels to assess the overall energy demand. Wessel et al. (2021) provide an analysis of energy demand due to scrap for a pilot line based on production data. The results show energy-intensive processes as critical when analysing energy demand due to scrap. Although the scrap rate has been considered in many works, it was usually limited to static values based on the average of production data. Thus, it is necessary to analyse closely the influence of scrap rates in battery production on the energy demand, considering process chain dynamics.

2.1.2 Economic Perspective

Material costs represent the largest share of battery production costs. Kwade et al. (2018) present in a cost breakdown that 74.9 % of the costs are caused by material and 3.1 % by energy demand. Duffner et al. (2021) show the share of the various costs for an optimization scenario with materials (77 %), machine depreciation (8 %), production scrap (6 %) and energy (3 %) being the largest ones. Due to the importance of material efficiency for a more competitive production, production scrap has been considered in different cost estimation models. A simulation-based approach to assess the importance of economy of scale on production costs is presented by Mauler et al. (2021) which considers production bottlenecks and end-of-line scrap rate. With regard to process-specific costs, Kwade et al. (2018) declare that processes further down the process chain are more cost sensitive, since they embody the value added by the previous processes. Duffner et al. (2021), on the other hand, mention an electrode production process (coating) as critical. The review on cost models presented by Duffner et al. (2020) lists many works which consider process-specific parameters in their estimations. However, none of them analyses process chain dynamics when defining scrap and energy related costs. Based on the relevance of the material efficiency to the battery cell costs, it is fundamental to consider the economic influence of different scrap rates.

2.2 Simulation Approaches for Process Chain Dynamics

Simulation is a consolidated approach to analyse different production scenarios and process chain performance (Schönemann et al. 2019). In the battery production

context, it has also been identified as an effective tool to assess and analyse energy demand for different production and machine configurations (Thomitzek et al. 2019b; Weeber et al. 2020). Discrete event simulation (DE) enables a better understanding and reproduction of material and energy flows within the production as well as provides insights on dependencies between processes. Agent-based simulation (AB) enables to describe elements, e.g. machines or products as unique agent, study their interactions, and store specific data. The use of DE and/or AB to analyse production throughput, machine availability, and process-specific energy demand in the battery context was already proposed by different works (Weeber et al. 2020; Thomitzek et al. 2019b; Schönemann et al. 2019). When considered, scrap rate is described as a process characteristic based on static data to support analysis of input and output flows between processes. Therefore, there is a lack of works with focus on the production scrap rate and its influence on the process chain dynamics.

3 Methodology

A simulation-based methodology was developed to study the influence of different production scrap rates on the process chain dynamics with focus on energetic and economic perspectives, as described in Figure 1.

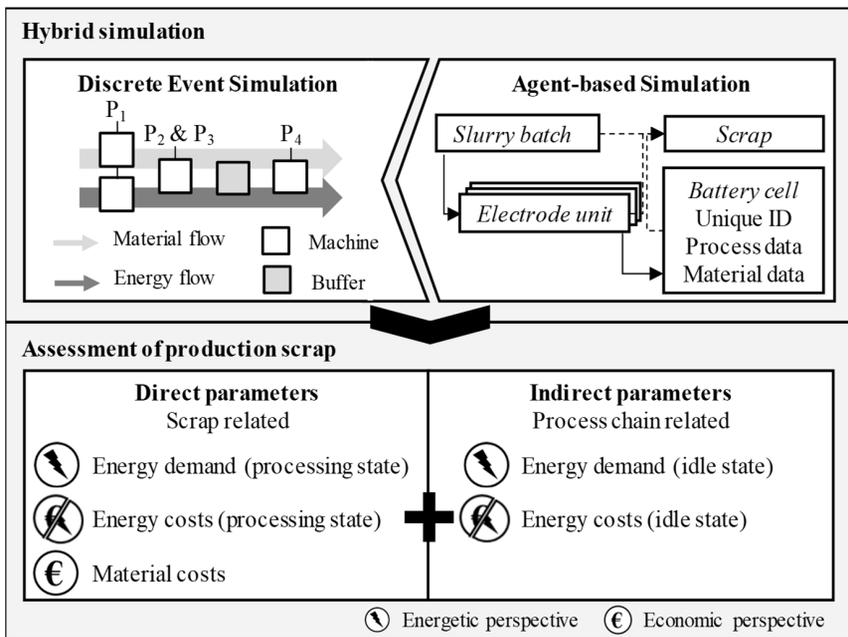


Figure 1: Simulation-based methodology to assess the effects of production scrap.

3.1 Hybrid Simulation

The first methodology part is a python-programmed hybrid simulation which combines DE and AB approaches. The focus of the DE is to reproduce material and energy flows of a process chain, consisting of the following elements: machine,

process and buffer. A process can be executed by more than one machine and a machine can be associated to more than one process. In addition, it is possible to have buffers to store finished parts. Otherwise, the finished part is temporarily stored in the machine, until it is taken to the next process. A machine presents five states: *off*, *ramp-up*, *idle*, *processing*, and *failure*. *Off* is the machine state either at the beginning of the simulation or after breakdowns. The *ramp-up* state starts after the machine is switched on until it is ready to produce. A machine is in *idle* state before processing, i.e. waiting for input material and machine availability. The *processing* state represents the production itself and, in some cases, the execution of sub-processes and storage of finished parts. Lastly, a machine may fail during processing. Average power consumption and duration are inputs defined by the user and associated to each machine state. An overview on the conditions for state changes and power consumption over time are shown in Figure 2.

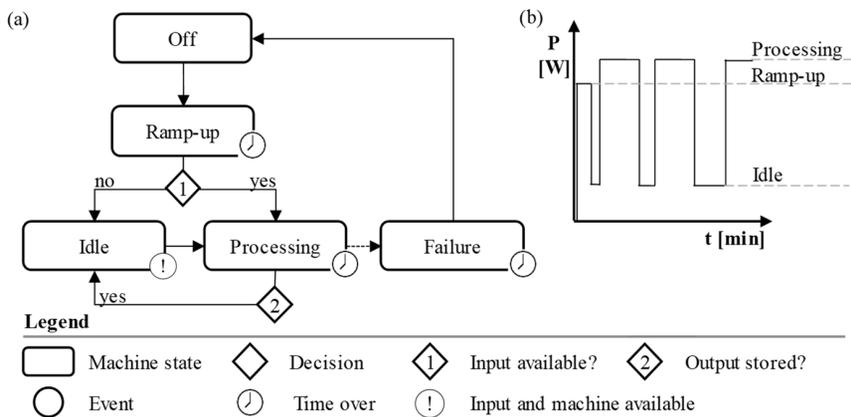


Figure 2: (a) Machine state chart (b) Machine energy profile based on the duration and average power consumption of the different states.

The conditions for each state change are represented in Figure 2a. With exception of the off state, all state changes are triggered by an event. *Ramp-up* and *failure* events are time-regulated, based on the user inputs regarding the average and variation of the process duration. The *processing* state is time-regulated and additionally considers the storage of finished materials. The *idle* state is controlled by two events: input and machine availability. The last condition is especially relevant for machines associated to more than one process. The timestamp of changes in the machine states as well as power consumption values result in the energy profile shown in Figure 2b.

The AB simulation focuses on the agents, e.g. slurry batches, electrodes and battery cell. During the simulation, agent-specific information regarding the process (e.g. timestamp and energy demand) and the material (e.g. input, output and scrap ratios) is stored. The interaction between agents is achieved by the possibility to combine them. For example: a battery cell contains various cathodes, these cathodes originate from the same slurry batch. The agents are either located in a buffer or a machine, which provides the integration of both DE and AB approaches. A timestamp is stored whenever a state change in the DE triggers a change in the agent location. Further process and material-specific data, e.g. scrap and output amount as well as energy

demand are also stored within each agent. The integration of both simulation approaches provides knowledge regarding the conditions under which each agent is produced and the associated energy demand. The main program functions responsible for this integration are described in Figure 3.

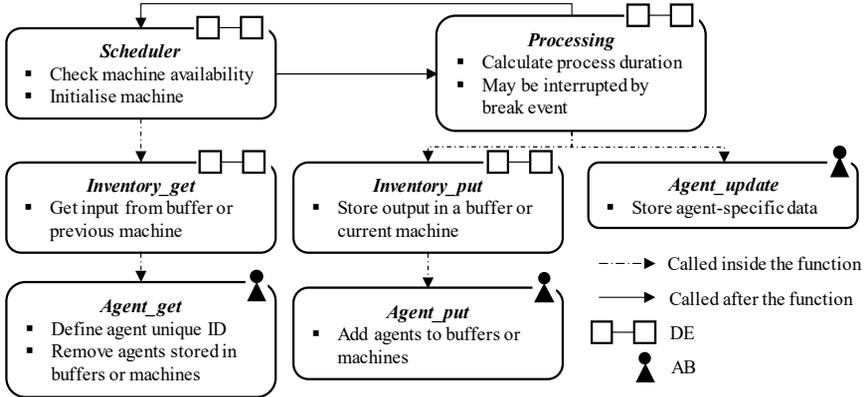


Figure 3: Program main functions for the DE and AB simulation approaches.

Scheduler is one of the main functions, responsible for initialising the machines at the simulation start. It is also called before and after processing to check the machine and input availability. The acquisition of input material and storage of finished parts are executed by the *inventory_get* and *inventory_put* functions. These functions are based on the Python package SimPy which enables an allocation of materials in a virtual container and provides, for example, the possibility to wait until a material is available. Lastly, the functions *agent_get*, *agent_put*, and *agent_update* support the AB simulation by managing the creation and location of agents as well as data storage.

3.2 Assessment of Production Scrap

The simulation results are used to assess the energetic and economic influences of different production scrap rates, considering direct and indirect parameters. Different power consumption values are associated to the machine states ramp-up, idle and processing. Energy demand during processing results from the average consumption and duration, and may be directly associated to a scrap agent. As consequence, energy demand during processing state is classified as *direct* parameter. Parameters affected by scrap on a process chain level are classified as *indirect*. Production scrap may cause changes in the material flow and affect the duration of waiting times and energy demand of machines. Therefore, energy demand in idle state is considered an *indirect* parameter. In the battery production, TBS is a major energy consumer, responsible for maintaining adequate production conditions. Since these conditions must always be achieved, independently of the throughput and scrap rate, TBS energy demand is constant and, therefore, not considered in this assessment.

A complete estimation of production costs includes fix and variable costs. Fix costs are associated to investments (e.g. machine acquisition), building, maintenance and overhead. Variable costs comprehend material, energy and labour. Since the fixed costs are strictly dependent on the production scale and are constant regardless of the production throughput and scrap rates, they are not considered in this work. Moreover,

for constant working hours and number of shifts, labour costs also remain the same. Thus, material and energy are the only costs considered in this assessment. Material and processing energy costs are classified as *direct* since they are calculated based on agent-specific information, e.g. amount of scrap and energy demand. *Indirect* parameters comprehend the ones affected by scrap on a process chain level, i.e. energy costs related to idle states.

4 Use Case: Battery Cell Production

The proposed methodology was applied to the pilot line of the Battery LabFactory Braunschweig (BLB). The energy and process parameters for the pouch cell production were automatically acquired via the SCADA system described by Turetsky et al. (2020). Since material prices for a pilot line are not consistent to the ones for a larger production scale, this use case considered the prices described in the BatPac cost model (Nelson et al. 2019). An around-the-clock production with the BLB machine capacities was simulated to investigate the dependencies and dynamics between processes, e.g. share of each machine state as well as material and energy flows. Moreover, differently from the BLB pilot line, the simulation considered separate production lines for cathode and anode production, as shown in Figure 4.

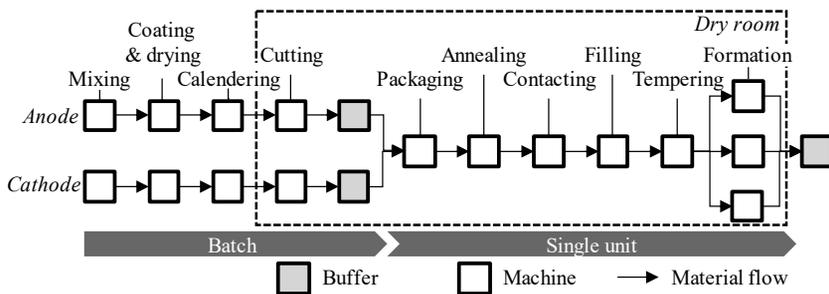


Figure 4: Simulated processes adapted from the BLB production line.

First, a one-month production with no scrap was simulated as a base scenario. Subsequently, the simulation was repeated in four scenarios with scrap rates ranging from small to large scale productions (1 %, 5 %, 10 %, and 15 %). In each scenario, the same scrap rate was considered for every process which represents, for example, a yield of 90.4 % for the 1 % scenario. For batch processes, scrap is a share of the produced batch. For single unit processes, scrap represents a complete unit.

4.1 Results and Discussion

The simulation results of all the five scenarios were assessed according to the direct and indirect parameters described in the methodology. First, the influence of scrap rate on a process chain level was evaluated by measuring the variation of indirect parameters for each scenario. The energy cost for the entire process chain associated with production scrap was calculated based on the energy demand in idle state per finished (good and scrap) part [kWh per part] of each process and the electricity price for business in Germany of 0.237 \$ per kWh. To provide a better identification of the variations, the costs for idle and processing states are compared in Figure 5.

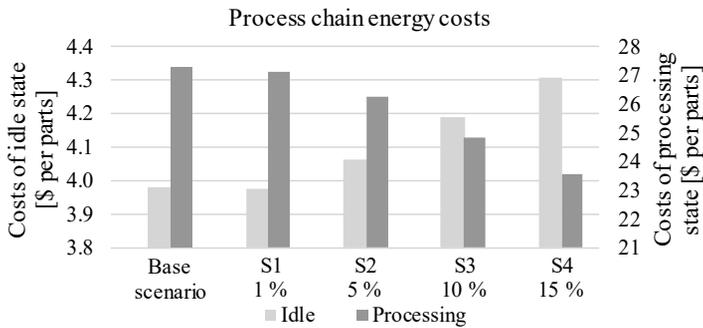


Figure 5: Costs associated with energy demand in idle and processing states for a one-month production for the base scenario and different scrap rates.

These results reinforce that a variation in the scrap rate is responsible for a change in the process chain dynamics. Since the processes are rigidly interlinked and the throughput of each single unit process is reduced by an increase in the scrap rate, processes down the process chain have to wait longer for input material. This increase in waiting times leads to higher idle state costs. The reduction of throughput at each single unit process also leads to fewer processed parts in one month and, consequently, to a reduction in processing state costs. It is also important to emphasize that these effects are not proportional to the scrap rate: for the 15 % scenario, the costs of idle state increase 8.2 % while the costs of processing state decrease 13.6 % in comparison to the base scenario. After assessing the indirect parameters, i.e. influence of scrap rate on a process chain level, the direct parameters were analysed. For this purpose, the energy demand associated with processing and material costs for each scrap agent at each process were calculated. Table 1 presents the total average of material and energy costs per finished (good and scrap) parts for each simulation scenario.

Table 1: Direct material and energy costs associated with the different scrap rates.

	S1 - 1%	S2 - 5%	S3 - 10%	S4 - 15%
Material [\$ per part]	11.05	79.90	205.36	289.16
Energy [\$ per part]	0.19	1.39	3.05	4.43
Formation finished parts	390	338	260	208

As expected, a scrap rate increase is directly related to higher material and energy costs associated with the production of these parts. Although a uniform increase is seen on the total costs per part, a deeper analysis shows variations in the critical processes for the different scenarios, as represented in Table 2. While some processes are more critical from an energetic perspective, others are more sensitive to material costs. For a 5% scrap rate, cathode calendaring and formation are the most critical processes from an energetic perspective. Cathode mixing and formation, on the other hand, are more critical with regard to material costs. Considering the combination of both aspects, the most critical ones are formation and cathode calendaring. For the highest simulated scrap rate (15%), cathode calendaring and mixing are the most critical processes from an energetic perspective; cathode mixing and filling from a material one. In total, cathode mixing and calendaring are the most critical ones.

Table 2: Comparison of the most critical processes considering direct parameters for 5 and 15 % scrap rates.

	Mixing Cathode (5%)	Calendering Cathode (5%)	Formation (5%)	Mixing Cathode (15%)	Calendering Cathode (15%)	Formation (15%)
Energy [\$]	0.07	0.17	0.19	0.50	1.24	0.31
Material [\$]	12.25	1.44	12.21	108.31	11.00	20.57
Total [\$]	12.32	1.61	12.40	108.81	12.24	20.89

Variations on a process level change in different proportions for different scrap rates. For batch processes, the number of produced batches remains the same, however, the percentage of scrap is higher. Therefore, electrode production processes are more sensitive to higher scrap rates. Moreover, cathode and anode production present different variations, since cathode production is more intense from both energy and material cost aspects. Overall, the results show that different scrap rates have dynamic effects on the process chain, altering the material flow and the shares in processing and idle times. An analysis on process level shows that processes are affected differently from both an energetic and economic perspective. The intensity of these effects is influenced by the process characteristics (e.g. batch or single unit), position in the process chain, material costs and energy demand.

5 Summary and Outlook

Material efficiency is fundamental for a more cost competitive and environmentally sustainable battery production. Current works on energy and cost estimations consider production scrap rates as static values derived from historical data and do not assess their effect on the process chain dynamics. To tackle this challenge, this work proposed a simulation-based methodology to dynamically study the effect of different scrap rates on a process chain level and provide the identification of critical processes from an energetic and economic perspective. First, a discrete event and agent-based simulation studied the material and energy flows of a one-month battery production. The results for different scenarios were analysed with focus on indirect parameters on a process chain level (e.g. energy demand and costs for idle states) affected by production scrap. In addition, parameters with a direct relation to production scrap (e.g. material costs and processing energy) were assessed. The assessment showed that the effects are different for each process and influenced by various factors, e.g. process characteristics, position in the process chain, material costs and energy demand. Future works will study the effect of process-specific scrap rates to define acceptable tolerances and support the planning of quality gates.

Acknowledgement

The authors gratefully thank the Federal Ministry of Education and Research for funding this work in the project Sim4Pro (grant number 03XP0242B) within the ProZell Cluster.

References

- Drachenfels, N.; Engels, P.; Husmann, J.; Cerdas, F.; Herrmann, C. (2021): Scale-Up of Pilot Line Battery Cell Manufacturing Life Cycle Inventory Models for Life Cycle Assessment. In: *Procedia CIRP* 98, S. 13–18.
- Duffner, F.; Mauler, L.; Wentker, M.; Leker, J.; Winter, M. (2021): Large-scale automotive battery cell manufacturing: Analyzing strategic and operational effects on manufacturing costs. In: *International Journal of Production Economics* 232, S. 107982.
- Duffner, F.; Wentker, M.; Greenwood, M.; Leker, J. (2020): Battery cost modeling: A review and directions for future research. In: *Renewable and Sustainable Energy Reviews* 127, S. 109872.
- Erakca, M.; Baumann, M.; Bauer, W.; Biasi, L.; Hofmann, J.; Bold, B.; Weil, M. (2021): Energy Flow Analysis of Laboratory Scale Lithium-Ion Battery Cell Production. In: *iScience*, S. 102437.
- Kwade, A.; Haselrieder, W.; Leithoff, R.; Modlinger, A.; Dietrich, F.; Droeder, K. (2018): Current status and challenges for automotive battery production technologies. In: *Nat Energy* 3 (4), S. 290–300.
- Mauler, L.; Duffner, F.; Leker, J. (2021): Economies of scale in battery cell manufacturing: The impact of material and process innovations. In: *Applied Energy* 286, S. 116499.
- Nelson, P. A.; Ahmed, S.; Gallagher, K. G.; Dees, D. W. (2019): *Modeling the Performance and Cost of Lithium-Ion Batteries for Electric-Drive Vehicles*, Third Edition.
- Schönemann, M.; Bockholt, H.; Thiede, S.; Kwade, A.; Herrmann, C. (2019): Multiscale simulation approach for production systems. In: *The International Journal of Advanced Manufacturing Technology* 102 (5-8), S. 1373–1390.
- Schünemann, J. (2015): *Modell zur Bewertung der Herstellkosten von Lithiumionenbatteriezellen*. Zugl.: Braunschweig, Techn. Univ., Diss., 2015. 1. Aufl. Göttingen: Serke (16).
- Thomitzek, M.; Cerdas, F.; Thiede, S.; Herrmann, C. (2019a): Cradle-to-Gate Analysis of the Embodied Energy in Lithium Ion Batteries. In: *Procedia CIRP* 80, S. 304–309.
- Thomitzek, M.; Drachenfels, N.; Cerdas, F.; Herrmann, C.; Thiede, S. (2019b): Simulation-based assessment of the energy demand in battery cell manufacturing. In: *Procedia CIRP* 80, S. 126–131.
- Turetskyy, A.; Thiede, S.; Thomitzek, M.; Drachenfels, N.; Pape, T.; Herrmann, C. (2020): Toward Data-Driven Applications in Lithium-Ion Battery Cell Manufacturing. In: *Energy Technol.* 8 (2), S. 1900136.
- VDMA Battery Production (2020): *Roadmap Battery Production Equipment 2030*. Update 2020.
- Weeber, M.; Wanner, J.; Schlegel, P.; Birke, K. P.; Sauer, A. (2020): Methodology for the Simulation based Energy Efficiency Assessment of Battery Cell Manufacturing Systems. In: *Procedia Manufacturing* 43, S. 32–39.
- Wessel, J.; Turetskyy, A.; Cerdas, F.; Herrmann, C. (2021): Integrated Material-Energy-Quality Assessment for Lithium-ion Battery Cell Manufacturing. In: *Procedia CIRP* 98, S. 388–393.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

A modular, discrete-event simulation framework for modelling free ranging transportation vehicles in intralogistics

Ein modulares, ereignisdiskretes Simulationsframework zur Modellierung freinavigierender Transportfahrzeuge in der Intralogistik

Karl-Benedikt Reith, Sebastian Rank, Thorsten Schmidt, TU Dresden, Dresden
(Germany), karl_benedikt.reith@tu-dresden.de

Abstract: This publication presents a novel simulation framework for modelling fleets of autonomous, free ranging transportation vehicles in intralogistics. It aims at overcoming current shortcomings of popular, commercial DES tools in connection with free ranging vehicles like possibilities to apply sophisticated routing strategies. It is driven by the needs of fast simulation times and an accurate reproduction of the vehicles' free ranging behaviour. So, the presented framework combines elements from manufacturing and robotics simulation tools. In general, it can be applied for planning and assessing intralogistics systems to be equipped with Autonomous Intelligent Vehicles and e.g. to evaluate the performance of different fleet control strategies—especially in connection with large vehicle fleet sizes. This publication presents the general, structure of the discrete-event framework, discusses specific advantages and disadvantages and shows a possible connection to a reinforcement learning approach as an example.

1 Introduction

Opposed to traditional guided vehicles, free ranging transportation vehicles in intralogistics do not rely on predefined, static roadmaps and hence, can use the entire available space for moving. These vehicles are able to localize themselves in a familiar environment with the help of sensors like laser scanners, for example due to “natural navigation” (De Ryck et al. 2020). Furthermore, modern vehicles are usually autonomous, which means they can make decisions (e.g. which route to take towards a destination) individually (Fottner et al. 2021). As distinct from guided vehicles (AGV), within this contribution we call these autonomous, free ranging transportation vehicles AIV, Autonomous Intelligent Vehicles.

Though transportation systems based on AGV are still industry standard, we expect fleets of AIV to be increasingly used in the future. As AIV are able to travel on the direct link the transportation distances and consequently the transportation times are

supposed to be shorter in general (Xin et al. 2015). Furthermore, AIV can deviate from planned routes more easily in order to avoid apparent conflicts and circumnavigate obstacles, which results in more flexible transportation systems. However, the higher degree of freedom of each vehicle leads to more complex systems from a control perspective; especially for large fleets (Güller et al. 2018).

Generally speaking, simulation is a very important tool with regard to material flow analyses of vehicle-based transportation systems. Analytical approaches can hardly display numerous interdependencies and high system dynamics, such as vehicle kinematics, vehicle congestion etc. However, a low simulation runtime and simulating large fleets of autonomous, free ranging vehicles in the intralogistics context, while ensuring a “proper” level of detail are contradicting objectives—hardly surprising there is a lack of commercial tools allowing simulation of AIV in accordance to our demands. So, in this publication, we present a discrete-event framework that

- is capable of simulating large AIV fleets,
- is able to display free ranging behaviour (especially in connection with conflict and collision avoidance) in a “sufficiently” level of detail, and
- shows “satisfactory” simulation execution times.

The publication is structured as follows: Section 2 briefly describes the general structure of routing algorithms for free ranging vehicles. Section 3 gives a quick overview of existing simulation tools. This includes software from the fields of manufacturing and robotics. Section 4 describes the general idea and structure of our AIV-simulation framework. Section 4 also exemplarily shows how a reinforcement learning approach applied for routing can be implemented. Section 5 discusses general advantages and disadvantages and possible fields of application.

2 Routing of free ranging vehicles

The routing of AIV is a challenging task—especially with increasing fleet sizes and its corresponding system dynamics. As AIV routing was also our motivation for applying simulation techniques and lead us to develop a suitable simulator, a short introduction and overview follows:

Regarding the routing of vehicles, many approaches and surveys present a similar dichotomy (see e.g. Taghaboni-Dutta 1995, Duchoň et al. 2014, Gasparetto 2015, De Ryck et al. 2020): A first subtask, which we name *route planning*, is responsible for finding a general, feasible route (global route) towards a destination taking only static layout information into account. Here, most of the approaches apply shortest-route-algorithms. A second subtask, which we name *motion coordination*, adjusts the vehicle’s movements during the execution of the global route with regard to mechanical limitations, e.g. in order to avoid and circumnavigate emergent obstacles. The specific routing strategy has significant impact on the number of routing conflicts, possible congestion in the system and hence, the overall transportation system performance.

3 State of the art in simulation of free ranging vehicles

In order to plan, design and analyse transportation systems based on AIV fleets, from a material flow perspective an appropriate simulation tool needs to fulfil the following requirements:

- A modularized structure with standardized interfaces that allows the implementation of different routing approaches that approximate the routing (route planning and motion coordination, see section 2) of free ranging vehicles.
- The ability to model different transportation patterns, stress scenarios, etc.
- An evaluation that calculates the most important logistics indicators, such as throughput, travel times, due dates or an overall density map, with regard to congestion in the system.
- The capability to perform extensive simulation studies for large fleet sizes (> 50 vehicles), i.e. “empirical investigation of a model’s behaviour by a set of simulation runs with a systematic variation” (VDI standard 3633-1).
- A visualization of simulation runs.
- Open source availability.

In general, there are two fields, where the simulation of AIV is important: “manufacturing” and “robotics” (Demesure et al. 2016). Similarly, the VDI standard 2710-3 distinguishes the two simulation types “logistical, dynamic simulation” and “emulation”. Also, ter Mors (2010) discusses a similar subdivision into the problems of “multi-agent route planning” and “robot motion planning”.

The following subsections describe both fields briefly and shows that each addresses aspects which are not considered by the other (Demesure et al. 2016). There are different points of interest and hence different characteristics of the corresponding simulation tools with respect to issues that can be modelled and analysed.

3.1 Manufacturing simulation tools

Compared to the robotics perspective (see subsection 3.2), typical manufacturing simulators have a focus on the higher-order manufacturing/production/warehouse system. Consequently, the transportation system is usually seen as a part of a greater system and therefore displayed with a significant degree of abstraction. With regard to vehicle-based transportation systems in general, the following describes typical tasks for the application of simulation tools with a focus on manufacturing:

- An evaluation of the general transportation system design with regard to specific intralogistics indicators such as overall throughput or vehicle utilization, or a sensitivity analysis of specific vehicle parameters, e.g. the fleet size or the handling time (see e.g. Um et al. 2009, Ben-Salem et al. 2017).
- The influence and effect of specific vehicle control strategies, such as the vehicle dispatching or the route planning (see e.g. Caridá et al. 2015, Fanti et al. 2018).
- The visualization of a production system/warehouse (see e.g. Rohrer 2000).

Popular simulation tools with a manufacturing focus can be found in Law (2013). Such tools usually provide a variety of predefined components and performance indicators concerning (intra)logistics and manufacturing systems in general. Depending on the model’s complexity, these tools can be very fast (simulation time compared to real time), which allows simulation studies with multiple thousand runs.

Regarding free ranging vehicles, it is possible to model the routing behaviour in a more or less abstract manner, depending on the exact focus of the specific tool. This ranges from workarounds with predefined grid-based roadmaps to vehicles that are capable of moving roughly on the air-line between two defined points.

3.2 Mobile robotics simulation tools

Opposed to the high-level abstraction of the manufacturing field (see subsection 3.1), simulation tools from the robotics field have a very detailed perspective on the behaviour of each individual vehicle, or rather robot, in its environment. In general, this simulation domain supports the development, design and testing of mechanical components and the corresponding control systems. In order to be close to the real world, these tools emulate realistic sensor feedback in detail, or include realistic physics engines. Furthermore, typical tools are capable of incorporating real world control algorithms, which results in a close to real-world behaviour of simulated robots. This allows users to focus on a specific aspect of a mobile robot while the other parts are accurately displayed by default (Žlajpah 2008). For example, simulation tools can be used for the following tasks:

- The in-depth development and testing of the robot motion control with regard to specific hardware setups (see e.g. Al Mamun et al. 2018).
- The examination of the exact behaviour of a robot in challenging environments, e.g. with humans (see e.g. Liu et al. 2017) or in an uneven terrain (see e.g. Wang et al. 2017).
- The evaluation of localization and mapping approaches (see e.g. Miah et al. 2018).
- The exact visualization of the robot behaviour (Žlajpah 2008).

All in all, there are multiple tools with different individual key aspects. A survey regarding open source simulation software can be found in Torres-Torriti et al. (2016). Since the tools accurately simulate the real world and the behaviour of (in most of the cases single) robots, it is hardly possible to execute a simulation much faster than real time, due to a high calculation effort on the one hand and due to the incorporation of a control that was not designed for that on the other hand. Hence, an adequate simulation study of AIV fleets against the background of intralogistics questions is rather unpractical.

4 Simulation framework

This section gives a motivation for the design of a new simulation framework (4.1), describes the developed framework in detail (4.2) and shows an example application (4.3). In general, the following subsections show the framework rather from a user perspective instead of a software engineering perspective.

4.1 Motivation

In order to evaluate fleets of free ranging transportation vehicles in the intralogistics context with special focus on different routing strategies, we could not find a suitable simulation tool on the market. On the one hand, there are well established tools with focus on manufacturing systems (see subsection 3.1). Unfortunately, they either do not provide the opportunity to adequately implement AIV and their complex free ranging routing algorithms at all or it is very time consuming to implement them into

the corresponding simulation tools' infrastructure (Schneider et al. 2019). On the other hand, typical tools from the field of robotics (see subsection 3.2) with their in-depth focus on an exact vehicle behaviour and environment are not designed for conducting extensive simulation studies of large fleet sizes and automatically deriving and analysing logistics performance indicators.

4.2 Structure and modules

Our application relies on the python discrete-event simulation framework *simpy* (Matloff 2009). In general, the discrete-event paradigm is beneficial for a fast execution of the simulation (Law 2013). Python was chosen since it is a widely used language in science and industry, has an active community support and contains useful packages and frameworks, e.g. for data analytics or machine learning (Tambad et al. 2020). This allows to combine the simulation framework for example with reinforcement learning approaches, as will be shown in subsection 4.3.

Figure 1 displays the general structure of our framework. In an *experiment environment* the user can define all input parameters of a scenario, such as the layout, vehicle characteristics, etc. Furthermore, the user can define a specific experiment, which is defined by the parameters to investigate in a sensitivity analysis, e.g. the fleet size. A *design of experiment* module automatically creates the corresponding input parameter files for all necessary runs and finally starts the simulation runs.

The *core* of the simulation is responsible for keeping all events synchronized (especially the vehicle movements), updating the system status, generating new transportation tasks and creating the necessary log files for a subsequent analysis.

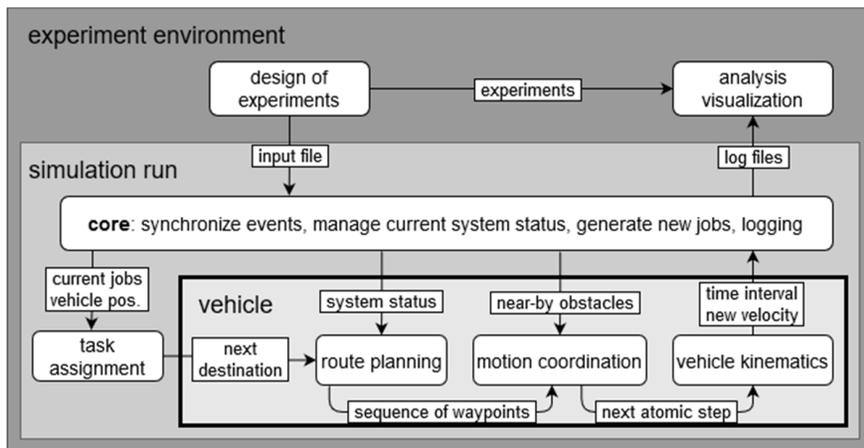


Figure 1: general structure and components of the simulation framework

The *task assignment* module is responsible for assigning transportation tasks to specific vehicles or vice versa assigning vehicles to specific tasks. The module's input are the current system status regarding open transportation tasks and vehicle positions in the layout. It is possible to apply very basic dispatching rules (“nearest job first”)

or even a sophisticated vehicle scheduling algorithm with a high look-ahead period. All in all, this module permanently determines the next destination for each vehicle.

Each vehicle individually executes the following three modules (route planning, motion coordination and kinematics), which represent a vehicle's autonomous routing behaviour. Based on the current destination as defined by the task assignment module, the *route planning* module determines a global route (see section 2). Different strategies can be applied in this module, like a shortest distance route (if a roadmap exists) or standard approaches for free ranging vehicles (e.g. rapidly exploring random trees). The route planning approach can theoretically make use of the entire global information (especially current position and destination of the other vehicles) or can just rely on local information or the vehicles individual experience (past travel times). It results in a sequence of (way)points that lead a vehicle to its destination.

The *motion coordination* module determines atomic steps between the waypoints taking a vehicle's local environment into consideration (see section 2). If there are dynamic obstacles near-by, such as other vehicles, the motion coordination of a vehicle will adapt the next steps in order to avoid a collision and resolve the routing conflict. Hence, physical collisions are prohibited. An example for a typical motion coordination approach is a so-called potential field method that calculates the next atomic direction with regard to attractive and repulsive forces in the local environment. The result of this module is an atomic step that does not violate the safety distance to any static or dynamic obstacle nearby.

The *vehicle kinematics* module transfers an atomic step as determined by the motion coordination into a corresponding atomic time interval taking the current vehicle velocity, possible acceleration/necessary deceleration, and the past and new direction of a step into account. This time interval and the coordinates of the next step are the main feedback to the simulation core, which keeps track of the vehicles' positions and movements. After the vehicle was delayed accordingly, the next atomic step can be calculated based on the new system status. For each vehicle, as long as there is a destination (and therefore next waypoints) defined, the simulation infinitely executes the loop of motion coordination, vehicle kinematics, and corresponding delays.

With regard to the design of experiments, the simulation saves relevant log files (global routes, atomic vehicle positions and steps, routing conflicts and congestion, job time stamps, etc.) for an *analysis*, or a subsequent *visualization* of a run. Furthermore, diagrams with predefined structure are created and stored automatically.

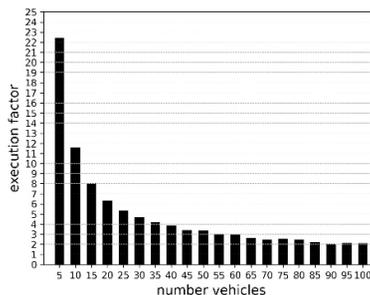


Figure 2: Average simulation execution factor (simulated time divided by real time).

In order to give a rough idea about the reachable simulation speed, Figure 2 displays the average simulation execution factor (how much faster is the simulation runtime compared to real time) with regard to the simulated fleet size as computed on an intel core i5, 2.5GHz, 2 cores. Large fleets with about 50 vehicles are simulated approximately 3 times faster than the reality, fleets with about 100 vehicles lead to an execution factor of about 2. However, the exact speed for a specific use case is also dependent on the layout and various simulation parameters.

4.3 Example: route planning using reinforcement learning

In order to give a deeper insight into the simulation framework, this example demonstrates its application and some core procedures. Figure 3 displays a layout (grey background) with two locations (crosses) and 6 obstacles (dark grey boxes). From a taxicab geometry perspective, there are 16 possible routes between both locations, whereas 10 routes can be considered as candidate routes (without detour). The best route to take in a specific situation depends on the current system status, i.e. the position and routes of other vehicles. It is possible to apply various algorithms for the route planning, i.e. for the identification of possible routes and the choice of a best route exemplarily. The following describes a reinforcement learning approach.

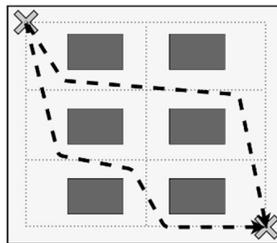


Figure 3: Example layout with two defined locations, an underlying grid roadmap and two alternative candidate routes as they are executed by AIV.

Figure 4 displays the elements of the reinforcement learning approach as included in the route planning module (see Figure 1). Before the simulation starts, a determination of possible global routes between different locations has to be done, e.g. finding all possible paths between two locations. With the information of the next destination (result of task assignment module), an artificial neural network (ANN) determines a specific route index (classification problem). Afterwards, the resulting index can be exactly matched to a specific predefined route and therefore a corresponding sequence of waypoints, which is the result of the route planning.

In the simulation, the vehicle executes the sequence of waypoints, while avoiding apparent routing conflicts with other vehicles (motion coordination). After reaching the final destination, the exact transportation time (reward) can be determined retroactively. A batch containing the particular basis for the decisions (observations), the choices for a route (actions) and the corresponding transportation times (rewards) trains the ANN.

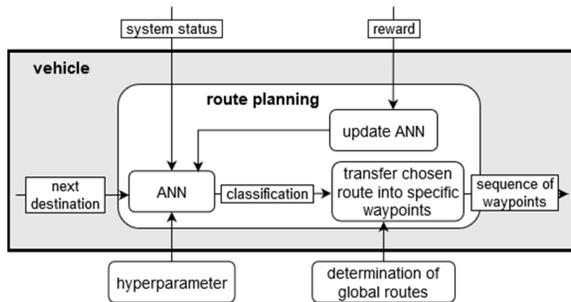


Figure 4: Example of a reinforcement learning approach as route planning.

Generally speaking, it is possible to use an ANN to select a vehicle’s route without taking the current system status (position and destination of other vehicles) into consideration. This leads to a relatively simple problem, where vehicles individually learn which routes between different points are beneficial in general. Over time, opposing traffic flows avoid themselves, resulting in one-way directions in bottlenecks. Opposed to that, with the ANN taking the current system status into consideration, a more complex problem arises. Here, each vehicle learns over time which route is beneficial in specific situations, in order to reach the destination quickly while avoiding routing conflicts with other vehicles.

5 Discussion

We created a simulation framework with focus on modelling the routing procedure of a fleet of free ranging transportation vehicles in intralogistics. The framework allows to implement different vehicle (control) strategies for the modules of “task assignment”, “route planning”, “motion coordination” and “vehicle kinematics”. The strategies and parameters can be approximated in order to realize an individual compromise between an accurate AIV behaviour and fast simulation runtimes. The advantages and disadvantages of the framework can be derived from this compromise.

The tool is fast enough to perform extensive simulation studies even with larger fleet sizes. It is modular, so different routing or task assignment strategies can be interchanged and evaluated quickly. Therefore, it can be used to assess general control strategies in preliminary studies. The tool automatically designs, executes and analyses necessary simulations with regard to predefined experiment parameters. It allows an “out-of-the-box” implementation of general routing strategies for AIV and evaluation with regard to logistics indicators.

Since there is only an approximation of the free ranging behaviour, the simulation tool however is not suitable for the in-depth analysis of a real world control, or the examination of the vehicle behaviour in a realistic environment. Effects like inter-vehicle communication, a physics engine, realistic vehicle sensing are not included.

The validity of a simulation based on the described framework is dependent on the exact control strategies and parameters applied, e.g. in the modules of route planning, motion coordination and vehicle kinematics. It needs to be determined individually with regard to the chosen algorithms/parameters for these modules and the corresponding real world system.

6 Conclusion

This publication presented a simulation framework for the simulation of an intralogistics transportation system based on free ranging and autonomous vehicles. The framework aims at finding a compromise between a high simulation speed and the depiction of free ranging vehicle behaviour. The general idea and modular structure of the framework was presented as well as a combination with a reinforcement learning approach as an example.

In the future we plan to further improve the framework with regard to vehicle kinematics and implement standard algorithms for the route planning, motion coordination and task assignment. Also, we plan to apply simulations based on the presented framework in future research projects.

Acknowledgement

The work was carried out as a part of the IGF research project 20516 BR “SVAT” (Schwarmverhalten autonomer Transportsysteme) and supported by “Forschungsgemeinschaft Intralogistik/Fördertechnik und Logistiksysteme (IFL) e.V.”.

References

- Al Mamun, M. A.; Nasir, M. T.; Khayyat, A: Embedded system for motion control of an omnidirectional mobile robot. *IEEE Access* 6 (2018), pp. 6722-6739.
- Ben-Salem, A.; Yugma, C.; Troncet, E.; Pinaton, J.: A simulation-based approach for an effective AMHS design in a legacy semiconductor manufacturing facility. In: 2017 Winter Simulation Conference (WSC), 2017, pp. 3600-3611.
- Caridá, V. F.; Morandin, O.; Tuma, C. C. M.: Approaches of fuzzy systems applied to an AGV dispatching system in a FMS. *The International Journal of Advanced Manufacturing Technology* 79 (2015) 1-4, pp. 615-625.
- De Ryck, M.; Versteyshe, M.; Debrouwere, F.: Automated guided vehicle systems, state-of-the-art control algorithms and techniques. *Journal of Manufacturing Systems* 54 (2020), pp. 152-173.
- Demesure, G.; Defoort, M.; Bekrar, A.; Trentesaux, D.; Djemaï, M.: Navigation scheme with priority-based scheduling of mobile agents: Application to AGV-based flexible manufacturing system. *Journal of Intelligent & Robotic Systems* 82 (2016) 3-4, pp. 495-512.
- Duchoň, F.; Babinec, A.; Kajan, M.; Beňo, P.; Florek, M.; Fico, T.; Jurišica, L.: Path planning with modified a star algorithm for a mobile robot. *Procedia Engineering*, 96 (2014), pp. 59-69.
- Fanti, M. P.; Mangini, A. M.; Pedroncelli, G.; Ukovich, W.: A decentralized control strategy for the coordination of AGV systems. *Control Engineering Practice* 70 (2018), pp. 86-97.
- Fottner, J.; Clauer, D.; Hormes, F.; Freitag, M.; Beinke, T.; Overmeyer, L.; Gottwald S.; Elbert, R.; Sarnow, T.; Schmidt, T.; Reith, K.-B.; Zadek, H.; Thomas, F.: Autonomous Systems in Intralogistics–State of the Art and Future Research Challenges. *Logistics Research* 14 (2021) 2, pp. 1-41.
- Gasparetto, A.; Boscarriol, P.; Lanzutti, A.; Vidoni, R.: Path planning and trajectory planning algorithms: A general overview. *Motion and operation planning of robotic systems 2015*, pp. 3-27.

- Güller, M.; Karakaya, E.; Uygun, Y.; Hegmanns, T.: Simulation-based performance evaluation of the cellular transport system. *Journal of Simulation* 12 (2018) 3, pp. 225-237.
- Law, A.: *Simulation modeling and analysis*, 5. ed. New York: McGraw-Hill Education 2013.
- Liu, S. B.; Roehm, H.; Heinzemann, C.; Lütkebohle, I.; Oehlerking, J.; Althoff, M.: Provably safe motion of mobile robots in human environments. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017, pp. 1351-1357.
- Matloff, N.: *Introduction to discrete-event simulation and the simpy language*. Davis, CA. Dept of Computer Science. University of California at Davis, 2009, pp. 1-33.
- Miah, M. S.; Knoll, J.; Hevrdejs, K.: Intelligent range-only mapping and navigation for mobile robots. *IEEE Transactions on Industrial Informatics* 14 (2017) 3, pp. 1164-1174.
- Rohrer, M. W.: Seeing is believing: the importance of visualization in manufacturing simulation. In 2000 Winter Simulation Conference Proceedings, 2000, pp. 1211-1216.
- Schneider, G. et al.: *Mikroelektronik in Deutschland für Industrie 4.0 fit machen - Sem140. Schlussbericht*. Dresden: Infineon Technologies Dresden GmbH & Co. KG, 2019.
- Taghaboni-Dutta, F.; Tanchoco, J. M. A.: Comparison of dynamic routeing techniques for automated guided vehicle system. *International Journal of Production Research* 33 (1995) 10, pp. 2653-2669.
- Tambad, S.; Nandwani, R.; McIntosh, S. K.: Analyzing programming languages by community characteristics on Github and StackOverflow. arXiv preprint arXiv:2006.01351 (2020).
- Ter Mors, A. W.: *The world according to MARP*. Dissertation. Delft University of Technology, 2010.
- Torres-Torriti, M.; Arredondo, T.; Castillo-Pizarro, P.: Survey and comparative study of free simulation software for mobile robots. *Robotica*, 34 (2016) 4, pp. 791-822.
- Um, I.; Cheon, H.; Lee, H.: The simulation design and analysis of a flexible manufacturing system with automated guided vehicle system. *Journal of Manufacturing Systems* 28 (2009) 4, pp. 115-122.
- VDI-standard 2710-3: *Application of simulation for automated guided vehicle systems (AGVS)*. Berlin: Beuth-Verlag 2014.
- VDI-standard 3633-1: *Simulation of systems in material handling, logistics and production*. Berlin: Beuth-Verlag 2014.
- Wang, C.; Meng, L.; She, S.; Mitchell, I. M.; Li, T.; Tung, F.; Wan, W.; Meng, M.; de Silva, C. W: Autonomous mobile robot navigation in uneven and unstructured indoor environments. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017, pp. 109-116.
- Xin, J.; Negenborn, R.R.; Corman F.; Lodewijks, G.: Control of interacting machines in automated container terminals using a sequential planning approach for collision avoidance. *Transportation Research Part C: Emerging Technologies* 60 (2015), pp. 377-396.
- Žlajpah, L.: Simulation in robotics. *Mathematics and Computers in Simulation* 79 (2008) 4, pp. 879-897.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Referenzmodell zur wertstrombasierten Simulation von Unstetigförderern in der Grobplanungsphase von Produktionssystemen

Reference Model for the Value-stream-based Simulation of Discontinuous Conveyors in the Rough Planning Phase of Production Systems

Markus Rabe, TU Dortmund, Dortmund (Germany), markus.rabe@tu-dortmund.de
Walter Wincheringer, Tobias Sohny, Hochschule Koblenz, Koblenz (Germany),
wincheringer@hs-koblenz.de, sohny@hs-koblenz.de

Abstract: The application of simulation in the bidding phase offers the advantage that suppliers of production systems can secure their throughput performance and improve their competitiveness. The disadvantage of simulation, a time- and resource-intensive modelling effort, can be reduced by using reference models. Based on the well-known value stream method (VSM), this paper presents a suitable modelling concept for a simulation reference model. It analyses how discontinuous conveyors of a material flow system can be modelled with sufficient accuracy applying value stream simulation for the bidding phase. The result is presented in a data box by means of parameters based on the VSM. Dynamic aspects are taken into account.

1 Einleitung

Mit Abgabe eines Angebotes für ein kundenindividuelles Produktionssystem (PS) mit einem komplexen Materialfluss garantieren Anbieter eine definierte Durchsatzleistung (Friedland und Kühling 2000). Für den Fall, dass ein realisiertes PS nicht den zugesicherten Durchsatz erbringt, drohen finanzielle Risiken, beispielsweise durch die erforderliche Nachbesserung des Systems. Eine Überdimensionierung des Systems, mit der Konsequenz von Mehrkosten, geht zu Lasten der Wettbewerbsfähigkeit. Daher bedarf es einer Absicherung der garantierten Durchsatzleistung des Systems zum Zeitpunkt der Grobplanung in der Angebotserstellung.

Hierzu sind die Produktionsprozesse mit den dazugehörigen Materialflüssen auf potentielle Auslegungsfehler zu überprüfen. Für eine übersichtliche Darstellung der Prozesse und deren Evaluierung eignet sich die Wertstrommethode (WSM) (Erlach 2020). Dynamisch auftretende Ereignisse lassen sich jedoch nicht ausreichend mit dieser Methode berücksichtigen (Urnauer und Metternich 2019). Für eine Abbildung dieser Ereignisse ist die ereignisdiskrete Simulation erforderlich. Diese erlaubt,

ablauffähige Modelle über die Zeit zu betrachten (Gutenschwager et al. 2017). Eine in Produktion und Logistik übliche ereignisdiskrete Simulation, insbesondere die Erstellung des dafür erforderlichen Modells, ist für die Angebotsphase allerdings zu zeit- und kostenintensiv (Friedland und Kühling 2000). Dies hat zur Folge, dass eine Simulation erst nach einer erfolgreichen Auftragsakquisition durchgeführt wird. Potentielle Planungsfehler werden erst zu diesem Zeitpunkt erkannt und führen zu aufwändigen Anpassungen. Daher bedarf es bereits in der Angebotsphase der Simulation bei geringem Modellierungsaufwand für eine hinreichend hohe Planungssicherheit der Durchsatzbetrachtung.

Dieser Beitrag ergänzt die Entwicklung eines Referenzmodells basierend auf Wertstromsimulation (Rabe et al. 2020a/b), welches sich durch eine aufwandsarme Modellerstellung von kundenindividuellen Produktionssystemen für die Angebotsphase eignet. In den Vorentwicklungen wurde mit Fokus auf der Abbildung der technischen Komponenten eines Materialflusssystems der Reihen- und Fließfertigung bereits Produktionsprozesse, das Produkt, Zusammenführungen und Verzweigungen (Rabe et al. 2020a) sowie Stetigförderer formalisiert (Rabe et al. 2020b). Dieser Beitrag behandelt die Entwicklung eines Modellelements für die Modellierung von Unstetigförderern.

2 Wertstromsimulation und Referenzmodelle

Dem Vorteil der Simulation, experimentierbare Modelle zu erhalten, steht der Hauptnachteil einer zeitaufwändigen Modellerstellung gegenüber. Dieser kann mittels eines Referenzmodells (RM) auf ein akzeptables Niveau reduziert werden (Friedland und Kühling 2000).

Im Bereich der Simulation fungieren Referenzmodelle als Konstruktionsschemata für die Gestaltung aufgabenbezogener Simulationsmodelle (Klinger und Wenzel 2000). Das RM fokussiert sich auf die Semantik und verallgemeinert die Syntax (Schütte 1998). Es befindet sich dabei auf der gleichen semantischen Stufe wie das Modell, welches mit ihm modelliert wird. Die Beschreibung erfolgt unabhängig vom Simulationssystem (Klinger und Wenzel 2000). Ein Modulelement-Baukasten eines Simulationssystems ist folglich kein Referenzmodell, sondern die simulatorspezifische Implementierung eines Referenzmodells (Schütte 1998).

Die Kombination der statischen WSM mit der DES ist daher naheliegend und wurde mehrfach beschrieben (Rabe et al. 2020a). So weisen jüngste Analysen erneut darauf hin, dass der Nachteil der WSM als statische Momentaufnahme durch die Erweiterung um die DES ausgeglichen werden kann und zudem eine Abbildung komplexerer Materialflüsse ermöglicht (Urnauer und Metternich 2019). Den existierenden Ansätzen ist gemein, dass die Wertstrommodelle in Simulationssysteme übertragen oder um diese ergänzt wurden. Hierbei lag der Fokus auf der Wertstrommethode. Die Abbildung eines Materialflusssystems (MFS) wird dabei zu oberflächlich vorgenommen. So sagen Meudt et al. (2017), dass die Berücksichtigung unterschiedlicher Förderprozesse mit einer Wertstromsimulation weiterer Forschung bedarf. Für die Sicherung der Planungsqualität in der Angebotsphase mit Wertstromsimulation ist eine höhere Granularität der Eingangsdaten erforderlich (Rabe et al. 2020a).

3 Struktur des Referenzmodells

Die dargestellten Vorteile der WSM, der DES und der RMe sollen miteinander kombiniert werden. Dies beinhaltet die Transparenz der WSM, die Dynamik der DES sowie ein reduzierter Modellierungsaufwand durch ein RM (Rabe et al. 2020a). Für ein modulares Konzept sind die für ein Materialflusssystem relevanten Systemelemente (Produktionsprozesse, Förderer, Produkt, Materialflussteuerung, Bewertung) als Modellelemente abzubilden. Diese sind so zu formalisieren, dass sie miteinander verknüpft werden können (Abb. 1). Für eine hohe Transparenz erfolgt die Darstellung der Modellelemente des RMs angelehnt an die WSM in Form von Datenkästen mit geeigneten Kenngrößen unter Berücksichtigung des Verhaltens über die Zeit. Gemäß der Einordnung der RMe erfolgt der Aufbau des Konstruktionsschemas simulatorunabhängig und bietet nach Implementierung eine höhere Granularität von Materialflüssen als es gängige Software-Tools zur Wertstromsimulation ermöglichen. Angelehnt an die WSM wurden deren Grundelemente *Produktionsprozesse* und *Materialfluss* (Stetigförderer) analysiert und unter dem Aspekt für das RM formalisiert (Rabe et al. 2020a/b). So erfolgt beispielsweise die Abbildung der Zykluszeit (ZZ) als kapazitives Angebot eines Produktionsprozesses (Erlach 2020) als auch der Prozesszeit (PZ) zur Abbildung von Chargenprozessen (Erlach 2020) produkt- und prozessspezifisch (Abb. 1).

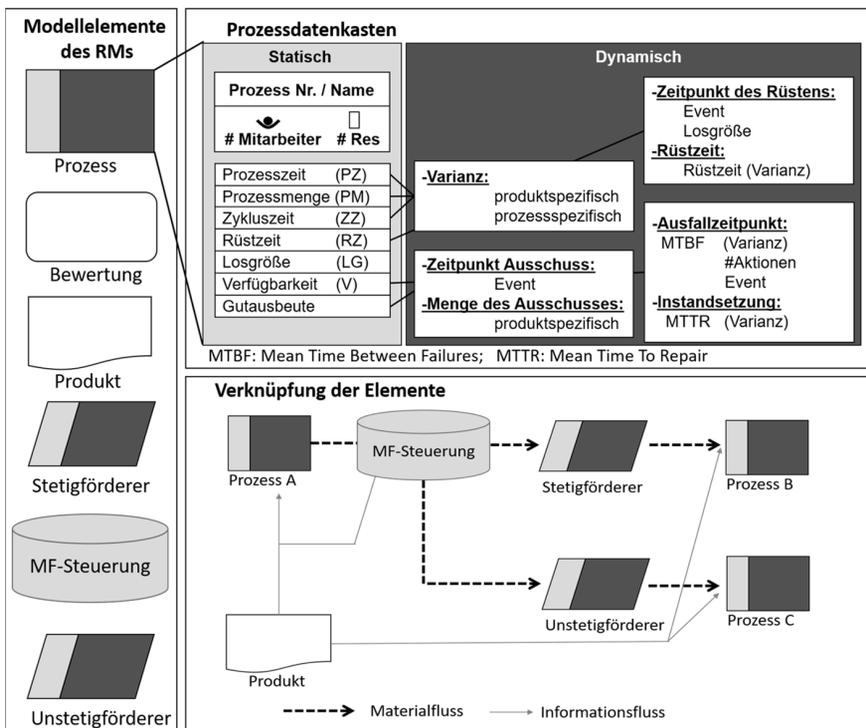


Abbildung 1: Übersicht des Referenzmodells (Rabe et al. 2020a, S.374)

3.1 Informationen zu Unstetigförderern in der Angebotsphase

Die Repräsentation des Materialflusses (Förderer) erfolgt bei der WSM einzig durch Pfeilsymbole und Materialbestände in Form von Dreiecken zwischen Produktionsprozessen (Erlach 2020). Förderer erzeugen einen kontinuierlichen oder diskontinuierlichen Materialfluss (MF), was eine Unterscheidung in Stetigförderer (z. B. Förderband) und Unstetigförderer (z. B. Fahrerloses Transportsystem, FTS) bedingt (ten Hompel et al. 2018). Angelehnt an die WSM bedarf es der Entwicklung von Kenngrößen unter Berücksichtigung von dynamischen Aspekten, welche die Unstetigförderer hinreichend genau beschreiben. Ferner soll die Darstellung in einem Datenkasten erfolgen, welcher sich zur Abbildung aller Unstetigförderer eignet. Für die Abbildung von Unstetigförderern bedarf es zuerst der Bestimmung, welche Daten und Informationen zum Zeitpunkt einer Angebotsphase vorhanden sind. Hieraus lassen sich Erkenntnisse zur Bestimmung der Kenngrößen ableiten, welche bei der Abgrenzung, was im RM betrachtet wird und was nicht, unterstützen. Ausgang bildet hierbei die Materialflussplanung (MFP). Der MF und das Layout sind miteinander verknüpft (Kesselring 2017). Das Layout ermöglicht die graphische Ausgestaltung der räumlichen Anordnung von Funktions- und Struktureinheiten sowie von Verbindungselementen (Schenk et al. 2014) und hat dabei Einfluss auf die Flexibilität eines MFS (Rabe 2003). Die MFP untergliedert sich in die Phasen der *Ideal-*, *Real-* und *Detailplanung*, wobei die Ideal- und Realplanung zusammengehörig als Grobplanung bezeichnet werden (Arnold und Furmans 2019). Die Inhalte einer Angebotsphase sind hierbei mit denen der Grobplanung gleichzusetzen (Ebert et al. 2015).

- Die *Idealplanung* liefert die technisch und organisatorisch optimale Lösung eines MF-Konzeptes, wobei bestehende Restriktionen, beispielsweise wie Fluchtwege (Schenk et al. 2014), keine Beachtung finden (Arnold und Furmans 2019). In Form eines Blocklayouts erfolgt aus Materialflusssicht die räumlich optimale Anordnung (geringster Förderaufwand) der Funktionselemente (Grundig 2018).
- In der *Realplanung* erfolgt die Ausarbeitung der optimalen MF-Variante. Bestehende Restriktionen werden dabei berücksichtigt. Ferner erfolgt u. a. die Bestimmung des geeigneten Fördermittels (FTS, Hängebahn) sowie die Wahl und Anordnung der Betriebsmittel (Arnold und Furmans 2019). Betreffend der Förderanlagen liegt das Hauptaugenmerk auf der Durchsatzbetrachtung, der Bestimmung der Anzahl der Fördermittel in Abhängigkeit des Streckenlayouts sowie einer grob definierten Streckenführungsstrategie als Navigation zur Bestimmung der Position des Fördermittels und der Strecke zur Zielposition (Trenkle 2018, Kesselring 2017). Das Groblayout visualisiert dabei eine realisierbare räumliche Anordnung der Funktionseinheiten unter Berücksichtigung von funktionalen, materialflussseitigen und flächenabhängigen Einflussfaktoren (Grundig 2018).

Abbildung 2 stellt die einzelnen Inhalte der Materialfluss-, Layout- und Fördertechnikplanung in Bezug zu der Ideal- und Realplanung einer Angebotsphase dar.

4 Formalisierung der Unstetigförderer

Entscheidend für das RM in der Angebotsphase ist die Wahl des Detailierungsgrads. Da diese Festlegung ein erhebliches Erfahrungswissen voraussetzt (Gutenschwager et al. 2017), wurde die VDI-Richtlinie 4465 (VDI 2016) herangezogen, welche mittels Leitfragen unterstützt. Relevant sind hierbei insbesondere die Fragestellungen, ob (i) die Modellierungsgenauigkeit sinnvoll mit der verfügbaren Informationsbasis

Inhalte in der Angebotsphase	Idealplanung	Materialfluss- und Layoutplanung	Fördertechnikplanung
	Realplanung	Blocklayout inkl. Materialflussmatrizen	Groblayout unter Berücksichtigung: <ul style="list-style-type: none"> • verbleibender Restriktionen • Wahl des geeigneten Fördermittels

Abbildung 2: Übersicht der Planungsinhalte einer Angebotsphase

gewählt wurde in Verbindung mit der Frage, ob (ii) das Weglassen von Aspekten das Modell noch glaubwürdig erscheinen lässt (VDI 2016).

Bezugnehmend zu (i) wäre die Nichtbetrachtung von Aspekten, welche nach der Grobplanung erfolgen, methodisch richtig, da das RM für die Angebotsphase konzipiert wird. Ausgehend von der dargestellten Analyse der Informationsbasis zum Zeitpunkt einer Angebotsphase ist diese allerdings zu grob, sodass gemäß (ii) der entstehende Auslegungsfehler zu groß wird und der Vorteil der Simulation, das auszulegende System mittels ablauffähiger Modelle hinreichend genau abzubilden, nicht mehr gegeben wäre. Entsprechend bedarf es der ganzheitlichen Betrachtung der Aspekte der Unstetigförderer und einer Analyse, ob und wie der betreffende Aspekt zu berücksichtigen ist. Im Unterschied zu bestehenden Ansätzen wie dem von Kesselring (2017), der ein virtuelles Konfigurationsmodell zur Planung und Konfiguration von Förderanlagen mittels mehrerer Modellelemente und Streckenabschnitte für alle Planungsphasen vorstellt, soll in diesem Beitrag die Abbildung der Unstetigförderer simplifiziert mit einem Datenkasten angelehnt an die WSM für die Angebotsphase erfolgen. Für eine hinreichend genaue Abbildung der Unstetigförderer sind die in Abbildung 3 dargestellten Aspekte zu analysieren.

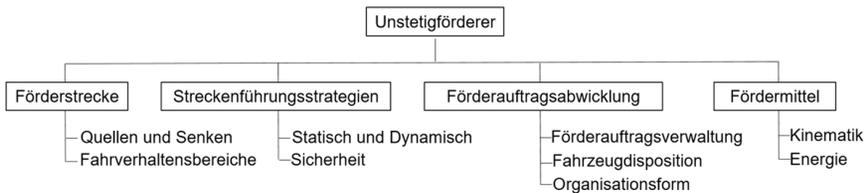


Abbildung 3: Analyse der Unstetigförderer (in Anlehnung an Kesselring 2017, S.96)

4.1 Förderstrecke

Dem Layout entsprechend bewegen sich Fördermittel auf einer Förderstrecke von Lastaufnahme- (Quellen Q) zu Lastabgabestationen (Senken S) als Start- und Endpunkte eines Materialflusses (ten Hompel et al. 2018). Diese können mit dem vorgestellten Datenkasten für Produktionsprozesse (Abb. 1) abgebildet werden (Rabe et al. 2020a/b). Betreffend der Förderstrecke sind Angaben zum Fahrverhalten zu spezifizieren. Hierzu zählen Geschwindigkeit und Beschleunigung, wodurch auch Abstandsregelungen der Fördermittel zueinander erforderlich werden (Kesselring 2017). Darüber hinaus existieren Überhol- und Ausweichbereiche sowie Sperrbereiche. Die Sperrbereiche sind Förderstreckenabschnitte, in welchem Fördermittel mit entgegengesetzten Fahrtrichtungen physisch nicht aneinander vorbeifahren können. Einzeln sind diese Bereiche in beide Richtungen passierbar (Kesselring 2017).

Abstandsregelungen, Ausweichregeln sowie Entscheidungen, wann welches Fördermittel Vorfahrt hat, erfolgen beim Detaillayout und werden in Folge dessen in der Angebotsphase nicht betrachtet.

Die Einsatzformen der Förderstrecke untergliedern sich in den Fließlinien- und Taxibetrieb (Kesselring 2017). Durch ihre geringe Anzahl an Verzweigungen und Zusammenführungen bildet der Fließlinienbetrieb ein Layout mit geringerer Komplexität (Kesselring 2017). Hierbei kommen meistens spurgebundene Unstetigförderer zum Einsatz. Dem gegenüber ist der Taxibetrieb mit einer größeren Anzahl an Verzweigungen und damit verbundenem komplexeren Layout charakterisiert. Hierbei kommen meistens nicht spurgebundene Fördermittel zum Einsatz (Kesselring 2017).

Bzgl. der spurgebundenen Fördermittel ergeben sich Streckenführungen. Neben Geraden und Kurven existieren Zusammenführungen und Verzweigungen, welche einen Materialfluss in verschiedene Richtungen ermöglichen (Arnold und Furmans 2019). Sowohl die Zusammenführungen als auch die Verzweigungen können mit dem vorgestellten Prozessdatenkasten (Abb. 1) (Rabe et al. 2020a) abgebildet werden, beispielsweise der Zeitbedarf für eine Weichenumschaltung als Zykluszeit (ZZ).

4.2 Streckenführungsstrategien

Mit Betrachtung der Streckenführung bedarf es einer Streckenführungsstrategie, die sich in statische und dynamische Strategien aufteilt (Kesselring 2017). Bei einer statischen Streckenführung bedarf es einer Routingtabelle an jeder Verzweigungsstelle zur Bestimmung des entsprechenden Förderwegs (Kesselring 2017). Hierbei kann das Routing in Abhängigkeit des Förderguts oder der Zielposition erfolgen. Die Berücksichtigung des Routings zum Zeitpunkt der Angebotsphase ist entsprechend vom Groblayout abhängig. Für eine effiziente Anwendung des RMs, angelehnt an die Darstellungsform der WSM in einem Datenkasten, ist das Ziel, eine vom Layout abstrahierte Betrachtung mittels Kenngrößen zu entwickeln. Detaillierte und aufwendige Ansätze wie in Rabe (2003), welcher das Layout mittels Graphen und Knotenpunkten als geometrische Koordinaten abbildet, werden daher nicht betrachtet. Für die Umsetzung sollen die Fördergüter (Produkte) Attribute in Form von *Quellen (Lastaufnahme-)* und *Senken (Lastabgabepositionen)* als Produktionsablaufschema besitzen. Das Ablaufschema ermöglicht einen produktspezifischen Materialfluss und übernimmt das *Routing* (Ulrich und Albrecht 2019). Die Abfrage erfolgt bei der Aufnahme des Förderguts durch das Fördermittel.

Mit Kenntnis der Quellen- und Senkenposition lassen sich in Abhängigkeit des Layouts die Distanzen (kürzeste Förderstrecke) bestimmen. Diese Beziehungen sind in einer *Kürzesten-Wege-Matrix* zu hinterlegen, sodass in Abhängigkeit der produktspezifischen mittleren *Fördergeschwindigkeit v* die benötigte *Förderzeit* im System ermittelt wird. Hierdurch ist die Abbildung eines Fließlinienbetriebs möglich. Ferner sind Schaltzeiten bei Verzweigungen und Zusammenführungen additiv zu berücksichtigen (Abb. 4).

Die dynamische Streckenführung ist dadurch definiert, dass die Förderstrecke während des Betriebs bestimmt wird (Kesselring 2017). Hierbei erfolgt die Berechnung der Förderstrecke mittels einer hinterlegten Strategie, beispielsweise für die schnellste oder kürzeste Strecke. Eine Ausarbeitung einer dynamischen Streckenführung ist herstellerepezifisch (Trenkle 2018). Zudem benötigen Unstetigförderer ohne Spurbindung einen hohen Planungsaufwand, welcher der Detailplanung zuzuordnen ist. Folg-

lich findet die Betrachtung einer dynamischen Streckenführung keine Berücksichtigung für ein RM in der Angebotsphase. Mit der Kürzesten-Wege-Matrix kann deren Abbildung allerdings abstrahiert werden. Unter Berücksichtigung der kürzesten Quellen- und Senkenbeziehungen, ausgehend von einem Groblayout, ist mit einer imaginären Spurbindung eine hinreichend genaue Abbildung der Förderstrecke möglich. Hierdurch kann auch der Taxibetrieb abgebildet werden.

von \ nach	Position 1	Position 2	Position 3	...
Position 1		20 m	X	...
Position 2	20 m		72 m	...
Position 3	45 m	72 m		...
...

X: Keine Verbindung

Abbildung 4: Darstellung der Kürzesten-Wege-Matrix

Des Weiteren werden Detailanalysen der Sicherheitsaspekte, wie das Ausweichen oder ein Stillstand, um eine Kollision zu vermeiden, sowie sicheres Bremsen, sodass das Fördergut nicht vom Fördermittel fällt, ausgeklammert (Kesselring 2017). Diese sollen mittels herstellerepezifischer Erfahrungen bei der Förderzeit berücksichtigt werden (Ullrich und Albrecht 2019). Im Fokus steht eine Durchsatzbetrachtung bei einer frühen Planungsphase in Abhängigkeit der Anzahl von Unstetigförderern. Ein einzelnes Fördermittel ist bei einer solch frühen Planungsphase nicht ausschlaggebend. Entscheidend ist eine hinreichend genaue Abbildung der Förderdistanzen, sodass Durchsatzbetrachtungen in Abhängigkeit der Förderzeiten möglich sind.

4.3 Förderauftragsabwicklung

Für die Zielfestlegung bei einer Auftragsabwicklung bedarf es der Verarbeitung von Förderaufträgen und der spezifischen Zuweisung zu einem freien Fördermittel. Dies lässt sich wie folgt untergliedern (Ullrich und Albrecht 2019):

1. Auftrag generieren: Die Förderaufträge werden nach Fertigstellung eines Prozesses an einer Quelle (Lastaufnahme-position) generiert.
2. Förderauftragsverwaltung: Die generierten Aufträge werden in der Förderauftragsverwaltung nach dem FIFO-Prinzip angeordnet (Ullrich und Albrecht 2019). Eine ideale Reihenfolgebildung der Aufträge, ausgehend von einem übergeordneten PPS-System (Ullrich und Albrecht 2019), liegt nicht im Fokus des RMs und wird nicht weiter betrachtet. Das erfordert eine Differenzierung in die Förderauftragsverwaltung und Fahrzeugdisposition.
3. Fahrzeugdisposition: Die Fahrzeugdisposition bestimmt das freie Fördermittel für die Ausführung des Auftrags (Ullrich und Albrecht 2019). Für diese Ermittlung kommen unterschiedliche herstellerepezifische Strategien zum Einsatz, beispielsweise zur Vermeidung von Staus (Ullrich und Albrecht 2019). Durch die *Kürzeste-Wege-Matrix* ist die aktuelle Position der Fördermittel an Quellen und Senken bekannt. Um eine hinreichend genaue Abbildung zu erzielen, soll der Auftrag dem freien Fördermittel zugewiesen werden, welches zur Lastaufnahme-position die kürzeste Distanz aufweist. Hierbei erfolgt keine Änderung der Reihenfolge der Aufträge; diese werden nach Auftritt ihrer Generierung verarbeitet (FIFO-Prinzip). Für den Fall, dass keine Förderaufträge vorliegen, verharret das

Fördermittel an seiner letzten Position, was bei einer Wiederaufnahme eine Leerfahrt zur Quelle bedingt. Die Auftragsvergabe und damit verbundene Anschlussfahrt (Last- oder Leerfahrt), soll hierbei ausschließlich an der Quelle oder Senke erfolgen und nicht auf dem Pfad (während der Förderzeit) sobald das Fördermittel leer ist. Ferner wird eine Priorisierung hinterlegt (Attribute), sodass die Versorgung der Prozesse mit Material eine größere Gewichtung hat als die Entsorgung von Leergut.

Als Organisationsform erfolgt die Betrachtung einer direkten Verbindung (1:1) von Q zu S. Milkruns (1:m bzw. n:1) und Routenzüge (n:m) werden nicht abgebildet.

4.4 Fördermittel

Zuletzt bedarf es der Abbildung des Fördermittels. Hierbei ist die produktspezifische *Kapazität* als Fassungsvermögen der Fördermittel zu beachten. Ferner sind die *technische Verfügbarkeit* (MTBF) in Abhängigkeit von Aktionen oder einem Event und die Reparaturzeit (MTTR) zu berücksichtigen. Für den Fall, dass der Be- und Entladevorgang durch das Fördermittel aktiv erfolgen kann, sind die produktspezifischen Prozesszeiten (PZ) $PZ_{Beladen}$ und $PZ_{Entladen}$ des Fördermittels zusätzlich zur Förderzeit zu berücksichtigen. Neben einer reinen Lastaufnahme und -abgabe können Unstetigförderer zudem in einen Produktionsprozess integriert sein (ten Hompel et al. 2018), beispielsweise in eine Roboterschweißzelle. Entsprechend ist die PZ zu berücksichtigen, für welche das Fördermittel in den Produktionsprozess eingebunden ist und nicht für weitere Fahraufträge zur Verfügung steht. Ferner können bei Montagetätigkeiten mehrere Unstetigförderer benötigt werden, um die erforderliche Prozessmenge (PM) zu erhalten. Entsprechend ist eine produktspezifische *Logik* zu etablieren, an welchen Prozessen und in welcher Anzahl die Unstetigförderer als Teil des Prozesses zur Verfügung stehen müssen. Eine Aufladung des Energiespeichers des Unstetigförderers kann u. a. induktiv oder an Lade- bzw. Akkutauschstationen erfolgen (Ullrich und Albrecht 2019). Eine induktive Ladung, die bei einem einfachen Layout während der Fahrt erfolgt, bedarf keiner weiteren Betrachtung (vgl. Ullrich und Albrecht 2019). Bei Ladestationen und Akkutauschstationen bedarf es einer gezielten Steuerung, welcher Unstetigförderer welche Station bei welchem Energielevel ansteuern darf (Trenkle 2018). Da diese Steuerung vom Hersteller abhängt und erst in der Detailplanung erfolgt, wird dies für das RM in der Angebotsphase nicht betrachtet. Aus Materialflusssicht ist allerdings die Verfügbarkeit des Fördermittels, als Relation der Betriebszeit zur Ladezeit, von entscheidender Bedeutung und wird folglich als *Betriebsstunden pro Ladezeit* berücksichtigt. Betreffend einer Durchsatzbetrachtung haben die layoutbedingten Förderstrecken sowie die Anzahl an Fördermitteln einen Haupteinfluss. Physikalische Aspekte wie die Kinematik (Rangierparameter), Abmessungen sowie die Beschleunigung und Verzögerung in Relation zur Förderzeit sind vernachlässigbar klein und werden erst bei der finalen Auswahl des Fördermittels in der Detailplanung relevant (VDI 2016).

Aus der Analyse ergibt sich konkret der Datenkasten für Unstetigförderer mitsamt der *Kenngrößen*, der alle abzubildenden Eigenschaften ganzheitlich bündelt (Abb. 5). Dieser bildet das Modellelement Unstetigförderer für das Referenzmodell basierend auf der WSM. Die definierten *Eingabeparameter* sowie die *produktspezifische Routingtabelle* werden in der Simulation verarbeitet. In Abhängigkeit des produktspezifischen Materialflusses (*Routingtabelle*) erfolgt die Bestimmung der Förderstrecke mittels der hinterlegten *Matrix*. In Kombination mit der produktspezifischen

mittleren *Fördergeschwindigkeit* v wird die *Förderzeit* im System berechnet. Die *Prozesslogik* erlaubt die Abbildung möglicher Montagetätigkeiten.

Die *Förderauftragsabwicklung* weist dem freien Fördermittel mit der kürzesten Distanz zur Quelle den Auftrag zu (*Fahrzeugdisposition*). Dies erlaubt eine Abbildung von mehreren Unstetigförderern ($\#$ *Ressourcen*) eines Typs mit einem Datenkasten. Durch das Modellelement in Form eines Datenkastens soll neben der Abbildung der Unstetigförderer mit einer Simulation der Durchsatz in Abhängigkeit der Anzahl an Unstetigförderern durch den Anwender bestimmt werden (Was-wäre-wenn-Szenarien). Die Evaluierung erfolgt in einer weiteren Entwicklung.

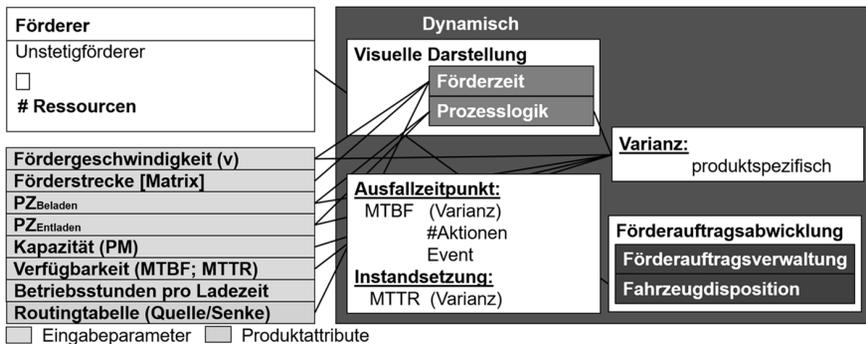


Abbildung 5: Datenkasten für Unstetigförderer

5 Zusammenfassung und Ausblick

Für Anbieter von kundenindividuellen Produktionssystemen mit einem komplexen Materialfluss bietet die Anwendung der Simulation zum Zeitpunkt der Angebotsphase den Vorteil einer exakten Angebotserstellung. Ein frühzeitiges Erkennen von Auslegungsfehlern steigert zudem die Wettbewerbsfähigkeit. Hierzu muss der Modellierungsaufwand möglichst gering ausfallen. Eine Kombination der WSM und Simulation verspricht eine transparente Darstellung aller erforderlichen Produktionsprozesse in Verbindung mit den entsprechenden Material- und Informationsflüssen unter Beachtung der Dynamik. Der bestehende Nachteil einer ressourcen- und zeitintensiven Modellerstellung von Simulationsmodellen wird mit einem Referenzmodell (RM), basierend auf einer Wertstromsimulation, reduziert. In diesem Beitrag wurden Kenngrößen unter Beachtung von dynamischen Aspekten definiert, welche Unstetigförderer hinreichend genau beschreiben und mit dem RM, basierend auf Wertstromsimulation, effizient in einer Angebotsphase abgebildet werden können.

In einer weiteren Entwicklung des RMs basierend auf einer Wertstromsimulation bedarf es der Untersuchung der Materialflusslogik. Hierzu sollen Standard-Logiken erarbeitet werden, welche eine Materialflusssteuerung hinreichend genau abbilden. Ferner sind Kenngrößen zu erarbeiten welche eine Bewertung eines Materialflusssystems ermöglichen. Die Gesamtkonzeption wird nach der Ausarbeitung in einem Prototyp simulatorspezifisch implementiert und im industriellen Umfeld hinsichtlich des Modellierungsaufwands und der Ergebnisdarstellung evaluiert.

Literatur

- Arnold, D; Furmans, K.: Materialfluss in Logistiksystemen. 7 Auflage. Berlin: Springer 2019.
- Ebert, R-E.; Schachmanow, J.; Wrobel, G.: Intelligente Planung von Produktionssystemen. ZWF (2015) 110 (3), S. 103- 106.
- Erlach, K.: Wertstromdesign. Der Weg zur schlanken Fabrik. 4. Auflage. Berlin: Springer 2020.
- Friedland, R.; Kühling, M.: Referenzmodelle für Fertigungssysteme. In: Wenzel, S. (Hrsg.): Referenzmodelle für die Simulation in Produktion und Logistik. Erlangen: SCS 2000, S. 133-150.
- Gutenschwager, K., Rabe, M., Spieckermann, S., Wenzel, S.: Simulation in Produktion und Logistik – Grundlagen und Anwendung. Berlin: Springer 2017.
- Grundig, C -G.: Fabrikplanung. Planungssystematik – Methoden – Anwendungen. 6. Auflage. München: Hanser 2018.
- Kesselring, J.: Prozessbegleitende Planung und Konfiguration von Fördertechnikanlagen unter Zuhilfenahme von virtuellen Konfigurationsmustern und Konfigurationsmodellen. Karlsruher Institut für Technologie, Dissertation 2017.
- Klinger, A.; Wenzel, S.: Referenzmodelle – Begriffsbestimmung und Klassifizierung. In: Wenzel, S. (Hrsg.): Referenzmodelle für die Simulation in Produktion und Logistik. Erlangen: SCS 2000, S. 13-29.
- Meudt, T.; Kaiser, J.; Metternich, J.; Spieckermann, S.: Wertstrommodellierung und –simulation im Zeichen von Digitalisierung und Industrie 4.0. ZWF (2017) 112 (12), S.865-868.
- Rabe, M.: Modellierung von Layout und Steuerungsregeln für die Materialfluss-Simulation. Technische Universität Berlin, Dissertation 2003.
- Rabe, M.; Wincheringer, W.; Sohny, T. (2020a) Reference model based on value stream simulation for the evaluation of production systems in the bidding phase. In: 13th International Conference of Research in Logistics and Supply Chain Management (RIRL), Le Havre: 7.-9. Oktober 2020.
- Rabe, M.; Wincheringer, W.; Sohny, T. (2020b) Referenzmodell basierend auf Wertstromsimulation zur Bewertung von Produktionssystemen in der Angebotsphase. In: Proceedings 25. Symposium Simulationstechnik, Online Conference, 14-15 Oktober 2020, S. 373-380.
- Schenk, M.; Wirth, S.; Müller, E.: Fabrikplanung und Fabrikbetrieb. Berlin: Springer 2014.
- Schütte, R.: Grundsätze ordnungsmäßiger Referenzmodellierung. Konstruktion konfigurations- und anpassungsorientierter Modelle. Wiesbaden: Springer 1998.
- ten Hompel, M.; Schmidt, T.; Dregger, J.: Materialflusssysteme. Förder- und Lagertechnik. Berlin: Springer 2018.
- Trenkle, A.: Entwurfsmuster für Fahrerlose Transportsysteme. Karlsruher Institut für Technologie, Dissertation BAND 93 2018.
- Ullrich, G.; Albrecht, T.: Fahrerlose Transportsysteme. Eine Fibel – mit Praxisanwendung zur Technik – für die Planung. 3. Auflage. Wiesbaden: Springer 2019.
- Urnauer, C.; Metternich, J.: Die digitale Wertstrommethode. ZWF (2019) 114 (12), S. 855-858.
- VDI: Richtlinie 4465 Modellierung und Simulation, Blatt 1 – Modellbildungsprozess. Berlin: Beuth 2016.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Application of discrete-event simulation for factory planning – A case study

Anwendung einer ereignisdiskreten Simulation im Fabrikplanungsprozess – Eine Fallstudie

Meike Herbert, Andreas Selmaier, Felix Mühlmann, Jörg Franke, Institute for Factory Automation and Production Systems, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen (Germany), meike.herbert@faps.fau.de, andreas.selmaier@faps.fau.de, felix.muehlmann@fau.de, joerg.franke@faps.fau.de

Jens Fürst, Siemens Healthcare GmbH, Erlangen (Germany),
jens.fuerst@siemens-healthineers.com

Abstract: During factory planning, production areas are dimensioned, structured and designed. Material flow simulation is part of the digital tool chain and intends to support factory planning processes. In this paper, a case study of planning a new factory in the medical technology industry is presented. A concept for the consistent use of digital data as a database for simulation is pointed out improving the information flow and transparency. Main objectives of the simulation study are the dimensioning of resources, optimization of sequences and processes as well as providing a three-dimensional visualisation. It is shown, how the discrete-event simulation study is integrated into the factory planning process. The application comprises the planning of the production area of high voltage tanks and the results of the simulation study are presented. Two examples are described in detail dealing with the impact of the shift work model and a comparison of different assembly concepts.

1 Introduction

Sustainable production and logistics systems are characterized by modularity, flexibility and adaptability. These characteristics have to be considered during planning processes in manufacturing industries. In greenfield scenarios, this planning process has less constraints than in existing factories and therefore, opens up opportunities for new production concepts. Besides Industry 4.0 and the transformation towards a digital industry, classical lean management is still the basis for factory planning processes.

Digital tools improve and shorten planning and development processes significantly. Research in the field of simulation has increased over the past years and many

applications exist, for example in the area of factory layout planning (FLP) and material flow simulation (Mourtzis et al. 2014).

Designing factory layouts is a multidisciplinary and knowledge-intensive task. Fast changes in current global markets require agile planning, designing or reconfiguration of factory layouts (Mourtzis et al. 2014). In the planning phase, additional simulation tools are applied to e.g. simulate material flow systems, work systems and warehouse systems. The tools are used, for example, to plan human resource deployment, examine planning parameters or increase the efficiency with regard to the dimensioning, throughput or bottlenecks (VDI 3633 Part 1 2014).

Data integration into the factory model comprises data from different areas such as plant information, product and production information and sequence control information. During simultaneous engineering, where product and production development takes place in parallel, simulation tools that are integrated in the information technology environment are beneficial (VDI 3633 Part 1 2014).

During factory planning, the use of discrete-event simulation has to be considered to support decision-making processes. An integrated database should be used and a procedure for transferring the results of the simulation study into the layout planning process has to be applied.

In the presented case study, the use of material flow simulation for planning a new factory is shown for an application in the medical technology industry. A concept for tool integration is pointed out which enables improved information flow and data exchange. The simulation study is conducted for the existing production as well as for the production area in the new factory. Finally, the results of two specific issues are explained exemplary. In the simulation study of the existing production, the impact of the shift work model on the dimensioning of the capacities is analysed. Furthermore, the impact of different assembly concepts on the lead times is evaluated for the newly planned production area. The results of the simulation study are integrated into the layout planning process.

2 Discrete-event simulation for factory planning processes

Procedure models for the execution of a simulation study are described in VDI 3633 Part 1 (2014) or by Rabe et al. (2008) with focus on verification and validation. Data collection and preparation is necessary for each simulation study. Input data is generated manually, automatically or semi-automatically. Manual data collection and preparation is often time-consuming and error-prone. Therefore, automatic approaches are preferred with direct or indirect access to the basic data of the planning systems or IT systems. Data consistency and quality are still issues to be considered in practice. However, standardized interfaces are implemented and the data consistency is improved with the digital transformation (Mieschner and Mayer 2020).

Common simulation tools provide an integrated 3D visualisation, which helps to improve the understanding, identify critical processes and enhance the communication to non-experts in simulation (Süß et al. 2009). It has also advantages for the layout planning process. Sender and Wanner (2013) present an integrated planning tool for capacity and layout planning, but the application is of the maritime

industry considering the specific requirements. Dombrowski and Ernst (2013) present a simulation approach for layout planning comparing and evaluating different future scenarios with a case study of factory redesigning in a medium-sized company and changing requirements. Optimization algorithms for layout planning are integrated into the simulation, as for example described in an overall concept by Krüger (2019).

There are many different applications for discrete-event simulation in the literature, but only few examples deal with planning a new factory. The dependencies between simulation and layout planning have to be pointed out clearly. Furthermore, the database and consistency to other engineering and planning tools is of particular importance in the factory planning process.

A case study of a simulation for planning a new factory is presented in the following going into the aspects of the required database and data consistency as well as the dependencies to the layout planning process. Based on the findings of the application in the medical technology industry, the main objectives, procedures and results of the simulation in the factory planning process are presented.

3 Case study

The main objectives of factory planning are, in summary, ensuring economic efficiency, flexibility and adaptability as well as attractiveness, lately complemented by the objective of energy and resource efficiency (Grundig 2018). These objectives are also pursued in the planning and design of a new plant, such as the technology center of Siemens Healthineers in Forchheim, Germany. The so-called High Energy Photonic (HEP) Center is designed as a lean, digital factory, where X-ray tubes and high voltage generators will be produced (InFranken 2020).

A systematic analysis and evaluation of the assembly and production operations using a discrete-event simulation supports the production planning and optimization for the new factory. In the following, the application, procedure and results of the simulation study of a defined production area are pointed out.

3.1 Requirements and objectives

While there are confined space conditions in existing production areas that result in restrictions of the layout design, new opportunities for space allocation and layout design arise in the planning of the future plant. This enables an optimal arrangement of the equipment in accordance with the material flow. In addition, it opens up the possibility of further extending production capacities.

The application in this case study comprises a defined production area of medical technology products. The production area and the corresponding products are characterized by the following characteristics:

- 19 different products of two product groups that are high voltage tanks (HVT) and single tank units (STU) are assembled in a serial production with small quantities.
- The operations are structured in pre-assembly, assembly, technology zone and final assembly operations.

- The production area is characterized by long assembly times of several hours, high lead times of up to a few days due to long drying times. Requirements on cleanliness and quality are high. Sequences are less transparent due to multiple machine operation in the technology zone.
- The process sequences vary strongly for each product.
- Changes in the product portfolio, increasing quantities and shorter assembly times characterize the changing conditions in the next years.

The main objectives of the simulation study are dimensioning, optimization and visualisation, as depicted in Figure 1. System dimensioning concerns all resources that are needed in the future scenario such as workstations, machinery, equipment and buffers, but also the human resource deployment. Utilizations are extracted and bottlenecks can be identified. The results of the analysis are integrated into the layout planning process. The consideration of flexibility and adaptability in the layout is crucial for a sustainable future-oriented design. Therefore, different scenarios with increasing quantities and changing product portfolio are evaluated.

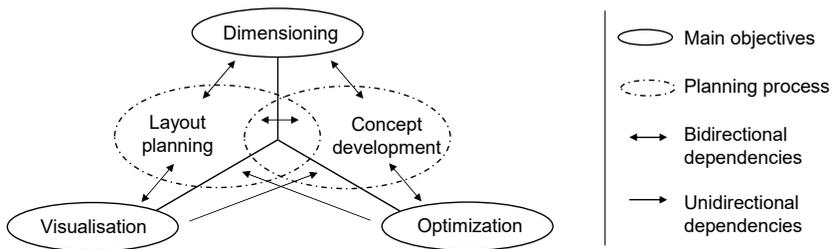


Figure 1: Overview of the objectives for the simulation study and their effect on the planning process

The development of the production concept is also a part of the planning process. This comprises the determination of all parameters that are relevant for the production. Two examples are shown in this case study (cf. chapter 3.3) analysing working time models and comparing different assembly concepts. The concept development has an effect on the dimensioning and layout planning, but the analysis mainly aims at an optimization of the system. The results of the optimization are transferred to the layout design. In general, the optimization process is performed to improve chosen (key) performance indicators (KPI).

The visualisation increases the comprehension of processes and different production concepts, especially for non-experts in simulation. Furthermore, the three-dimensional visualisation is connected to the layout planning process and shows the arrangement of resources in 3D, which is beneficial for a newly planned production area. In order to obtain a realistic visualisation of the production area, not only standard objects of the simulation tool are used, but also 3D objects that are modelled using external computer aided design (CAD) tools.

3.2 Procedure of the simulation study

The simulation study is performed in the factory planning process. The integration of engineering tools effecting the generation of a database for the simulation is

described. Moreover, the dependencies between simulation and layout planning are pointed out in detail.

3.2.1 Simulation database

Material flow simulation is a part of the digital tool chain comprising CAD, factory and line design, material flow optimization, robotics and automation simulation and further engineering tools. The use of these tools is based on a production data management (PDM) system to ensure data consistency in the whole planning process. The data is structured in bill of materials (BOM), bill of resources (BOR) and bill of processes (BOP). In this context, Figure 2 shows the structure of the simulation database, the future concept of an integrated tool landscape and the data sources of the case study.

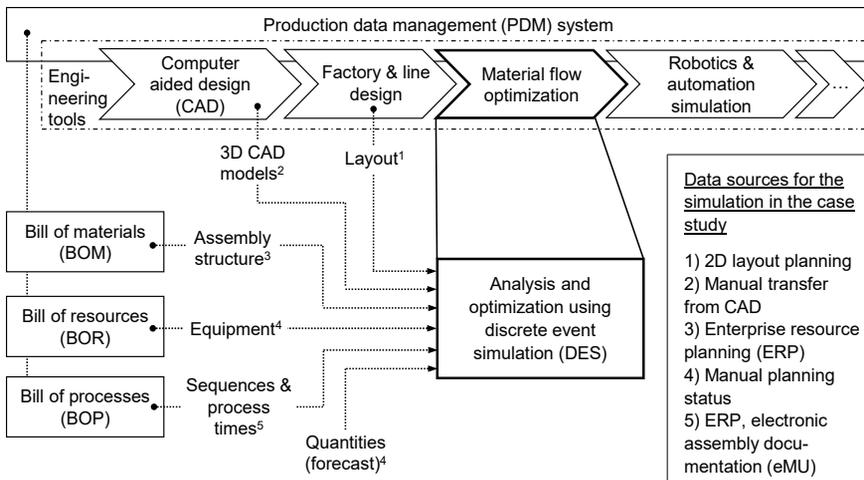


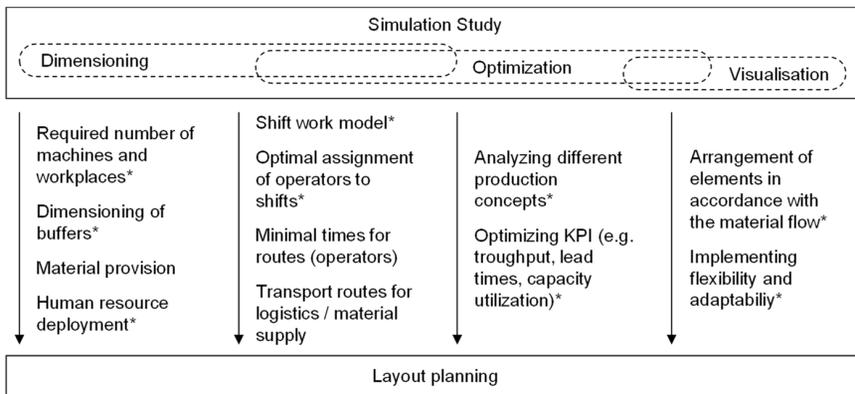
Figure 2: Towards consistent use of digital data as a basis for material flow simulation

3D CAD models of facilities and equipment are used to enable 3D visualisation in the simulation. While layout planning is performed in a 2D tool today, an integrated engineering tool for 3D line design will be used in the future. Furthermore, BOM, BOR and BOP provide important data for the simulation that are information on the assembly structure, equipment, sequences and process times. In this case study, data is collected from other planning tools such as Enterprise resource planning (ERP), or documents such as the electronic assembly documentation. Furthermore, the manual planning is performed using spreadsheet programs.

Data collection is characterized by high manual effort for the case study, but the factory planning process will be improved by an integrated tool landscape. Due to the central provision of technical data and models, the transparency and continuous reuse of digital data is increased. Tool and process discontinuities with manual data transfers are avoided and the simulation study is performed with an enhanced integration and visualisation.

3.2.2 Supporting layout planning by simulation

In factory planning, the use of simulation supports the layout planning process. A systematic analysis and evaluation of different scenarios provide the basis for decision-making in designing layouts. In simulation, the dimensioning of resources is conducted, an optimization of the processes is performed and a three-dimensional visualisation is provided, which improves layout planning. The results of the simulation study are integrated into the layout design. Specific aspects that are examined in the simulation study for factory planning are depicted in Figure 3.



* Aspects examined in the case study

Figure 3: Aspects investigated in the simulation study effecting layout planning

The aspects include the dimensioning of resources such as machines, workplaces, buffers and systems for material provision. For human resource deployment, the number of required operators has to be determined. The shift work model is analysed and the optimal assignment of operators to the shifts are determined depending on the process sequences. The routes and distances that the workers have to walk in the production area are analysed. Furthermore, logistics aspects are analysed such as the transport routes and further aspects concerning material supply. Various scenarios are examined e.g. comparing different production or assembly concepts. Optimization measures are based on an evaluation of KPI. A main aspect for layout planning is the arrangement in accordance with the material flow and also the implementation of flexibility and adaptability in layout designs. The simulation (and visualisation) of processes support the understanding of complex sequences in production and helps to evaluate future scenarios (e.g. with different products, processes and quantities).

3.3 Application, results and discussion

The simulation study is carried out for the production area considering the approaches described above. The tool Tecnomatix Plant Simulation is used for the material flow simulation. Different scenarios are evaluated based on KPIs such as throughput, lead times or capacity utilisation. In the following, two specific issues will be addressed as an example that are a comparison of the shift work models and

a comparison of different assembly concepts. Afterwards, the results of the whole case study will be summarized and discussed.

3.3.1 *Impact of the shift work model on the dimensioning of the capacities*

The simulation study enables the analysis of the current state. After parameter variation and evaluation of the results for different scenarios, conclusions are drawn and transferred to the factory planning process. The comparison of different shift work models is one issue of interest, which can also provide insights of the current situation in the production area.

A short-term change in the shift work schedule became particularly relevant for many companies of the manufacturing industry in 2020 due to the Covid-19 pandemic. The challenge was to adapt existing processes with the aim of providing greatest possible safety for their employees, in particular in production. This resulted in a short-term switch from 1-shift to 2-shift work schedule in order to isolate groups of people from each other and to be able to keep greater distances to each other. It is possible to examine the impact of the shift work model in the simulation without having to make changes in the production itself. Furthermore, an analysis and evaluation is equally relevant for the current situation and for the future factory.

Therefore, the impact of the shift work model is examined in the case study. Two scenarios are compared using the simulation of the production area with constant amount of workers and quantities of products. Table 1 shows a comparison of the utilization of four workstations / machines for 1-shift and 2-shift work model. The listed workstations or machines reach their capacity limit of up to 100 % in the first scenario of 1-shift work model. This would result in the necessary measure of doubling the capacities to ensure the completion of orders and balanced production processes.

Table 1: Comparison of capacity utilisation for 1-shift and 2-shift work schedule

Workstation	Utilisation (1 shift)	Utilisation (2 shift)	Measure for 1-shift scenario
1	96 %	56 %	Doubling of capacities
2	100 %	84 %	Doubling of capacities
3	95 %	66 %	Doubling of capacities
4	100 %	88 %	Doubling of capacities

However, the analysis encourages the planner to look for the reasons of the high utilisation rate of several stations in the 1-shift scenario. On the one hand, bottlenecks in the production area are identified. On the other hand, the results provide indications of possible deficiencies such as modified operations. One example is that, in practice, an operation might not be performed at the assigned workstation. It happens that the operator performs the task at a different place to avoid congestion at the designated workstation. Such deficiencies are revealed by the simulation study for the current situation and should be considered in resource planning.

In 2-shift operation, the utilisation is significantly lower. In this scenario, bottlenecks of the listed workstations or machines do not occur. This effects resource planning and dimensioning significantly; unnecessary investments are avoided.

The simulation study enables the identification of bottlenecks, in the current state and in future scenarios. The impact of the shift work model on resource planning is shown. In the planned factory, there will be 2-shift work. For further planning, increasing quantities have to be considered which will lead to higher utilisation of the workstations. Therefore, further scenarios should be analysed for the dimensioning of resources.

3.3.2 Comparison of different assembly concepts

A second example of the simulation study deals with the comparison of different assembly concepts. The assembly operations are mainly performed on individual workstations. In the current state, only some components are pre-assembled on different workstations. The pre-assembled components are supplied to the workstation (consumption-controlled). New orders with batch size 1 are initialized at the main assembly workstations.

Analysing the assembly processes, further sub-assemblies are identified that can be pre-assembled and supplied to the main workstations. In this new scenario, the pre-assembly is extended for most products, which leads to a reduction of main assembly times and to a higher utilisation rate of the workstations for pre-assembly operations. This scenario is depicted in Figure 4.

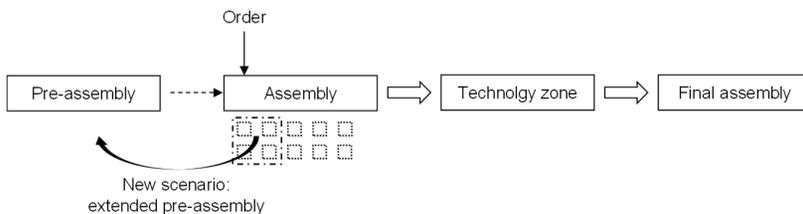


Figure 4: Overview on the operations showing the scenario of extended pre-assembly

The lead time, that is the time between the initialization at the assembly operation and the completion of the order, is reduced. Figure 5 shows the expected reduction of lead times obtained in the simulation. A reduction of lead times is expected for almost all products of up to 23.5 %. Lead times are reduced by 6.5% on average for all products.

In the current state, the lead times are not precisely recorded. However, it is possible to determine the lead times und to compare the different assembly concepts using simulation.

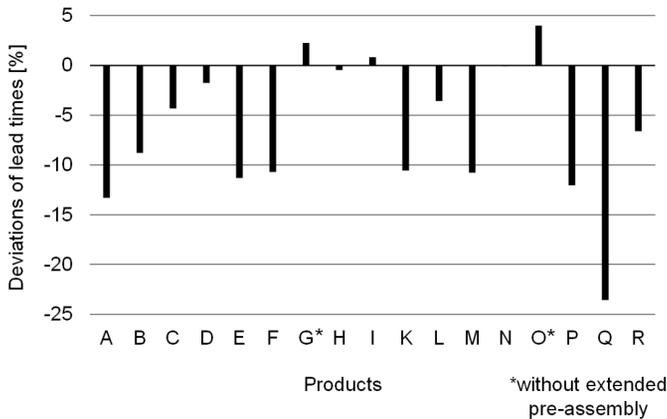


Figure 5: Deviation of lead times between scenario 1 and 2 for all products

3.3.3 Conclusion and Discussion

The examined production area is a complex system and therefore, suitable for an analysis using simulation. There are different assembly operations, high variation of sequences and multiple machine operations resulting in less transparency. Simulation is a method to manage the complexity in the production area. Bottlenecks are identified and the amount of required resources are determined. The simulation study supports manual planning and reveals errors in the planning process. The manually planned capacities are validated and unnecessary investments are avoided. Different scenarios can be compared easily without having to make changes in the production itself. In addition, the simulation tool provides a 3D visualisation of the production area, as depicted in Figure 6.

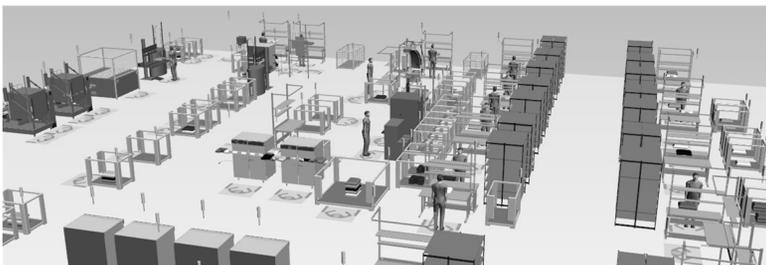


Figure 6: Visualisation of the production area in Plant Simulation

The results of the simulation study are integrated into the planning process. Due to the high percentage of manual assembly operations, the recorded times are fluctuating in practice, for example due to the dependence on the level of experience of the workers. Such deviations have to be taken into account when evaluating and interpreting the results. The generation of an adaptable simulation model creates long-term benefits. If modifications are made during planning, the changes can be tested and evaluated with little effort using the adaptable simulation model.

4 Summary and Outlook

In this paper, a case study on simulation for factory planning is presented. It comprises the planning of a new factory in the medical technology industry. Requirements, objectives, procedures and results of the simulation study of a production area are presented. The database is pointed out and a concept for integrating engineering tools is described. A central provision of technical data and models improves the information flow and data exchange avoiding manual data transfers. The simulation study supports the layout planning process providing a basis for dimensioning of resources, optimization of processes and a 3D visualization. The results of the simulation study are presented exemplary showing the impact of shift work model on the dimensioning and a comparison of different assembly concepts.

Layout planning will be performed for further production lines supported by simulation studies. The interfaces within the integrated tool landscape are under development and will be evaluated in future applications. Future research will be on the implementation of optimization algorithms.

References

- Dombrowski, U.; Ernst, S.: Scenario-based simulation approach for layout planning. *Procedia CIRP* 12 (2013) pp. 354-359.
- Grundig, C.-G.: *Fabrikplanung – Planungssystematik Methoden Anwendungen*. München: Hanser 2018.
- Krüger, T.: *Entwicklung einer Gesamtmethodik zur Kombination von mathematischer Anordnungsoptimierung und Materialflusssimulation für die Produktionslayoutplanung*. Düren: Shaker 2019.
- Mieschner, M.; Mayer, G.: Bereitstellung und Verwaltung von Simulationseingangsdaten. In: Mayer, G.; Pöge, C.; Spieckermann, S.; Wenzel, S. (Eds.): *Ablaufsimulation in der Automobilindustrie*. Berlin: Springer 2020, pp. 319-332.
- Mourtzis, D; Doukas, M., Bernidaki, D.: Simulation in manufacturing: review and challenges. *Procedia CIRP* 25 (2014) pp. 213-229.
- InFranken: Siemens Healthineers baut 24 Meter hohes Technologiecenter in Forchheim, 2020. <https://www.infranken.de/lk/forchheim/siemens-baut-24-meter-hohes-technologiecenter-in-forchheim-art-5136063>, accessed May 5th, 2021.
- Rabe, M.; Spieckermann, S.; Wenzel, S.: *Verifikation und Validierung für die Simulation in Produktion und Logistik - Vorgehensmodelle und Techniken*. Berlin: Springer 2008.
- Sender, J; Wanner, M.-C.: Integrierte Kapazitäts- und Layoutplanung. *Zeitschrift für wirtschaftlichen Fabrikbetrieb* 108 (2013) 3, pp.109-112.
- Süß, T.; Fischer, M.; Huber, D.; Laroque, C.; Dangelmaier, W.: Ein System zur aggregierten Visualisierung verteilter Materialflusssimulationen. In: Gausemeier, J.; Grafe, M. (Ed.): *Augmented & Virtual Reality in der Produktentstehung*, Band 252, Paderborn HNI Verlagsschriftenreihe 2009, pp. 111-126.
- VDI 3633 Part 1: *Simulation of systems in materials handling, logistics and production, Fundamentals*. Berlin: Beuth 2014.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Einsatz der ereignisdiskreten Materialflusssimulation bei Methoden aus der kontinuierlichen Verbesserung im Produktionsumfeld: Erfolg und Auswirkungen der Speedweek 4.0

***Use of discrete-event material flow simulation in continuous
improvement methods in the production environment: success and
impact of Speedweek 4.0***

Sarah Sitz, Maximilian Zerreis, Robert Bosch GmbH, Nürnberg (Germany),
sarah.sitz@de.bosch.com

Tobias Lechler, Jörg Franke, Lehrstuhl für Fertigungsautomatisierung und
Produktionssystematik (FAPS), FAU Erlangen-Nürnberg, Erlangen (Germany),
tobias.lechler@faps.fau.de, joerg.franke@faps.fau.de

Abstract: Derived from Kaizen, there are various approaches to improvement work, such as Lean or the Toyota Production System (Chiarini 2012; Alukal 2006). The further development of improvement systems is subject to constant challenges, currently by the challenge of the digital transformation. A method of continuous improvement work, known as Kaizen Event or also Speedweek, focuses very strongly on the on-site improvement work in production. The method is also timely and thematically limited and thus only allows a limited scope of improvements. Based on this method, it is made evident how the use of event-discrete material flow simulation can increase efficiency as well as raise potentials. Application examples from Robert Bosch GmbH are used to demonstrate effects. Subsequently, success criteria and risk factors from the adaptation of the improvement method with event-discrete material flow simulation are classified and evaluated.

1 Einleitung

Mit der Entwicklung der industriellen Massenproduktion hat sich auch die Systematik der Produktionssteuerung und -verbesserung kontinuierlich entwickelt. Im Zuge zunehmender Digitalisierung ergeben sich neue Herausforderungen für die Produktion von heute. Die „Digitale Fabrik“ und „Digitale Zwillinge“ sind in aller Munde, weshalb sich die Frage stellt, wie dies die bestehenden Methoden der Verbesserungsarbeit beeinflusst und zur Weiterentwicklung beiträgt. Insbesondere

der Einsatz simulativer Methoden wie der ereignisdiskreten Materialflusssimulation nimmt in der Produktion und Logistik stark zu, ist aber aktuell noch wenig in den systematischen Ansätzen aus der kontinuierlichen Verbesserungsarbeit verankert (Lechler et al. 2020). Dieser Beitrag beleuchtet daher den Einsatz der ereignisdiskreten Materialflusssimulation in fokussierter, produktionsnaher Verbesserungsarbeit in der Produktion der Automobilzuliefererindustrie.

2 Stand der Technik

Die Grundlage der systematischen und kontinuierlichen Verbesserungsarbeit in der Produktion (auch engl. Continuous improvement (CIP)) stellt die aus dem Japanischen übernommene Idee des Kaizen dar. Hosono et al. (2020) definieren Kaizen als einen inklusiven Ansatz für die kontinuierliche Verbesserungsarbeit mit einem Set an Methoden, das auf einer japanischen Grundphilosophie aufbaut. Verschiedene Managementsystematiken in der Produktion beinhalten einen Kaizen-Ansatz oder bauen direkt auf der Grundsystematik auf. Hierzu zählen: Japanese Total Quality Control (JTQC), Total Quality Management (TQM), Business Process Reengineering (BPR), Lean Thinking und Six Sigma sowie das Toyota Production System (TPS) (Hosono et al. 2020; Hodge et al. 2011; Chiarini 2012). Letzteres ist nach der Entwicklung in Japan in abgeleiteter und angepasster Form in vielen Ländern im Einsatz (Hosono et al. 2020). So baut auch die in der Robert Bosch GmbH eingesetzte Systematik, das Bosch Production System (BPS), auf dem TPS auf und beinhaltet viele Aspekte aus dem Lean.

Die verschiedenen Ansätze unterscheiden sich durchaus in Bezug auf die Verbesserungsarbeit. Ein detaillierter Vergleich findet sich in (Chiarini 2012). Die grundlegenden Erfolgsfaktoren sind neben der starken Kundenorientierung bei der Verbesserungsarbeit die Eingebundenheit der Mitarbeiter sowie der Stil des Managements (Chiarini 2012). Kulturelle Faktoren, vor allem des Kulturraums an dem sich das Unternehmen orientiert, spielen eine große Rolle bei der erfolgreichen und langfristigen Auswirkung (Chiarini 2012). So ist Six Sigma ein Ansatz mit starker hierarchischer Ausrichtung und einem Fokus auf externe Motivation der Mitarbeiter im Kontrast zu den ursprünglichen Ansätzen aus dem asiatischen Kulturraum. Der Aspekt der konzentrierten, schnellen Verbesserungsarbeit („Speed“) ist ebenfalls stark mit den Ansätzen des Lean Managements sowie Six Sigma aufgekommen. Im kombinatorischen Ansatz Lean Six Sigma ist zusätzlich die Eingebundenheit der Mitarbeiter betont (Chiarini 2012; George 2002).

In allen Ansätzen der kontinuierlichen Verbesserungsarbeit haben sich verschiedene Vorgehensweisen entwickelt, wie die Verbesserungsarbeit kontrolliert methodisch ablaufen kann. In diesem Aspekt ähneln sich viele der systematischen Ansätze. Alle setzen auf Problemlösungsmethoden, sind dabei jedoch unterschiedlich stark formalistisch in den Vorgaben zu Einsatz und verwendeten Methoden.

Sowohl in stark vorstrukturierten als auch in eher freieren Ansätzen der kontinuierlichen Verbesserung findet sich ein wiederkehrendes Vorgehen. So gibt es immer zeitlich begrenzte Projekte oder Vorhaben, die meist eine Woche andauern und die Mitarbeiter einbinden. Diese sind je nach zugrundeliegender Systematik als Kaizen Event, Kaizen Week, Kaizen Blitz bekannt und unterscheiden sich leicht im Einsatz und in der Darstellung in der Literatur (Chiarini 2012; Alukal und Manos

2006; Manos 2007; Coffey and Thornley 2006). Aus diesen leitet sich auch die Speedweek ab, deren Umfang weiter unten detailliert beschrieben wird.

Neben der zeitlichen Eingrenzung auf zwei bis fünf Tage beinhalten Kaizen Events einen thematischen Fokus, an dem teambasiert gearbeitet wird. Die Teams sind in der Regel funktionsübergreifend und für die Dauer des Projektes nur diesem zugeordnet. Grob werden drei Schritte unterschieden: Planung des Events, Durchführung und Follow-ups (Gan et al. 2014). Diese Schritte können weiter ausdetailliert werden, Beispiele hierzu finden sich etwa in (Gan et al. 2014; King 2019).

In der Literatur gibt es verschiedene Einsatzbeispiele. Neben der Anwendung in der Automobilindustrie zeigt der Einsatz von Lean im britischen Gesundheitssystem NHS (National health system), dass mit Einsatz eines Tools wie Kaizen Blitz Verbesserungen erzielt werden können (Radnor et al. 2012). Viele Studien beschäftigen sich mit den Erfolgsfaktoren. So ist das Bekenntnis (engl. commitment) des Managements ein elementarerer Faktor (García et al. 2014). Ebenso wird die Eingebundenheit der Mitarbeiter genannt, so sei es wichtig neben der methodischen Vorgehensweise auch einen Sinn zu vermitteln, um eine längerfristige Auswirkung zu erzielen. (Radnor et al. 2012) Die langfristige Wirksamkeit von Verbesserungsmaßnahmen ist allgemein als kritischer Faktor bekannt und wirkt sich bei zeitlich begrenzten Vorhaben besonders aus (Radnor et al. 2012; Hines et al. 2004).

Simulationen finden zunehmend in der kontinuierlichen Verbesserungssystematik ihren Einsatz, mit einem starken Anstieg seit 2005. Neben klassischen Nachstellungen über sogenannte Mock Ups können auch digitale Lösungen wie mathematische Modelle oder die ereignisdiskrete Materialflusssimulation zum Einsatz kommen. Einen Überblick bietet u.a. Goienetxea Uriarte et al. (2020). Die Schwerpunkte beim Einsatz sind vor allem Verbesserungsmethoden wie das Value Stream Mapping (VSM), aber auch die Kanbanmethodik oder die Taktzeitsimulation. Erste Veröffentlichungen zeigen, dass auch in einer begrenzten, kurzen Zeit mit ereignisdiskreter Simulation gearbeitet werden kann sowie deren positive Effekte (Baril et al. 2016; Robinson et al. 2012). So wird in Studien im Gesundheitssektor eine Steigerung der Performance von 15% mit Unterstützung der ereignisdiskreten Materialflusssimulation erzielt (Robinson et al. 2012). Auch in anderen Industrien gibt es entsprechende Studien, insbesondere dort, wo die Simulation stark verbreitet ist, wie im Gesundheitswesen oder in der Verpackungsindustrie. Stand der Technik ist der strukturierte Einsatz allerdings bisher keineswegs (Goienetxea Uriarte et al. 2020).

3 Beschreibung der Methodik „Speedweek“ als angepasste Systematik in der kontinuierlichen Verbesserung innerhalb des Bosch Production Systems (BPS)

Die Methode „Speedweek“ ist eine konzentrierte Verbesserungsmethode, bei der ein Team aus erfahrenen Mitarbeitern ohne oder nur mit geringem Bezug zum Thema (Fresh Eyes) innerhalb einer Woche an einem Fokusthema arbeitet. Ähnliche Methoden sind in der Literatur bekannt als Speed Kaizen, Kaizen Blitz oder auch Kaizen Event (Chiarini 2012).

Elemente wie die direkte Beteiligung von Mitarbeitern aus verschiedenen Hierarchieebenen an den Verbesserungsmaßnahmen und die Arbeit vor Ort als Go to Gemba sind als weitere Kernpunkte der Speedweek enthalten. Go to Gemba oder auch Gemba Walk bezeichnet die Beobachtung von Arbeit, wie diese von den Mitarbeitern an ihrem Arbeitsort stattfindet und leitet sich ab aus dem japanischen Wort “gembutsu”, “echter Ort” ab (Dalton 2019). Eine Speedweek hat einen festen Ablauf und eine begrenzte Dauer von fünf Arbeitstagen. Innerhalb dieser Zeit arbeiten einzelne Teams an einem oder zwei Schwerpunktthemen. Die Zusammensetzung der Teams ist anhand von Rollen festgelegt. Neben einem Coach, der durch die Methodik führt, moderiert und sich um die Beseitigung organisatorischer Hindernisse kümmert, gibt es einen oder mehrere Teamleiter. Diese sind verantwortlich für den inhaltlichen Fortschritt der Arbeiten und leiten die Teams an. Innerhalb der Teams gibt es einen möglichst hierarchieübergreifenden Mix aus verschiedenen technischen Disziplinen. Ebenso muss ein Teil der Teilnehmer als sogenannte Fresh Eyes ohne oder nur mit geringen Fach- und Vorkenntnissen teilnehmen, um neue Perspektiven einzubringen. Außerdem empfiehlt es sich, eine entsprechende technische Supportstruktur zu schaffen, etwa um kurzfristig technische Daten einzusehen oder um Änderungen vornehmen zu können. Die Zusammensetzung der Teilnehmer sowie der grundlegende Ablauf einer Speedweek inklusive der Vor- und Nacharbeiten ist in Abbildung 1 dargestellt und wird im Folgenden kurz beschrieben.

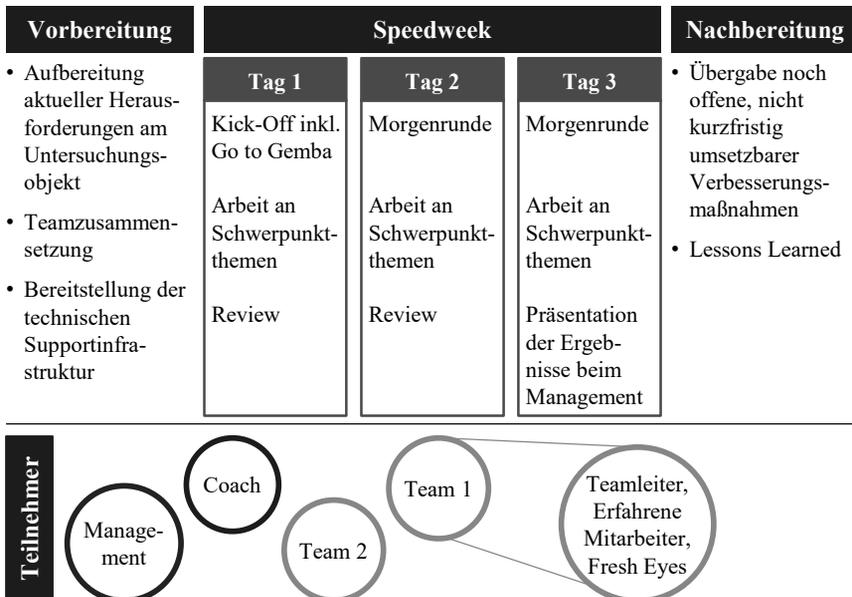


Abbildung 1: Systematischer Ablauf und Teilnehmer einer Speedweek

Nach einer Einführung, etwa in die Problemstellung, die Fertigungslinie, aber auch das gefertigte Produkt oder die Methodik allgemein, beginnt zunächst eine Aufnahme. Anschließend gehen die Teams in die Analyse und/oder in die direkte Umsetzung von Maßnahmen. Dabei können auch Fehleranalysemethoden wie 8D oder DMAIC

(Akronym für Define, Measure, Analyze, Improve and Control (Chiarini 2012)) eingesetzt werden. Der Fokus liegt jedoch auf Sofortmaßnahmen und schneller Umsetzung dieser. Es gibt tägliche Berichtsrunden, etwa am Ende oder Anfang eines jeden Tages, um den Arbeitsfortschritt zu verfolgen. Zum Ende der Woche werden die Resultate dem Management vorgestellt und nicht kurzfristig umsetzbare oder noch offene Aufgaben werden systematisch an die Auftraggeber übergeben.

Die wesentlichen Erfolgsfaktoren im Rahmen dieser Methode sind ein klarer Fokus auf umsetzbare, relevante Fragestellungen, die Arbeit unmittelbar an der Fertigungslinie und ein diverses Team mit einer strukturierten Teamleitung. Die Fragestellung muss in der Vorbereitung der Speedweek festgelegt werden, ebenso wie das Team-Staffing. Ohne die Möglichkeit, vor Ort in der Fertigungslinie aktiv zu werden, ist der Einsatz der Methode „Speedweek“ sehr genau zu prüfen. Ist das nicht gegeben, sollten andere Methoden in Betracht gezogene werden. Der Aspekt der langfristigen Wirksamkeit der Methode ist ebenfalls ein wichtiger Punkt. Durch nachträgliche und regelmäßige Erfolgskontrollen wird sichergestellt, dass alle Maßnahmen umgesetzt und deren Wirksamkeiten gemessen werden.

Die Speedweek kann gegenüber den in der Literatur beschriebenen Methoden Kaizen Event oder Kaizen Blitz wie folgt eingeordnet werden: Alle wesentlichen Elemente sind sich ähnlich oder gleich, so dass die Speedweek ebenfalls als Kaizen Event bezeichnet werden könnte. Kleinere Unterschiede liegen in der starken Betonung der Arbeit vor Ort sowie in der thematischen Fokussierung und Strukturierung der Woche sowie in der klar definierten funktions- und hierarchieübergreifenden Teamstruktur. In der Literatur finden sich ähnliche Ausprägungen, jedoch keine so umfassende und eindeutige Definierung des Vorgehens, der Teamrollen und des Ablaufs (Chiarini 2012; García et al. 2014; Gan et al. 2014; King 2019).

4 Einsatz der ereignisdiskreten Materialflusssimulation als Erweiterung der Grundmethodik

Die oben beschriebene Methode der Speedweek ist innerhalb der Robert Bosch GmbH um den Einsatz der ereignisdiskreten Materialflusssimulation erweitert worden. Eine ereignisdiskrete Materialflusssimulation wird im Produktionsumfeld normalerweise zur Lösung von bestimmten Fragestellungen eingesetzt, etwa zur Layoutplanung oder um die optimale Entkopplung eines Produktionssystems zu finden. Der Simulationseinsatz findet in der Regel mit direktem Projektbezug über einen längeren Zeitraum statt. In der kontinuierlichen Verbesserung kann das etwa ein System CIP Projekt sein (Chiarini 2012).

Beim Einsatz der Simulationssoftware für die ereignisdiskrete Materialflusssimulation im Kontext der kontinuierlichen Verbesserungsarbeit ergeben sich Herausforderungen.

So ist etwa einer der Schüsselfaktoren in der kontinuierlichen Verbesserungsarbeit die Einbindung der Mitarbeiter. Auf Simulationsmethoden kann das jedoch nur begrenzt übertragen werden. Meist werden Simulationsmodelle nur von Experten aufgebaut und bedient, da entsprechende Fachkompetenzen nötig sind. Ebenso benötigt die Software entsprechende Hardware, dies widerspricht sich scheinbar mit dem Fokus auf Verbesserungsarbeit vor Ort in der Produktion. Zu diesen Aspekten wird es in Abschnitt 5 eine Einordnung geben.

Grundsätzlich ist Verbesserungsarbeit aus eigener Kraft und nicht durch einen Invest in neue, bessere Maschinen zu leisten (Hosono et al. 2020). Ebenso liegt ein Fokus auf die Vermeidung oder Reduzierung von Verschwendung (engl. Waste). Eine Simulation verursacht jedoch in der Regel initial Mehrkosten, ebenso bietet eine digitale Arbeit meist nur indirekte Wertschöpfung. Diese vermeintlichen Widersprüche sind exemplarisch in der Literatur untersucht worden. So zeigt eine Gegenüberstellung von Innovationen und Verbesserungsarbeit nach der Philosophie des Kaizens, dass es einen gegenseitigen positiven Einfluss gibt. So kostet eine Simulation zwar Geld, reduziert nur indirekt Verschwendung, hebt aber gleichzeitig ein hohes Potential durch Effizienzsteigerung und wirkt so inkrementell (Hosono et al. 2020). Ein gewichtiges Argument ist auch, dass bei Einsatz einer ereignisdiskreten Materialflusssimulation die Verbesserungsansätze am digitalen Modell erprobt werden können und somit Stillstände in der Produktion vermieden oder zumindest deutlich verringert werden können.

Dass der Einsatz von ereignisdiskreter Simulation zur Steigerung der Effizienz von Verbesserungsmaßnahmen beiträgt, wird auch in der Literatur bestätigt (Robinson et al. 2012).

5 Entwicklung zur Speedweek 4.0

Für die Weiterentwicklung zur Speedweek 4.0 ist die Integration zweier etablierter Methodiken erforderlich. Dabei steht auf der einen Seite der klassische Ablauf der Speedweek, auf der anderen Seite die Durchführung von Simulationsstudien nach der VDI 3633 (Verein Deutscher Ingenieure 2014). Dies bedeutet Eingriffe sowohl in die Vor- und Nachbereitung als auch in die eigentliche Durchführung der Verbesserungsarbeit.

In einer Speedweek 4.0 sind vor allem Anpassungen in der Teamzusammensetzung vorzunehmen. So ist ein Teammitglied zu ergänzen, das die Simulationssoftware bedienen kann und über ausreichende Problemlösekompetenz innerhalb der Simulation verfügt. Der Ablauf einer Speedweek 4.0 bleibt grundsätzlich unverändert gegenüber einer Speedweek, ebenso der Aspekt der Arbeit vor Ort.

Die Vorgehensweise bei der Speedweek 4.0 kann mit Vorgehen für Simulationsstudien wie folgt abgeglichen werden: Simulationsstudien beginnen mit einer Zielbeschreibung gefolgt von der Aufgabendefinition. Dies deckt sich mit der Vorbereitungsphase einer klassischen Speedweek. Während die Aufgaben und Ziele für das gesamte Projektteam festgelegt werden, gilt es aus Sicht der Simulation relevante Problemstellungen zu extrahieren. Das übergeordnete Ziel einer Speedweek basiert auf der Verbesserungsarbeit an einem Produktionssystem und richtet sich damit an dessen Leistungskennzahlen. Dazu gehören beispielsweise die Gesamtanlageneffektivität oder der Kundentakt einer Serienproduktion. Verbesserungen einzelner Abläufe oder Komponenten der ausgewählten Produktion in der klassischen Speedweek sollen eine Optimierung dieser Kennzahlen bewirken. Dieser Wirkzusammenhang, stellt das grundlegende Anwendungsfeld der Simulation in einer Speedweek 4.0 dar. Die Materialflusssimulation ermöglicht die quantifizierte Evaluierung von Optimierungsmaßnahmen hinsichtlich dieser Kennzahlen. Zusammen mit dem geschätzten Aufwand lässt sich eine Priorisierung der Optimierungsmaßnahmen vornehmen, was die Verbesserungsarbeit beschleunigt und somit das Produktionssystem bei definiertem Aufwand schneller optimiert. Die

Simulation liefert die notwendige Entscheidungsgrundlage für das Projektteam und hilft bei der Sicherstellung, dass an den richtigen Themen gearbeitet wird. Um während der Durchführung möglichst schnell Aussagen zu treffen, ist es notwendig den Aufbau des Simulationsmodells in der Vorbereitungsphase abzuschließen. Die Phasen einer Simulationsstudie bis einschließlich zur Implementierung nach der VDI 3633 finden somit in der Vorbereitung der Speedweek 4.0 statt. Für eine reibungslose Integration der Simulationsergebnisse in die Verbesserungsarbeit im Projektteam ist in der Vorbereitung auf ein vorhandenes Grundverständnis der Beteiligten mit Blick auf die Materialflusssimulation zu achten. Dieses kann mit einem Einführungskurs aufgebaut werden.

Während der Durchführung der Speedweek 4.0 befindet sich die Simulationsstudie somit in der Phase der „Experimente und Analyse“. Der Simulationsexperte ist Teil des Projektteams und in ständigem Austausch mit dem Teamleiter sowie den anderen Teilnehmern. Zu Beginn der Speedweek 4.0 erfolgt das Go to Gemba. Hierbei entstehen Ideen, wo die Verbesserungsarbeit ansetzen kann. Diese Ideen können durch das Team priorisiert und in der Simulation getestet und bewertet werden. Die Ansätze mit der optimalen Auswirkung werden umgesetzt. Diese Zusammenarbeit zwischen Team und Simulationsexperte bedeutet kurze Zyklen zur Bewertung und Entscheidung über das weitere Vorgehen. Neben der Priorisierung von aussichtsreichen Optimierungsmaßnahmen trägt der Einsatz der Simulation auf diesem Weg zudem zur Vermeidung von Verschwendung durch nicht effektive Maßnahmen bei. Auf Basis der bewerteten Optimierung werden Ressourcen optimal eingesetzt und Zeit eingespart. Die Simulation erlaubt außerdem die Rechtfertigung von Aufwand oder Investitionen durch die quantitative Bewertung des Nutzens. Außerdem muss erwähnt werden, dass die dreidimensionale Visualisierung von dynamischen Abläufen, welche heutige Simulationsprogramme bieten, den Mitarbeitern den Zugang zur Simulation ermöglicht. Sie können die Vorgänge am Modell gut mit dem vergleichen, was an der realen Linie zu sehen ist und lernen die Einsatzmöglichkeiten (und Grenzen) kennen.

Nach Abschluss der Speedweek 4.0 folgt in der Nachbereitung die Übergabe noch offener Maßnahmen an die Verantwortlichen, ebenso wie die dokumentierten Ergebnisse aus der Simulation. Diese ergänzen den Maßnahmenplan um bewertete Potentiale und sorgen somit auch über die Speedweek 4.0 hinaus für den nachhaltigen Einsatz der Ressourcen. Die organisatorischen Randbedingungen sowie die diskutierten inhaltlichen Ergänzungen sind in Abbildung 2 zusammengefasst dargestellt.

Vorbereitung	Speedweek			Nachbereitung	
<ul style="list-style-type: none"> • Aufbereitung aktueller Herausforderungen am Untersuchungsobjekt • Teamzusammensetzung • Bereitstellung der technischen Supportinfrastruktur • Aufbau des Simulationsmodells anhand definierter Aufgabenstellung 	Tag 1	Tag 2	Tag 3	<ul style="list-style-type: none"> • Übergabe noch offene, nicht kurzfristig umsetzbarer Verbesserungsmaßnahmen • Lessons Learned • Übergabe des Simulationsmodells an Fachbereich 	
	Kick-Off inkl. Go to Gemba	Morgenrunde	Morgenrunde		
	Arbeit an Schwerpunktthemen	Arbeit an Schwerpunktthemen	Arbeit an Schwerpunktthemen		
	Modellvalidierung und -verifikation	Simulative Evaluierung von Verbesserungsmaßnahmen	Simulative Evaluierung von Verbesserungsmaßnahmen		
	Review	Review	Präsentation der Ergebnisse beim Management		

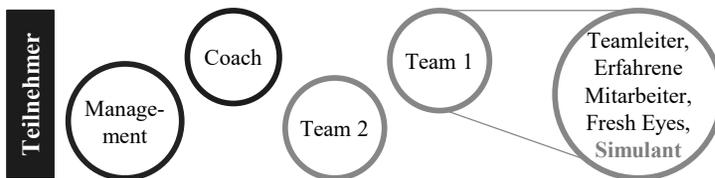


Abbildung 2: Anpassungen bei der Speedweek 4.0

6 Anwendung der Methodik in der Produktion: Anwendungsbeispiele aus dem Fertigungswerk Nürnberg der Robert Bosch GmbH

Im Folgenden wird die Durchführung einer Speedweek 4.0 anhand von vier Anwendungsbeispielen beschrieben. Alle Speedweeks 4.0 haben eine Performancesteigerung zum Ziel, zeichnen sich aber durch unterschiedliche Problemstellungen aus. Die Use Cases haben gemein, dass ein Simulationsmodell, welches den aktuellen Status der Produktionssysteme vereinfacht abbildet, in der Vorbereitungsphase aufgebaut wird. Das Modell stellt für einen geplanten Produktionszeitraum den Materialfluss des jeweiligen Produktionssystems auf Prozessebene dar. Als Parameter werden hierfür die gemittelten Taktzeiten und Verfügbarkeiten der Prozesse mit einbezogen. Damit ist die schnelle Bewertung von grundlegenden Optimierungsmaßnahmen sichergestellt.

Use Case A: Hauptmontagelinie

Gegenstand dieses Use Case ist ein Produktionssystem bestehend aus einem Montage- und Prüfbereich. Zwei Teams arbeiten während der Speedweek unter dem Gesamtziel der Performancesteigerung an zuvor konkretisierten Fokusthemen. Dazu gehört zum einen die Steigerung der Gesamtanlagenperformance, insbesondere die Reduzierung

der Anlaufprobleme nach den Rüstvorgängen im Montagebereich. Zum anderen steht der Prüfbereich im Fokus. Dieser unterteilt sich in mehrere parallel arbeitende Prüfstände, die durch unterschiedliche Performance aufgefallen sind. Es werden Ursachen analysiert und mittels Maßnahmen Ausschuss und Störungen reduziert. Der Prüfbereich stellt ein komplexes System dar, welches einen hohen Analyseaufwand zur Ableitung von Optimierungsmaßnahmen benötigt. Die Simulation unterstützt diese Analyse mit dem Aufzeigen wesentlicher Ursache-Wirk-Zusammenhänge. Auf Basis des Simulationsmodells wird der Einfluss unterschiedlicher Performance-Niveaus analysiert sowie die Auswirkungen beim Abschalten einzelner Prüfstände bewertet. Während einzelne Prüfstände bisher unter Einbezug der aktuellen Ausschussrate abgeschaltet wurden, können jetzt auch die Auswirkungen auf die Produktionskennzahlen in die Entscheidung mit integriert werden. Dies ermöglicht die Erstellung eines zielgerichteten Reaktionsplans und damit die Optimierung der operativen Produktionssteuerung, was sich über die Speedweek 4.0 hinaus positiv auf die Performance auswirkt.

Use Case B: Vormontage

In Use Case B wird eine Speedweek 4.0 an einer hochautomatisierten Vormontagelinie mit sieben über Fördersysteme verbundenen Stationen durchgeführt. Da der Linientakt gegenüber der zugehörigen Hauptmontagelinie abweicht, fokussieren sich die Teams auf eine Taktzeitoptimierung sowie auf eine Reduzierung von technischen Störungen, um den Linientakt der Vormontage anzugleichen. Das Simulationsmodell ist im Voraus aufgebaut worden und spiegelt auf Basis von aufbereiteten Parametern den aktuellen Status der Linie. Ideen aus den beiden Teams zur Optimierung einzelner Prozesse werden innerhalb der ersten beiden Tage der Speedweek 4.0 am Modell getestet. Diese umfassen neben Reduktionen der Taktzeiten auch verbesserte Verfügbarkeiten. Die Simulation wird zur Bewertung und zum Nachweis der Nachhaltigkeit der Optimierungen genutzt. Im Fokus dieser Speedweek 4.0 stehen zudem Mitarbeiteraufgaben, die Störungsbehebung, Materialversorgung und Teileverpackung umfassen. Eine vereinfachte Umsetzung der Aufgabensteuerung in der Simulation zeigt einen hohen Einfluss der Mitarbeiterauslastung auf die Gesamtanlageneffektivität. Auf Basis dieser Erkenntnis werden Optimierungsmaßnahmen abgeleitet. Im konkreten Fall kann eine wesentliche Entlastung des Mitarbeiters durch ein zusätzliches Puffersystem beim Verpacken erzeugt werden. Die wirkt sich gleichzeitig reduzierend auf den Linientakt aus. Auf Basis der quantitativen Bewertung des Potentials von 4 - 6 % wird diese Investition vom Management bewertet und unterstützt und führt zu einer nachhaltig verbesserten Performance. Das Beispiel zeigt den sinnvollen Einsatz von Investitionen und wirkt sich nachhaltig aus. Das Simulationsmodell wird auch nach der Speedweek 4.0 für die Bewertung weiterer Ideen genutzt.

Use Case C: Hauptmontagelinie

Dieser Use Case fokussiert auf eine Performancesteigerung beim Rüsten. Gegenstand der Verbesserungsarbeit ist ein hochautomatisiertes Produktionssystem mit starr verketteten Prozessen. Die beiden Teams arbeiten an der Reduzierung der Rüstlücke, sowie an der Beseitigung der Anlaufstörungen nach dem Rüsten. Der Großteil der Rüstvorgänge wird manuell von drei Mitarbeitern durchgeführt. Im Simulationsmodell werden die Aufgaben systematisch zugeteilt. Potentiale zur Reduzierung einzelner Aufgaben aus einer Neubewertung durch die Arbeitsplanung

und Zeitwirtschaft werden direkt in der Simulation implementiert und hinsichtlich der Rüstlücke bewertet. Fortlaufend wird die Aufgabenverteilung überprüft und neue Szenarien getestet. Die Simulationsergebnisse werden kurzzyklisch in die Verbesserungsarbeit des Rüstablaufs einbezogen. Die 3D-Visualisierung ermöglicht es zudem, den abschließend optimierten Rüstablauf grafisch verständlich darzustellen. Hierdurch können die betroffenen Mitarbeiter anhand der Simulation in den neuen Ablauf eingeführt werden und diesen nachvollziehen. Außerdem werden kurze Videosequenzen erstellt, die auch nach der Speedweek 4.0 zur Weiternutzung zur Verfügung stehen. Das Simulationsmodell der Linie steht außerdem zur Weiternutzung zur Verfügung. Insgesamt konnte durch die Verbesserungsmaßnahme in Kombination mit der Simulation die Rüstlücke um etwa 30 - 40 % reduziert werden.

Use Case D: Zentrenfertigung Zerspanung

Im Use Case D wird eine Speedweek 4.0 in einer Zentrenfertigung mit Zerspanungsmaschinen durchgeführt. Im Fokus stehen die Reduzierung der Ausschussrate sowie ein Angleichen des Nutzungsgrads einzelner Maschinen auf den aktuellen Benchmark mit dem Ziel einer Performancesteigerung der Produktion. Die Simulation wird zur Bewertung der Ausschussrate und der Störungen hinsichtlich der relevanten Produktionskennzahlen, vor allem dem Nutzungsgrad eingesetzt. Mit der Simulation wird der Einfluss der Mitarbeiterauslastung auf den Nutzungsgrad analysiert. Zusammen mit dem Team werden Optimierungsmaßnahmen diskutiert und abschließend bewertet. Es konnte ein negativer Einfluss durch weite Laufwege bei den Mitarbeitern aufgedeckt werden, der wesentlich den Nutzungsgrade beeinträchtigt. Entsprechende Optimierungen durch neue Positionen verschiedener Elemente und Arbeitsplätze konnten mithilfe der Simulation evaluiert und durch das Team umgesetzt werden. Daneben machen sich die Anpassungen auch direkt beim Mitarbeiter bemerkbar, da die Laufwege pro Mitarbeiter und Schicht um etwa 2 bis 4 km reduziert werden konnten.

Evaluierung der Use Cases

Die beschriebenen Use Cases werden anhand von vier Kriterien eingeordnet und evaluiert. Diese Kriterien orientieren sich an den in der Literatur genannten Erfolgsfaktoren (Chiarini 2012). Mit Inklusion wird die Einbindung von Mitarbeitern beschrieben. Diese sollte funktions- und hierarchieübergreifend sein, mit einem möglichst klaren Bekenntnis des Managements. Qualifizierung beschreibt den Umstand, dass bei einem Verbesserungsvorhaben nicht nur das Produktionssystem im Vordergrund steht, sondern dass auch die Mitarbeiter mit neuen Tools und Kenntnissen in Berührung kommen und sich so direkt oder indirekt weiterentwickeln können (Hosono et al. 2020). Dieser Aspekt wird gesondert aufgeführt, da er im Kontext des aktuellen Wandels in der Automobilindustrie als elementar angesehen wird. Potentiale bezieht sich auf die Wirksamkeit der Verbesserungsmaßnahmen auf das Produktionssystem und dessen Peripherie, während Nachhaltigkeit diese Wirksamkeit in einer mittelfristigen Zeitspanne nach der Speedweek 4.0 beschreibt. Die Einordnung ist in Tabelle 1 dargestellt.

Table 1: Bewertung der Anwendungsbeispiele

Use Case	Inklusion	Qualifizierung	Potentiale	Nachhaltigkeit
A	Begrenzt (Fokus eng, nur Kontakt mit wenigen Mitarbeitern)	Niedrig	Hoch (Ausbringungssteigerung und Lösung komplexer technischer Probleme)	Hoch (Reaktionsplan und weitere Maßnahmen)
B	Hoch (Beide Teams und direkte Kommunikation von Ideen von allen Teilnehmern)	Mittel (Erster Kontakt mit Software und deren Möglichkeiten)	Hoch (Ausbringungssteigerung)	Hoch (weitere Nutzung des Simulationsmodells, weitere Folgemaßnahmen identifiziert)
C	Hoch (direkte Auswirkung, Test und Einbringung von Ideen im Team)	Hoch (Ablaufsimulation unterstützt Schulungsmaßnahmen)	Hoch (Ausbringungssteigerung)	Hoch (weitere Nutzung des Simulationsmodells)
D	Hoch (direkte Auswirkung, Tests und Einbringung von Ideen im Team)	Mittel (Erster Kontakt mit Software und deren Möglichkeiten)	Hoch (Ausbringungssteigerung sowie reduzierte Laufwege für Mitarbeiter)	Hoch (weitere Verfolgung von Maßnahmen)

Zusammenfassend kann herausgestellt werden, dass der Fokus einer Speedweek 4.0 sich auch auf den Erfolg für den Simulationseinsatz auswirkt. Handelt es sich um eine eng gefasste, technisch komplexe Fragestellung, so erreicht dies nur einen Teil des Teams mit entsprechenden (Fach-)Kenntnissen. Die Reichweite der Simulation ist dann ebenfalls begrenzt, was die Zielerreichung einer Speedweek 4.0 nicht grundsätzlich verhindert. Ist es jedoch aufgrund der Fragestellung möglich, die Simulation in einem oder beiden Teams vorzustellen und mit den Ideen aller Teilnehmer zu befüllen, steigt die Effizienz deutlich. Neben den Auswirkungen auf die Ausbringungssteigerung entwickeln sich „Added Value Benefits“, wie die Beispiele zeigen. So entstehen zusätzliche Qualifizierungsmaßnahmen für die Mitarbeiter oder es entwickeln sich zusätzliche Potentiale wie die Reduktion der gelaufenen Kilometer pro Mitarbeiter und pro Schicht im Beispiel D. Diese Verbesserung könnte ohne Simulation nur sehr aufwendig bewertet und erprobt werden, jedoch nicht innerhalb von einem Zeitrahmen von 5 Tagen. Allein der erste Kontakt mit der Software stellt ebenfalls einen Erfolgsfaktor dar, da diese Software und ihre Möglichkeiten in der täglichen Arbeit für die Mitarbeiter direkt erfahrbar

gemacht werden und nicht nur anhand von (Management-)Präsentationen. Auch hier ist ein Kontakt zu möglichst allen Teammitgliedern unabhängig von der Hierarchie ideal und die Akzeptanz vor allem dann hoch, wenn ein direkter visueller Vergleich zwischen realer Produktion und 3D-Modell möglich ist.

Ebenfalls positiv beeinflusst wird die Nachhaltigkeit. Dieser Aspekt wird sehr oft kritisch bewertet. In der Regel findet eine Weiternutzung der Simulationsmodelle statt, dies bedingt auch gleichzeitig eine deutlich vereinfachte Nachverfolgung von Maßnahmen und erleichtert eine Erfolgskontrolle. Die Modelle werden für andere Fragestellungen weitergenutzt, so dass sich auch Folge-Speedweeks 4.0 mit verändertem Fokus anschließen können.

Ein Transfer der Erkenntnisse ist möglich, da die wesentlichen Punkte gut umsetzbar sind. Die Kernaspekte der Speedweek 4.0 sind ein klarer Fokus mit umsetzbaren, relevanten Fragestellungen, die Arbeit unmittelbar an der Fertigungslinie und vor Ort mittels Go to Gemba und ein diverses Team mit einer strukturierten Teamleitung sowie eine klare und enge Strukturierung des Vorgehens, der Teamrollen und des Ablaufs. Eine Teamrolle ist mit einem Simulationsexperten besetzt, dieser sollte bereits bei der Vorbereitung mit eingebunden werden. Die Fragestellung ist idealerweise so zu wählen, dass eine Simulation von allen Teammitgliedern aktiv mit Ideen befüllt werden kann.

Während der Fokus in den beschriebenen Anwendungsbeispielen im Rest dieses Beitrags vor allem auf der ereignisdiskreten Materialflusssimulation liegt, schließt es sich nicht aus, dass auch weitere Simulationstools je nach Fragestellung zum Einsatz kommen können. Liegt etwa ein digitales Verhaltensmodell einer Maschine vor, so kann dies auch für steuerungstechnische Fragestellungen eingesetzt werden. Allerdings ist vermutlich hier die Reichweite durch den Teilnehmerkreis entsprechend begrenzt. Auch sind Simulationsmodelle etwa von Montageanlagen aufwändig zu erstellen, so dass auch der Zeitfaktor limitierend sein könnte. In ersten Tests konnte jedoch kein Ausschluss für den Einsatz von digitalen Maschinenmodellen in Speedweeks 4.0 gefunden werden.

7 Fazit und Ausblick

Die digitale Transformation bringt neue Werkzeuge und Herausforderungen für die Produktion mit sich. Eines der Werkzeuge, die ereignisdiskrete Materialflusssimulation, ist bereits begleitend zu längerfristigen Projekten im Einsatz und gut untersucht.

Gleichzeitig entwickeln sich die systematischen Grundmethoden der kontinuierlichen Verbesserungsarbeit weiter. In der Literatur vermehrt zu findend ist eine Methode, die als Kaizen Blitz, Kaizen Event oder in diesem Beitrag als Speedweek bezeichnet wird. Die Grundkriterien einer Speedweek sind dabei neben einem definierten Fokus der Einsatz am realen Arbeitsort mittels Go to Gemba und organisatorische Rahmenbedingungen wie ein diverses Team mit klaren Rollen.

Ein Einsatz von simulativen Methoden in einer Speedweek 4.0 bringt verschiedene Herausforderungen mit sich, etwa die Sicherstellung der Eingebundenheit möglichst vieler Mitarbeiter auf verschiedenen Ebenen und Funktionen mit einem Tool, dass per se eine hohe Expertise zur Bedienung benötigt. Diese Herausforderungen sind gleichzeitig auch die größten Erfolgspotentiale, wie anhand von Anwendungs-

beispielen gezeigt werden kann. So wird über ein Heranführen der interdisziplinären Speedweekteams an ein Simulationsmodell eine hohe Dichte an Verbesserungsideen eingeholt, während die Mitarbeiter gleichzeitig in Kontakt mit der Software kommen und sich so (in kleinen Schritten) weiterentwickeln können. In den vorgestellten Anwendungsbeispielen aus der Produktion der Robert Bosch GmbH ist dies mehrfach positiv bestätigt worden, indem neben einer Performancesteigerung weitere Potentiale gehoben wurden. Grundlegende Voraussetzungen sind die Notwendigkeit einer Passung der Fragestellung zur Simulation, aber auch zur Teamstruktur der Speedweek 4.0. Eine hohe Reichweite kann erzielt werden, indem möglichst viele Teammitglieder eingebunden werden können. Dies wirkt sich auf die Effizienz bei der Erzielung einer Performancesteigerung aus, steigert aber auch die zusätzlichen Potentiale. Die Nachhaltigkeit einer Speedweek kann durch Simulationseinsatz verbessert werden, da die Weiterarbeit mit den erstellten Simulationsmodellen auch die Nachverfolgung von offenen Punkten und Potentialen insgesamt erleichtert.

Dabei ist ein Transfer in andere Unternehmen und Industrien theoretisch gesehen einfach umsetzbar. In der Praxis sind jedoch einige grundlegende Punkte zu beachten. Neben der Fähigkeit in der Bedienung der Simulationssoftware sollte die kontinuierliche Verbesserungsarbeit nach Lean oder anderen Ansätzen bekannt und im Einsatz sein. Außerdem wird eine Organisationsstruktur benötigt, die Speedweeks vorbereiten, durchführen und abschließen kann. Dies könnte ebenso wie die Simulation jedoch auch als Dienstleistung erfolgen, wenn ein Unternehmen diese Strukturen nicht selbst leisten kann.

Literaturverzeichnis

- Alukal, G.; Manos, A.: *Lean Kaizen: a simplified approach to process improvements*. Milwaukee: ASQ Quality Press 2006.
- Baryl, C.; Gascon, V.; Miller, J.; Côté, N.: Use of a discrete-event simulation in a Kaizen event: A case study in healthcare. *European Journal of Operational Research* 249 (2016) 1, S. 327-339.
- Chiarini, A.: *From Total Quality Control to Lean Six Sigma: Evolution of the Most Important Management Systems for the Excellence*. Milan: Springer 2012.
- Coffey, D.; Thornley, C.: Automation, motivation and lean production reconsidered. *Assembly Automation*, 26 (2006) 2, S. 98-103.
- Dalton J.: *Gemba Walks*. In: *Great Big Agile*. Berkeley, CA: Apress 2019, S. 173-174.
- Gan, C. W.; Toh, M.H.; Lim, R.Y.; Ma, B.; Tan, P.S.; Bhullar, A.S.: Development of QuicKaizen™ technique for productivity execution management for Singapore SMEs. In: *2014 IEEE International Conference on Industrial Engineering and Engineering Management*, 2014, S. 448-452.
- García, J. L.; Maldonado, A.A.; Alvarado, A.; Rivera, D.G.: Human critical success factors for kaizen and its impacts in industrial performance. *The International Journal of Advanced Manufacturing Technology* 70 (2014) 9-12, S. 2187-2198.
- George, M.: *Lean six sigma: Combining six sigma quality with lean production speed*. New York: McGraw-Hill 2002.

- Goienetxea Uriarte, A.; Ng, A.H.; Urenda Moris, M.: Bringing together Lean and simulation: a comprehensive review. *International Journal of Production Research*, 58 (2020) 1, S. 87-117.
- Hines, P.; Holweg, M.; Rich, N.: Learning to evolve: a review of contemporary lean thinking. *International journal of Operations & Production Management* 24 (2004) 10, S. 994-1011.
- Hodge, G.L.; Goforth Ross, K.; Joines, J.A.; Thoney, K.: Adapting lean manufacturing principles to the textile industry. *Production Planning & Control*, 22 (2011) 3, S. 237-247.
- Hosono, A.; Page, J.; Shimada, G.: *Workers, managers, productivity: Kaizen in developing countries*. Singapore: Springer Nature 2020.
- King, P.L.: *Lean for the process industries: dealing with complexity*. New York: Productivity Press 2019.
- Lechler, T.; Zerreis, M.; Sitz, S.; Reinhard, M.; Franke, J.: (2020). Simulation hochautomatisierter Serienproduktion. *ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 115 (2020) 3, S. 123-126.
- Manos, A.: The benefits of Kaizen and Kaizen events. *Quality progress* 40 (2007) 2, S. 47.
- Radnor, Z. J.; Holweg, M.; Waring, J.: Lean in healthcare: the unfilled promise?. *Social science & medicine*, 74 (2012) 3, S. 364-371.
- Robinson, S.; Radnor, Z.J.; Burgess, N.; Worthington, C; SimLean: Utilising simulation in the implementation of lean in healthcare. *European Journal of Operational Research*, 219 (2012) 1, S. 188-197.
- Verein Deutscher Ingenieure: VDI 3633, Blatt 1: *Simulation von Logistik-, Materialfluss und Produktionssystemen – Grundlagen*. Düsseldorf: Beuth 2014.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Methoden und Konzepte des Digitalen Logistikzwillings der AG der Dillinger Hüttenwerke (Dillinger)

Application of a Digital Twin in Production Logistics at Dillinger (Rolling Mill)

Alexander Herzog, TU Clausthal, Clausthal-Zellerfeld (Germany),
alexander.herzog@tu-clausthal.de,

Jan Necil, Martin Pollet, Heike Busch, AG der Dillinger Hüttenwerke, Dillingen
(Germany), jan.necil@dillinger.biz, martin.pollet@dillinger.biz,
heike.busch@dillinger.biz

Abstract: For many optimisation problems in industrial environment, a digital twin is the appropriate tool. In this paper, the concrete challenges for planning on different time scales are described using the example of a rolling mill. Whereas on the long-term level questions arise concerning investment decisions, on the medium-term level it is a question of a product mix that is optimal in terms of yield, and on the short-term level it is a question of the optimal sequence of individual tasks. To be able to investigate these problems, various simulation models using different tools have been developed, focusing on different aspects. One of the simulation programs used is available as open source and is not limited to modelling processes in rolling mills. Therefore, the functionality of this program is discussed in detail.

1 Einführung

Die simulationsunterstützte Planung von Fabrikanlagen stellt ein in der Forschung gut durchdrungenes Gebiet dar, siehe z. B. Bracht et al. (2018). Angestrebt wird hier, die einzelnen Planungsprozesse von der Gebäudeplanung über das Produktionslayout, den Logistik- und Materialflussplänen bis hin zu der Planung der konkreten Fertigungsprozesse so gut wie möglich zu verzahnen. Dies bedeutet insbesondere, dass sich z. B. das Layout der Halle auf Basis der einzusetzenden Maschinen ergibt und die Auswahl der Maschinen wiederum auf Basis der zu fertigenden Produkte.

In vielen praktischen Anwendungsfällen besteht die Aufgabe jedoch nicht darin, eine vollständig neue Fabrik zu entwerfen, sondern Optimierungen der Produktion unter vorgegebenen Rahmenbedingungen durchzuführen. In den folgenden Überlegungen wird hierzu exemplarisch das Walzwerk von Dillinger betrachtet. Das

heute betriebene Walzwerk 2 am Standort Dillingen wurde 1971 gemäß der damaligen Nachfrage für die Produktion von großen Losen an Dickblechen und Pipeline-Blechen mit einer verhältnismäßig geringen Fertigungstiefe konzipiert. Seither ist die Fertigungstiefe stetig gestiegen, vgl. Busch (2015). So erhöhte sich die Fertigungstiefe in den vergangenen 20 Jahren um etwa 54 %. Auch die durchschnittliche Anzahl an Blechen pro Auftrag hat sich seit dem Jahr 2000 nahezu halbiert. 2020 umfassten ca. 35 % der Fertigungslose lediglich noch eine Größe von 1. Da ein vollständiger Neubau eines Walzwerkes, welches maßgeschneidert für diese neuen Anforderungen entworfen werden könnte, nicht nur aus finanziellen Erwägungen, sondern auch vor dem Hintergrund fehlender freier Flächen in der Umgebung keine Option darstellt, müssen die neuen Anforderungen so gut wie möglich in dem bestehenden Walzwerk abgebildet werden. Es können Investitionsentscheidungen zur Anschaffung bestimmter Maschinen getroffen werden, eine generelle Veränderung des Hallen-Layouts ist jedoch aufgrund der Größe und des Gewichts der zu transportierenden Bleche und der damit verbundenen Anforderungen an die Transportwege und -mittel ausgeschlossen.

Konkret werden in dem in dieser Arbeit beschriebenen Forschungsprojekt folgende Fragestellungen zur Optimierung von Durchsatz, Liefertreue und Ertrag untersucht:

- *Auftragsannahme:*
In den meisten Fertigungsprozessen ist jeder Auftrag mit einem spezifischen Ertrag und mit verschiedenen Kosten an unterschiedlichen Maschinen verbunden (Kilger 2013). Während der Ertrag leicht zu ermitteln ist, hängen die Kosten häufig von den Verarbeitungsschritten, die ein Werkstück durchläuft, ab. Werden mehrere Aufträge in dem Werk gleichzeitig bearbeitet, so können die Kosten, die an einer Maschine für einen bestimmten Auftrag entstehen, von dem gesamten Systemzustand abhängen. Erfolgt z. B. ein Erwärmvorgang in einem Ofen, so ist es von Vorteil, wenn möglichst viele gleichartige Werkstücke in direkter zeitlicher Folge bearbeitet werden, um zusätzliche Aufheiz- und Abkühlzeiten zu vermeiden. Die Frage, welche Aufträge zu welchen Preisen zu welchen Zeitpunkten anzunehmen sind, um einen maximalen Gewinn zu erzielen, kann daher nur auf Basis von dynamischen Simulationsmodellen untersucht werden (Klein und Steinhardt 2008).
- *Produktportfolio:*
Überführt man die Abwägung von Auftragsannahmen auf einen strategischen Zeithorizont, so ergibt sich die Frage nach dem zukünftig anzubietenden Produktportfolio. Während die operative Entscheidung über einzelne Aufträge die Fertigung als solches als unveränderlich annimmt, kann die Optimierung des zukünftigen Produktportfolios mit Investitionsentscheidungen gekoppelt werden.
- *Bestandsentwicklung:*
Die dynamische Simulation einer Fertigung ermöglicht es zu untersuchen, vor welchen Maschinen sich zu welchen Zeitpunkten welche Bestandsmengen ergeben werden. Bei der Analyse der Bestände muss domänenspezifisches Wissen zu den jeweiligen Werkstücken berücksichtigt werden: Während in einer Assembly-Line häufig nur Pufferrestriktionen als solches eingehalten werden müssen, führen z. B. bei der Verarbeitung von Dickblechen hohe Bestände dazu, dass die Bedienung nicht mehr in Ankunftsreihenfolge (First in first out, FIFO) erfolgen kann, sondern die wartenden Werkstücke einen Stapel bilden, auf den Werkstücke nur von oben aufgelegt und entnommen werden können (Last in first

out, LIFO). LIFO führt zu einer drastischen Steigerung des Variationskoeffizienten der Durchlaufzeiten (DLZ), vgl. Shortle et al. (2018), und damit zu einer schlechteren Vorhersagbarkeit der Liefertermine. Da im Fall des Walzwerks in Dillingen der Transport der Bleche zu einem Großteil über Hallenkräne, die jeweils für mehrere Stationen zuständig sind, erfolgt, führt ein hoher Bedarf an Kraninteraktionen an einer Station zu einer verringerten Kranverfügbarkeit an anderen Stationen.

2 Digitaler Logistik Zwilling des Walzwerks

Zur Analyse der genannten Problemstellungen wurde das Walzwerk von Dillinger in Form eines Warteschlangenmodells abgebildet und dieses Modell in zum Teil speziell für die Fragestellung entwickelten ereignisorientierten stochastischen Simulationswerkzeugen simuliert. Das Walzwerk umfasst ca. 150 Bedienstationen, an denen ca. 600 verschiedene Fertigungsschritte ausgeführt werden können. Über die jeweils auszuführenden Fertigungsschritte entscheiden dabei neben der Größe des Blechs vor allem die gewünschten Materialeigenschaften. Wenn mehrere unterschiedliche Anlagentypen für einen Fertigungsschritt genutzt werden können, müssen die jeweiligen Kosten sowie die unterschiedlichen Bearbeitungsauern berücksichtigt werden. Eines der Ziele der Produktionsplanung besteht dabei darin, Grobbleche mit ähnlichen Anforderungen zusammen zu führen, um so eine bessere Auslastung der Anlagen zu erreichen, sowie um Zeitverluste durch Rüstzeiten der Anlagen zu vermeiden.

Zur Analyse der verschiedenen Fragestellungen wurde auf zwei Modelle, die durch verschiedenartige Reduktionen und Idealisierungen der realen Fertigung entstanden sind, zurückgegriffen:

1. In dem ersten Modell wurden 22 Stationen, die unter Berücksichtigung der Hauptmaterialströme und auch auf Basis der Relevanz der jeweiligen Station aus warteschlangentheoretischer Sicht ausgewählt wurden, abgebildet. Eine Modellstation muss dabei nicht mit einer physischen Station übereinstimmen: Verzweigungen der Materialströme können im Modell durch Stationen abgebildet werden, auch wenn hierfür in der Realität kein Aggregat notwendig ist. Umgekehrt können reale Lagerflächen und Maschinen zu einer Modellstation zusammengefasst werden. Während die Anzahl der Stationen sehr gering gewählt war, erfolgte die Ausgestaltung der Stationen insbesondere in Bezug auf den verfügbaren Stauraum sehr detailliert: Die Warteschlangenkapazität vor einer Station kann in der Realität entweder durch die Anzahl an Blechen (z. B. bei der Stapelung), durch die Fläche (z. B. auf Abkühlflächen) oder durch das Volumen der Bleche (z. B. bei Bunkerprozessen) erfolgen. Soll ein Blech zu einer Station transferiert werden, so kann geprüft werden, ob an der Zielstation in Bezug auf die jeweils relevante Größe genügend Kapazität vorhanden ist. Auf diese Weise können sehr detaillierte Analysen zu Rückstauvorgängen usw. durchgeführt werden.
2. Das zweite Modell besteht aus insgesamt 128 Modellstationen, die ebenfalls nicht deckungsgleich mit einer physischen Station müssen. Im Gegensatz zum ersten Modell wird hier nicht das gesamte Walzwerk betrachtet, sondern nur die Prozesse der Endbearbeitung. D. h. in Bezug auf die Abbildung der Stationen ist dieses Modell wesentlich detaillierter. Dafür verfügen die Stationen über weniger

Detailinformationen zu den Blechen. Ziel dieses Modelles ist es, Stoffströme in der Fertigung detailliert und zugleich auf eine sehr zugängliche Art analysieren zu können. Der abgebildete Fertigungsplan enthält 35 mögliche Prozessschritte. Die Visualisierung der Fertigung ermöglicht es, diese Schritte auf Konsistenz in Bezug auf die realen Prozesse hin zu überprüfen.

Für die konkrete Umsetzung der beiden Modelle, die gemeinsam den Digitalen Logistik Zwilling bilden, wurde auf zwei Programme, die beide auf derselben Simulationsklassenbibliothek aufbauen, zurückgegriffen. Die Bibliothek ist dabei insbesondere auf einen hohen Durchsatz an Ereignissen ausgelegt. Auf die Funktionsweise ereignisorientierter Simulation (Discrete-event simulation, DES) soll an dieser Stelle nicht eingegangen werden. Einführungen zu dem Thema finden sich z. B. in Law und Kelton (2014), Gutenschwager et al. (2017) oder auch Waldmann und Helm (2016). Auf die für die Implementierung relevanten mathematischen Details wird u. a. in Kolonko (2008) eingegangen. An dieser Stelle sollen jedoch einige für die konkreten Fragestellungen relevanten Spezifika betrachtet werden:

2.1 Flexible Modellierung in Bezug auf die Eingabeparameter

Die Zwischenankunftszeiten am System, die Bedienzeiten usw. können in DES-Systemen prinzipiell auf zwei Arten modelliert werden: Entweder in Form von historischen Daten bzw. bekannten zukünftigen Daten über bereits angenommene Aufträge (und den daraus jeweils abgeleiteten Objektattributen der Aufträge bzw. Bleche) oder in Form von Wahrscheinlichkeitsverteilungen. Während es die Nutzung von Verteilungen ermöglicht, dass das Modell experimentierfähig ist, vgl. auch Sauerbier (1999), ermöglicht die Verwendung von konkreten, bekannten Daten die exakte Voraussage der Bestände usw. auf einem kürzeren zeitlichen Horizont.

Die verwendeten Simulationswerkzeuge ermöglichen eine flexible Mischung dieser beiden Varianten, so dass für die Vorhersage von DLZ, Beständen, Lieferterminen usw. zum einen bekannte Aufträge und zum anderen über Wahrscheinlichkeitsverteilungen geschätzte Werte in der rollierenden Planung verwendet werden können. Während für die strategische Planung meist vollständig auf Verteilungsfunktionen zurückgegriffen wird, werden für die operative Planung vorrangig bekannte, konkrete Werte verwendet.

2.2 Anbindung an Data-Warehouse-Systeme

Da IoT-Sensoren mittlerweile vielfach drahtlos arbeiten, ist auch eine Nachrüstung bestehender Fertigungsanlagen, sogenanntes Retro-Fitting, häufig mit geringem Aufwand möglich, vgl. auch Butun (2020) und Kumar et al. (2020). Auf dem Werksgelände in Dillingen sind über 200.000 IoT-Sensoren und -Aktoren im Einsatz, so dass eine sehr große Menge an Rohdaten für die Erzeugung von Simulationsmodellen zur Verfügung steht. Für eine langfristig stabile Nutzung dieser Daten kommt der Art der Anbindung eine besondere Rolle zu. Da Fertigungsprozesse Veränderungen von technischer und von kaufmännischer Seite unterworfen sind, ist eine direkte statische Anbindung an bestehende Datenbanken für ein übergreifend nutzbares Werkzeug nicht zielführend. Daher wurde eine Abstraktionsschicht in Form eines Modellgenerators entworfen. In diesem sind die Namen der Datenbanken und der relevanten Spalten hinterlegt. Auf dieser Basis

können über die Java Database Connectivity (JDBC) Abfragen durchgeführt werden und die Ergebnisse über das Jakarta XML Binding (JAXB) in den Simulator übernommen werden. Damit sind die Simulationswerkzeuge als solches bestmöglich von der technischen Implementierung der Datenbanksysteme abstrahiert.

2.3 Erfassung und Analyse der Kenngrößen

Während einer Simulation werden automatisch an allen Stationen nicht nur Durchschnittswerte und Streuungen, sondern auch zeitliche Verläufe erfasst, so dass eine stationäre und auch eine transiente Analyse möglich sind. Im Anschluss an eine Simulation stehen die Daten über eine grafische Oberfläche (als Texte, Tabellen und Grafiken) und auch zur maschinellen Weiterverarbeitung in Form von XML-Daten zur Verfügung. Es hat sich gezeigt, dass dieser fließende Übergang zwischen beiden Varianten für viele Nutzungsszenarien einen großen Vorteil darstellt.

3 Warteschlangensimulator

Für die im Folgenden beschriebenen Untersuchungen wurden zwei Werkzeuge, die beide auf derselben Plattform aufbauen, sich aber in der konkreten Ausgestaltung unterscheiden, verwendet.

Für die detaillierte Modellierung der Abläufe an den einzelnen Stationen (wenige Stationen, dafür mit hohem und stark domänenspezifischem Modellierungsgrad, siehe Modell 1 in Abschnitt 2) wurde ein Werkzeug entwickelt, welches exklusiv Dillinger zur Verfügung steht. Dieses besitzt, abgesehen von der Statistikanzeige, nur eine rudimentäre grafische Oberfläche. Die Stärke dieses Werkzeuges liegt in der unmittelbaren Anbindung an die Daten des Walzwerks in Dillingen. Auf dieses Werkzeug soll daher im Folgenden nicht im Detail eingegangen werden. Bei dem zweiten Programm (mit dem Modell 2 in Abschnitt 2 simuliert wurde) handelt es sich um einen als OpenSource verfügbaren ereignisorientierten stochastischen Simulator. Der Ebeneneinteilung von Gutenschwager et al. (2017) folgend handelt es sich bei dem obigen Dillinger-spezifischen Simulator um ein Werkzeug der Ebene 4 („Spezialsimulationswerkzeug, das spezifische Komponenten für ein abgegrenztes Teilgebiet eines Anwendungsbereiches enthält“). Der im Folgenden vorgestellte Warteschlangensimulator ist jedoch ein Programm der Ebene 2 („Simulationswerkzeug mit bereits implementierten allgemeinen Komponenten für beliebige Anwendungen“).

Vergleichbar mit dem Warteschlangensimulator sind z. B. AnyLogic, siehe Borshchev (2013), Rockwell Arena, siehe Kelton (2002), und ExtendSim, siehe Strickland (2013). Das Besondere dabei ist, dass es sich bei den genannten Programmen durchgängig um kostenpflichtige Closedsource-Programme handelt. Betrachtet man Swain (2017) oder Dagkakis und Heavy (2016), so stellt man fest, dass im OpenSource-Bereich fast nur Bibliotheken, auf deren Basis der Anwender selbst Simulatoren implementieren muss, zur Verfügung stehen. Der Warteschlangensimulator schließt diese Lücke: Das Programm ist kostenlos verfügbar, stellt jedoch ein direkt einsetzbares Werkzeug dar.

3.1 Programmüberblick

Eine vollständige Beschreibung der Funktionsweise des Warteschlangensimulators inkl. Modellierungsbeispielen aus Industrie und Logistik findet sich in Herzog (2021). An dieser Stelle soll nur ein kurzer Überblick über das Programm gegeben werden. Der Warteschlangensimulator ist ein plattformunabhängiges Java-Programm. Den Hauptbereich des Fensters füllt die Zeichenfläche. Auf dieser können einzelne Stationen abgelegt und verknüpft werden. Die Funktionsweise orientiert sich an den zuvor genannten Simulationswerkzeugen bzw. an vektoriiellen Zeichenprogrammen. Der Modellierungsumfang, d. h. die angebotenen Stationstypen, umfasst neben typischen Stationen wie Quellen, Bedienstationen und Senken auch Stationen für Wertzuweisungen, zur bedingungsgesteuerten Verzögerung von Kunden, zur Verzweigung, zur Batch-Bildung und -Auflösung, zum Transport und zur Daten-Ein- und -Ausgabe. Auch können in einem begrenzten Umfang zeitkontinuierliche Werte abgebildet werden. In Summe stehen mehr als 100 verschiedene Stationstypen zur Verfügung. Bei Bedienstationen können z. B. neben der jeweils notwendigen Bedienzeit auch eine Rüst- und eine Abrüstzeit angegeben werden, es kann konfiguriert werden, wie viele Bediener welcher Typen für eine Bedienung benötigt werden, welche Priorität die Station in Bezug auf eine Bedienerzuordnung erhalten soll und welche Prioritäten die einzelnen wartenden Kunden innerhalb der Station erhalten sollen.

3.2 Modellparameter

Zeitdauern können über Wahrscheinlichkeitsverteilungen oder über Rechenausdrücke definiert werden. Der Warteschlangensimulator verfügt über 30 Wahrscheinlichkeitsverteilungstypen. In Rechenausdrücken stehen neben mathematischen Funktionen insbesondere alle Systemparameter zur Verfügung. So kann die Wartezeit in nicht-linearer Weise in die Priorität eines Kunden einfließen oder aber die Warteschlangenlänge kann sich auf die Bedienzeit auswirken. Zur Ermittlung von passenden Wahrscheinlichkeitsverteilungen auf Basis von Messwerten steht ein integriertes Input-Analyse-Werkzeug zur Verfügung, welches einen Abgleich der Messwerte mit insgesamt 22 verschiedenen Verteilungen vornehmen kann und die jeweils am besten passende vorschlägt. Neben der quadratischen Abweichung werden hierfür der Kolmogorov-Smirnov-, der χ^2 - und der Anderson-Darling-Anpassungstest zu Rate gezogen.

Ankünfte am System oder auch die jeweils verfügbare Anzahl an Bedienern (Schichtpläne) können optional aus externen Dateien oder auch aus Datenbanken geladen werden. Über eine konfigurierbare JDBC-Schnittstelle ist ein prinzipieller Zugriff auf beliebige Datenbanksysteme möglich.

Des Weiteren kann über Skript-Stationen direkt im Simulator nutzerdefinierter Javascript- oder Java-Code zu hinterlegen werden, der jeweils ausgeführt wird, wenn ein Kunde die entsprechende Station durchquert.

3.3 Statistikausgabe

Während einer Simulation werden sämtliche relevanten Kenngrößen automatisch erfasst, so dass keine a-priori Planung der Erfassung erfolgen muss. Im Anschluss an eine Simulation werden die Ergebnisse über die integrierte Statistikansicht angeboten. Die verfügbaren Texte, Tabellen und Grafiken können direkt über die

Zwischenablage exportiert werden und vom Simulator in Microsoft Office- und LibreOffice-Formaten gespeichert werden. Außerdem steht eine Reportgeneratorfunktion zur Verfügung, die ausgewählte Daten als docx-, pdf-, tex- oder html-Report zusammenfassen kann. Zur schnelleren Zuordnung, welche Daten sich auf welche Station beziehen, werden nach einer Simulation die zu den jeweiligen Stationen gehörenden Daten in Tooltips angezeigt und es können Kenngrößen als Heatmaps auf der Zeichenfläche visualisiert werden.

3.4 Parameterreihen und Optimierung

Mit der Parameterreihen-Funktion können basierend auf einem Ausgangsmodell ein oder mehrere Parameter in einem vorgegebenen Bereich variiert werden und es können jeweils bestimmte Ausgabeparameter ausgegeben werden. Auf diese Weise ist es mit minimalem Aufwand möglich, ein Verständnis über verschiedene Zusammenhänge zu gewinnen. Neben der Erhebung von Daten in einem bestimmten Parameterbereich steht auch eine Optimierungsfunktion zur Verfügung, die versucht mit so wenig wie möglichen Simulationsläufen einen bestimmten Zielwert für eine bestimmte Kenngrößen optional unter Einhaltung von Nebenbedingungen zu erreichen. Für die Optimierung stehen neben klassischen Hill-Climbing-Algorithmen (unter Verwendung von Tabu-Suche) auch Genetische Algorithmen zur Verfügung (Papageorgiou 2015).

3.5 Animationen

In der auf die Einbindung von Simulationssystemen in industrielle Prozesse ausgerichteten Literatur, z. B. Wenzel et al. (2007), Arnold und Furmans (2005) und Sauerbier (1999), wird wiederholt darauf hingewiesen, dass Simulationsergebnisse nicht nur korrekt sein müssen, sondern auch akzeptiert werden müssen. Die Animation von Modellen trägt in vielen Fällen erheblich zur Akzeptanz bei denjenigen, die die Ergebnisse später in der realen Fertigung umsetzen sollen, bei. Darüber hinaus stellt die Animation von Modellen ein leistungsstarkes Hilfsmittel zur Fehlersuche in Modellen dar. Im Warteschlangensimulator können Modelle animiert werden, ohne dass dafür eine gesonderte Konfiguration nötig wäre. Optional können dabei Stationen, Kunden-, Bediener- und Transportertypen individuelle Symbole zugewiesen werden. Animationen können jeder Zeit unterbrochen und im Einzelschrittmodus ausgeführt werden. Wird eine Animation im Einzelschrittmodus ausgeführt, so werden die jeweils ausgeführten Ereignisse angezeigt und Systemparameter können direkt eingesehen und verändert werden.

3.6 Automatisierung und Einbindung in IT-Systeme

Der Warteschlangensimulator ist gleichermaßen für die interaktive Nutzung und auch für den Betrieb eingebettet in eine größere IT-Landschaft ausgelegt. Aufgrund der offenen und dokumentierten XML-Dateiformate für Modelle, Statistikdaten usw. sind eine externe Modellgenerierung und Weiterverarbeitung der Ergebnisdaten einfach möglich. Ein Kommandozeilen-Betrieb des Simulators ist ebenso möglich wie eine Fernsteuerung über REST-, MQTT- und Socket-Schnittstellen oder per Webinterface.

3.7 Verfügbarkeit

Der Warteschlangensimulator steht als ausführbare Binärdatei als auch im Quellcode unter

<https://github.com/A-Herzog/Warteschlangensimulator>

zum Download zur Verfügung. Das Programm wird unter der Apache 2.0 Opensource-Lizenz zur Verfügung gestellt und ist damit uneingeschränkt auch im kommerziellen Kontext nutzbar. In dem Installationspaket befinden sich neben dem Programm auch sämtliche Online-Hilfen, pdf-basierte Referenzen und auch Tutorials. Die gesamte Dokumentation und auch die Programmoberfläche selbst stehen dabei auf Deutsch und auf Englisch zur Verfügung.

4 Untersuchungsergebnisse

Die beschriebenen Simulationswerkzeuge werden in der rollierenden Produktionsplanung und zur Analyse strategischer Fragestellungen eingesetzt:

- *Analyse der Produktionsprozesse / Investitionsentscheidungen:*
Die warteschlangentheoretische Abbildung der Produktion im Walzwerk von Dillinger gewährt neue Einblicke in die Prozesse und führt so zu einem besseren Verständnis der Zusammenhänge und Abhängigkeiten. Konkret wird auf Simulationsbasis untersucht, welche Auswirkungen eine Veränderung der Bedienkapazität an bestimmten Anlagen zur Folge hat, um so Empfehlungen für mögliche Investitionsentscheidungen abgeben zu können. Diese Ergebnisse fließen bereits in die taktische und strategische Ausrichtung der Produktion ein.
- *Retrospektive Analyse:*
Die retrospektive Analyse der Produktion mit einem mit historischen Daten parametrisierten Simulationsmodells erlaubt es, Engpässe oder Konstellationen in der Belegung von Anlagen zu identifizieren, die z. B. Liefertermine beeinflussen. So wurden bereits umsetzbare Handlungsempfehlungen abgeleitet, um den Durchlauf bestimmter zeitkritischer Aufträge durch die Produktion zu beschleunigen.
- *Revenue Management:*
In klassischen Verfahren zur Ertragsoptimierung des Produktmixes wird eine Fertigung in Form eines linearen Programms (LP) abgebildet und im Folgenden mit einem LP-Solver analysiert (Mellouli und Suhl 2006). Bei dieser Vorgehensweise geht jedoch die zeitliche Dimension bzw. die Abhängigkeit der Produktionskosten für ein bestimmtes Los vom aktuellen Systemzustand verloren. Mit Hilfe eines Simulationsmodells können solche Abhängigkeiten, die im vorliegenden Fall zu erheblichen Kostenunterschieden führen können, abgebildet werden.
- *Kapazitätsplanung:*
Für die kurz- bis mittelfristige Planung der Auftragsannahme und der Reihenfolgeplanung kann die Auslastung des gesamten Werkes sowie einzelner Stationen analysiert werden. Hierbei können konkrete, vorliegende Aufträge mit Schätzungen für die weitere Zukunft kombiniert werden und so Wahrscheinlichkeitskorridore bestimmt werden, siehe auch Abbildung 1. Hierauf kann heute zeitnah z. B. mit Anpassungen von Schichten und Arbeitsplänen reagiert werden.

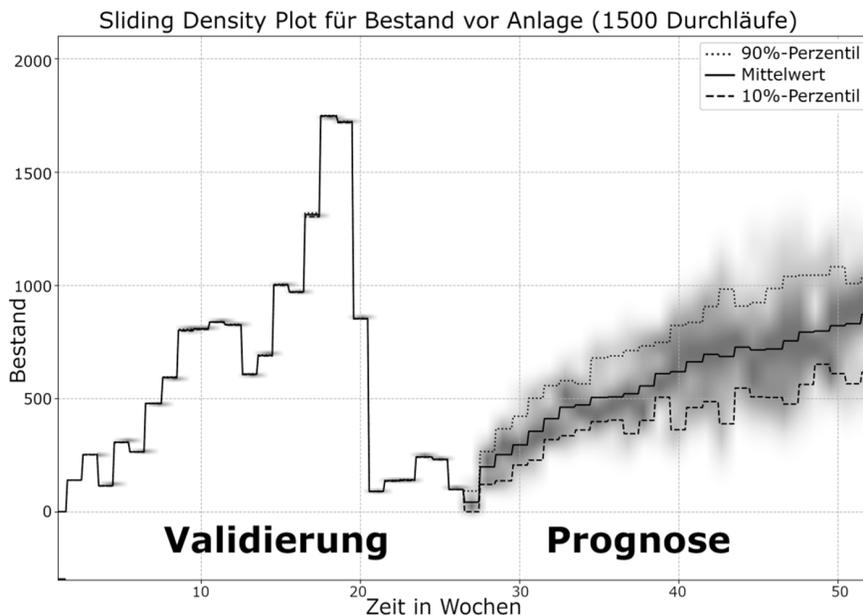


Abbildung 1: Simulierte Entwicklung der Bestände in der Produktion

5 Ausblick

Die beschriebenen Simulationswerkzeuge befinden sich bei Dillinger im Regelbetrieb und werden momentan bereits zur operativen Planung von Produktion und Auftragsannahme und zur Bewertung von strategischen Entscheidungen verwendet.

Da die Simulationsmodelle technische Details abstrahieren, sollen diese perspektivisch nicht nur in den technischen Fachabteilungen zum Einsatz kommen, sondern eine Brücke zwischen Marketing und Vertrieb auf der einen Seite und technischem Betrieb und strategischen Produktions- und Anlagenplanung auf der anderen Seite bilden. So ist es für den Vertrieb von hoher Bedeutung, auf möglichst zugängliche Weise und ohne selbst über umfangreiches Wissen zu Materialeigenschaften und Fertigungsschritten zu verfügen, den mit einem Auftrag verbundenen Ertrag abschätzen zu können. Außerdem beschränkt sich der Einsatz der Werkzeuge bislang auf den Bereich des Walzwerks. Das integrierte Hüttenwerk deckt jedoch die gesamte Produktionskette von der Koks- und Roheisenproduktion über die Stahlerzeugung bis hin zur Blechherstellung ab. Das nächste Ziel besteht daher darin, diese vertikale Integration auch in einem Simulationswerkzeug abzubilden, um so eine übergreifende Analyse und Optimierung der Prozesse zu ermöglichen.

Literatur

- Arnold, D.; Furmans, K.: Materialfluss in Logistiksystemen. 6. Auflage. Berlin, Heidelberg, New York: Springer 2005.
- Borshchev, A.: The big book of simulation modeling: multimethod modeling with anylogic 6. Chicago: AnyLogic North America 2013.
- Bracht, U.; Geckler D.; Wenzel, S.: Digitale Fabrik: Methoden und Praxisbeispiele. Berlin, Heidelberg: Springer 2018.
- Busch, H.: Strongly diversifying product mix in heavy plate rolling - A challenge in logistics. Stahl und Eisen, 135. S. 73-79, Verlag Stahleisen GmbH 2015.
- Butun, I.: Industrial IoT - Challenges, Design Principles, Applications, and Security. Springer International Publishing 2020.
- Dagkakis, G.; Heavey, C.: A review of open source discrete event simulation software for operations research. Journal of Simulation 10 (2016) 3, S. 193-206.
- Gutenschwager, K.; Rabe, M.; Spieckermann, S.; Wenzel, S.: Simulation in Produktion und Logistik. Berlin, Heidelberg: Springer 2017.
- Herzog, A. Simulation mit dem Warteschlangensimulator. Angenommen zur Veröffentlichung 2021 von Springer.
- Kelton, W. D.: Simulation with arena. New York: McGraw-Hill 2002.
- Kilger, W.: Optimale Produktions- und Absatzplanung: Entscheidungsmodelle für den Produktions- und Absatzbereich industrieller Betriebe. VS Verlag für Sozialwissenschaften 2013.
- Klein, R.; Steinhardt, C.: Revenue management: Grundlagen und mathematische Methoden. Berlin, Heidelberg: Springer 2008.
- Kolonko, M.: Stochastische Simulation – Grundlagen, Algorithmen und Anwendungen. Wiesbaden: Vieweg+Teubner 2008.
- Kumar, P.; Ponnusamy, V.; Jain, V.: Industrial Internet of Things and Cyber-Physical Systems: Transforming the Conventional to Digital. IGI Global 2020.
- Law, A.; Kelton, W D.: Simulation modeling and analysis. 5. Auflage. New York: McGraw-Hill 2014.
- Mellouli, T.; Suhl, L.: Optimierungssysteme: Modelle, Verfahren, Software, Anwendungen. Berlin, Heidelberg: Springer 2006.
- Papageorgiou, M.; Leibold, M.; Buss, M.: Optimierung: Statische, dynamische, stochastische Verfahren für die Anwendung. Berlin, Heidelberg: Springer 2015.
- Sauerbier, T.: Theorie und Praxis von Simulationssystemen. Braunschweig, Wiesbaden: Vieweg & Sohn 1999.
- Shortle, J. F.; Thompson, J. M.; Gross, D.; Harris, C. M.: Fundamentals of queueing theory. John Wiley & Sons 2018.
- Strickland, J.: Discrete event simulation using extensim 8. Morrisville: Lulu.com 2013.
- Swain, J. J.: Simulation software survey-simulation: new and improved reality show. OR/MS Today 44 (2017) 5, S. 38-49.
- Waldmann, K.-H.; Helm, W. E.: Simulation stochastischer Systeme. Berlin, Heidelberg: Springer 2016.
- Wenzel, S.; Weiß, M.; Collisi-Böhmer, S.; Pitsch, H.; Rose, O.: Qualitätskriterien für die Simulation in Produktion und Logistik: Planung und Durchführung von Simulationsstudien. Berlin, Heidelberg: Springer 2007.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Einsatz ereignisdiskreter Simulation zur simultanen Allokation von Bestands-, Kapazitäts- und Zeitpuffern in variabilitätsbeeinflussten Produktionssystemen

***Discrete-event simulation for simultaneous allocation of inventory,
capacity and time buffers in variability-influenced production systems***

David Lenze, TU Dortmund, Dortmund (Germany),
david.lenze@ips.tu-dortmund.de

Julian Schallow, Institut für Forschung und Transfer (RIF) e. V., Dortmund
(Germany), julian.schallow@rif-ev.de

Jochen Deuse, TU Dortmund, Dortmund (Germany) und University of Technology,
Sydney (Australia), jochen.deuse@ips.tu-dortmund.de

Abstract: In today's production systems, one of the main requirements is to manage the prevailing variability effects. The targeted use of buffers is a suitable approach for this purpose, whereby the combination of inventory, capacity and time buffers is most effective. This paper therefore describes an integrated approach to the simultaneous allocation of inventory, capacity and time buffers in order to achieve an optimised control of the variability-related influences in production systems. This framework combines the simulation-based evaluation of buffer allocations with statistical experimental design approaches. For a higher degree of practicability of the approach, the evaluative method module of the discrete-event simulation contains a (semi-) automatic simulation model generation. The description of a prototypical industrial application of the methodology underlines the suitability of the proposed approach.

1 Motivation

Die zunehmende Individualisierung der Märkte und eine stetig wachsende Erwartungshaltung der Kunden hinsichtlich Lieferfähigkeit und verkürzter Lieferzyklen stellen produzierende Unternehmen vor große Herausforderungen. Die aus diesen Anforderungen entstehenden Variabilitätseinflüsse sind eine Hauptursache für das Auftreten von Verschwendung innerhalb der Produktion, die in dieser Sichtweise als Symptom auftritt und nur über eine Reduzierung der Ursachen beeinflusst werden kann (Hopp und Spearman 2004). Da sich ein erhöhtes

Maß an Variabilität immer nachteilig auf die Leistungsfähigkeit eines Produktionssystems auswirkt, stellt Variabilitätsmanagement einen wesentlichen Aspekt der Produktionsforschung dar (Hopp und Spearman 2009). Für einen zielgerichteten Umgang mit der in der Produktion auftretenden Variabilität ist eine Unterscheidung zwischen verschiedenen Variabilitätsklassen erforderlich (Lorentzen et al. 2011). Während Leistungs-, Verfügbarkeits- und Qualitätsverluste gemeinhin als nicht-wertschöpfend gelten und eliminiert werden sollten, sind bspw. Wettereinflüsse oder Krankenstand zwar nicht-wertschöpfend, können jedoch in weiten Teilen nicht eliminiert werden, weshalb sie in ihren Auswirkungen beherrscht werden müssen (Wiegand et al. 2014). Darüber hinaus besitzt wertschöpfende Variabilität in Form individueller Kundenwünsche oder Liefertermine eine gesteigerte Bedeutung. Diese sollte nicht eliminiert werden, da ihre Existenz eine unmittelbare Steigerung des Kundennutzens bewirkt und daher mittelbar eine unternehmerische Wertschöpfung generiert (Maschek et al. 2014). Bedingt durch diese nicht zu eliminierenden Variabilitätseinflüsse besteht ein Bedarf an Ansätzen zur Beherrschung ihrer Auswirkungen, wobei der Einsatz von Puffern hinsichtlich dieser Zielstellung als besonders effizient zu bewerten ist (Palominos et al. 2009; Deuse et al. 2018).

2 Bestehende Ansätze zum Einsatz von Puffern

Grundsätzlich ist innerhalb von Produktionssystemen jede Form von Asynchronität zwischen der Nachfrage und dem internen Transformationsprozess ein Auslöser für die Entstehung von Puffern (Hopp und Spearman 2009). Diese verfolgen das Ziel, die negativen Auswirkungen der systemimmanenten Variabilität auf die Produktionsleistung abzufangen (Pound et al. 2014). In dieser allgemeinen Definition wird der Nutzen von Puffern stark betont, die Pufferausprägung hinsichtlich Art, Größe und Position bleibt jedoch offen. In der Produktion und Logistik ist eine nahezu ausschließliche Betrachtung von Bestandspuffern wahrzunehmen. Demgegenüber definieren Hopp und Spearman (2009) neben Bestands- ebenfalls Kapazitäts- sowie Zeitpuffer und weisen auf ihre Schutzwirkung für Produktionssysteme hin. Diese Sichtweise basiert auf der einfachen Annahme, dass in variabilitätsbehafteten Systemen immer Wartezeiten für Produkte (Bestandspuffer), Kunden (Zeitpuffer) oder Ressourcen (Kapazitätspuffer) entstehen (Pound et al. 2014). Die drei Pufferarten sind dabei in ihrer Wirkung auf Produktionssysteme untereinander nahezu vollständig substituierbar, weshalb eine integrierte Betrachtung der drei Pufferarten im Hinblick auf ihre Wirkung als zielführend zu bewerten ist. Somit resultiert hinsichtlich der Puffergestaltung die Fragestellung, wo, in welcher Größe und welche Kombination der drei Pufferarten die Leistung eines Produktionssystems besonders positiv beeinflussen.

Die bisherigen Entwicklungen im Bereich der Optimierungsansätze für Puffer sind unter dem Begriff des Pufferallokationsproblems zusammengefasst. Abstrahiert beschrieben ist das Pufferallokationsproblem ein kombinatorisches Optimierungsproblem zur Ermittlung von Pufferkonfigurationen, durch deren Implementierung innerhalb eines Produktionssystems eine vorab definierte Zielfunktion optimiert wird (Almomani 2012; Narasimhamu et al. 2014; Costa et al. 2015). Durch seine Klassifikation als stochastisches, i. d. R. nichtlineares, ganzzahliges, NP-schweres Optimierungsproblem ist das Pufferallokationsproblem

ein komplexes Problem der Produktionssystemgestaltung (Papadopoulos und Vidalis 1998; Chaharsooghi und Nahavandi 2003; Alon et al. 2005). Die Lösung derartiger Probleme erfordert insbesondere für komplexe Produktionssysteme Ansätze, deren Struktur aus zwei ineinandergreifenden Methoden, einer generativen und einer evaluativen Methode besteht. (Spinellis und Papadopoulos 2000; Demir et al. 2012). Dabei existieren vielzählige Ausprägungen der beiden Methodenbestandteile, die individuelle Vor- und Nachteile aufweisen. Auf Seiten der evaluativen Methoden besitzen exakte und analytische Lösungsverfahren deutliche Einschränkungen bei der Abbildung komplexer Lösungsszenarien, da diese in vielen Fällen nicht mit vertretbarem Aufwand modelliert werden können (Colledani et al. 2010). Zwar ermöglichen Näherungsverfahren die Abbildung eben dieser Szenarien, jedoch weisen sie deutliche Einschränkungen hinsichtlich der Abbildung dynamischer Effekte innerhalb der Produktion ab. Da diese aus den im Vorkapitel dargestellten Gründen eine wesentliche Rolle für eine praxistaugliche Anwendung spielen, sind die analytischen Evaluationsmethoden in der Praxis deutlich weniger verbreitet als die präferiert eingesetzten und in der Anwendung aufgrund ihrer guten Abbildungsfähigkeiten für tauglich befundenen Simulationsansätze. Unter den generativen Methoden weisen Suchverfahren entsprechend der analytischen Evaluationsverfahren nachteilige Eigenschaften hinsichtlich der Abbildbarkeit der Realität auf. Zwar beheben Metaheuristiken diese Nachteile speziell in hybriden Ansätzen, jedoch sind sie aufgrund der erforderlichen Parametrisierung zur Anpassung auf das jeweilige Pufferproblem schwieriger zu handhaben und erfordern somit Kompetenzen, die nicht jedem Anwender in der Praxis zur Verfügung stehen (MacGregor Smith und Cruz 2005). Standardisierter einsetzbar und aufgrund der zahlreichen Strategien zur Versuchsplanentwicklung verbessert auf das Problem anpassbar sind Ansätze der statistischen Versuchsplanung. Die gute Kombinierbarkeit mit Simulationsansätzen in der Evaluation macht eine Kombination der ereignisdiskreten Simulation und der statistischen Versuchsplanung zielführend. Bei einer Detailbetrachtung der in allen Klassen verfügbaren Ansätze ist deutlich festzustellen, dass der Betrachtungsfokus zur Lösung des Pufferallokationsproblems nahezu ausschließlich auf der Allokation von Bestandspuffern liegt. Eine simultane Betrachtung und Allokation von Bestands-, Kapazitäts- und Zeitpuffern bleibt hingegen bisher aus (Demir et al. 2014). Die drei Pufferkomponenten Kapazität, Bestand und Zeit wirken jedoch wie bereits beschrieben wechselseitig aufeinander ein, weshalb für eine optimierte Beherrschung der in industriellen Produktionssystemen vorherrschenden Variabilitätseinflüsse eine geeignete Kombination dieser einzusetzen ist. Die Bestimmung des optimalen, pufferbezogenen Betriebspunkts eines Produktionssystems erfordert somit eine simultane Berücksichtigung aller drei Pufferarten. Auf Basis einer Betrachtung der Gewichtung verfahrensspezifischer Vor- und Nachteile im Anwendungsumfeld (geringere Bedeutung der Lösungsgeschwindigkeit, mittlere Bedeutung der Lösungsgenauigkeit, hohe Bedeutung der Abbildungstreue, hohe Bedeutung der Einfachheit des Verfahrens) ist die Simulation als Evaluationsverfahren als gut geeignet zu bewerten. Aufgrund ihrer Fähigkeit auch *komplexe Produktionssysteme realitätsgetreu abbilden* zu können, ist sie in vielfältigen Praxisanwendungen einsetzbar (Can et al. 2008). Darüber hinaus erzeugt sie eine hohe Transparenz, was zu guter Akzeptanz in der praktischen Anwendung und einer einfachen Interpretierbarkeit der Ergebnisse führt. Die Schwachstelle einer geringen Lösungsgeschwindigkeit besitzt im Hinblick auf

die Problemstellung der Pufferallokation eine geringere Bedeutung. Die Pufferallokation ist zwar in regelmäßigen Zeitabständen zu wiederholen, bspw. um auf signifikante Verschiebungen im Produktionsprogramm zu reagieren, eine betriebsbegleitende Entscheidungsunterstützung nahe Echtzeit ist jedoch nicht erforderlich. Ansätze unter Verwendung der statistischen Versuchsplanung im generativen Teil weisen die geringsten Anwendungsvoraussetzungen auf und sind daher zur Puffergestaltung in der industriellen Praxis am besten geeignet. Darüber hinaus sind die experimentell ermittelten Ergebnisse einfach interpretierbar und ihre Anwendungskomplexität ist, verglichen mit Metaheuristiken oder Suchverfahren, deutlich geringer. Schwachstellen hinsichtlich des gering gewichteten Kriteriums der Lösungsgeschwindigkeit in der häufigen Kombination mit dem evaluativen Verfahren der Simulation und deren Laufzeiten können durch die Vermeidung einer vollständigen Lösungsraumuntersuchung hinreichend reduziert werden (Siebertz et al. 2017). Aufgrund ihrer spezifischen Vorteile ist die Anwendung von Ansätzen der statistischen Versuchsplanung zur Puffergestaltung in komplexen, industriellen Produktionssystemen somit zusammenfassend als zielführend zu bewerten.

3 Methodik zur simultanen Allokation von Bestands-, Kapazitäts- und Zeitpuffern

Vor dem Hintergrund des aufgezeigten Bedarfs einer simultanen Bestands-, Kapazitäts- und Zeitpufferallokation wurde die in Abbildung 1 dargestellte Methodik entwickelt. Diese folgt dem im Stand der Technik verbreiteten integrierten Ansatz zur Lösung des Pufferallokationsproblems. Sie setzt sich aufgrund der aufgezeigten Vorteile aus einem simulationsbasierten evaluativen Methodenbaustein (EMB) und einem generativen Methodenbaustein (GMB) zur statistischen Versuchsplanung zusammen.

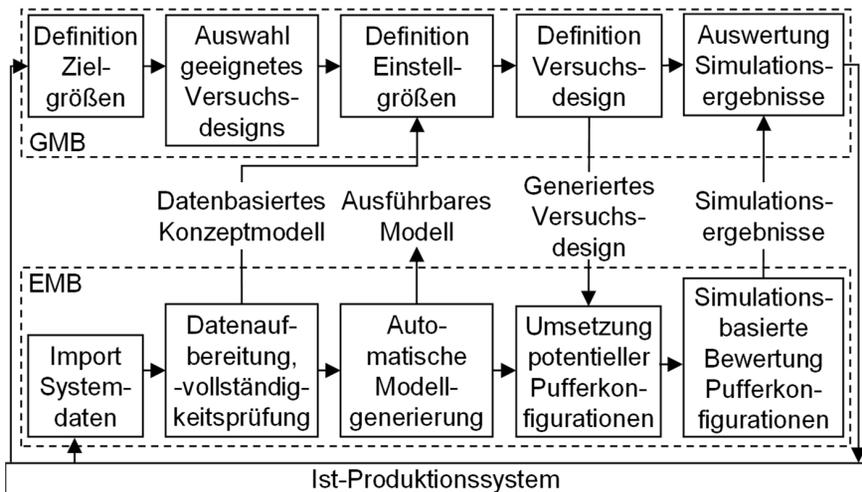


Abbildung 1: Konzept zur simultanen Allokation von Bestands-, Kapazitäts-, und Zeitpuffern

3.1 Evaluativer Methodenbaustein

Der evaluative Methodenbaustein umfasst sowohl die Generierung als auch den Einsatz des Materialflusssimulationsmodells des zu betrachtenden Produktionssystems, um potentielle Pufferkonfigurationen zu bewerten (Abb. 1). Die Generierung des Materialflusssimulationsmodells erfolgt dabei entsprechend dem Ansatz der automatischen, datengetriebenen Modellgenerierung. Spezifischer ist die zur Anwendung kommende Vorgehensweise der Klassifizierung nach Eckhardt (2002) folgend als parametrischer Ansatz zu definieren. In diesem Rahmen erfolgt zunächst der Aufbau einer datenbezogenen Repräsentation des zu betrachtenden Produktionssystems und die Spezifikation definierter Parameter. Für den Datenimport herrscht dabei eine enge Nutzerführung vor, indem der Anwender über Dialoge explizite Informationen erhält, welche Daten in welchem Format zu importieren sind. Dadurch ist die Methodik in der Lage die Qualität und Validität der zu importierenden Daten positiv zu beeinflussen. Im Anschluss erfolgt die Aufbereitung der importierten Daten, d. h. Datenformatierung und -anreicherung sowie eine Vollständigkeitsprüfung. Im Falle fehlender Datensätze erfolgt die standardisierte Information des Anwenders, dass Daten zu ergänzen und mit welchen Verfahren diese zu erheben sind. Die so erfassten Systemlast-, Organisations- und technischen Daten werden in einem datenbasierten Konzeptmodell zusammengefasst. Dieses besitzt eine standardisierte Datenstruktur und stellt eine geeignete Basis zur Automatisierung der Modellgenerierung sicher. Dazu werden basierend auf aus dem Konzeptmodell zu extrahierenden Parametern, in Bibliotheken gespeicherte Simulationsbausteine durch den Modellgenerator selektiert und konfiguriert. Dabei wird nicht ausschließlich auf bereits in der zur Anwendung kommenden Software Siemens Plant Simulation (v16) vorhandene Simulationsbausteine zurückgegriffen. Vielmehr werden weiterführende Bausteine, die eine parameterbasierte, automatische Generierung von dynamischen Steuerungseinflüssen (bspw. Auftragserzeugung und -freigabe) ermöglichen, entwickelt und bereitgestellt.

3.2 Generativer Methodenbaustein

Der generative Methodenbestandteil nutzt anschließend das effizient generierte Simulationsmodell zur Identifikation optimierter Pufferkonfigurationen, bestehend aus Bestands-, Kapazitäts- sowie Zeitpuffern, und übernimmt somit die generative Funktion innerhalb der Gesamtmethodik (Abb. 1). Dieser Baustein integriert Ansätze der statistischen Versuchsplanung und stellt somit eine effiziente Planung und Auswertung der durchzuführenden Simulationsexperimente bzw. -ergebnisse sicher. In einem ersten Schritt erfolgt zunächst die Definition der zu betrachtenden Zielgrößen. Der Stand der Technik zum Pufferallokationsproblem stellt verschiedene Zielsetzungen, wie die Maximierung der Systemausbringung oder die Minimierung des mittleren Bestandsniveaus bzw. der im Produktionssystem existierenden Bestandspufferplätze, vor (Demir et al. 2014). Jedoch weist bereits Park (1993) darauf hin, dass in praktischen Anwendungen, die den Fokus der vorliegenden Methodik darstellen, i. d. R. monetäre Zielgrößen wie Kosten, Gewinn oder Umsatz heranzuziehen sind. Zudem eignen sich monetäre Zielgrößen als einheitliche Bewertungsgrundlage für die drei Pufferfaktoren Bestand, Kapazität und Zeit. Daher wird in der vorliegenden Methodik die Zielgröße der Gesamtpufferkosten genutzt, um die Wirtschaftlichkeit potentieller

Pufferkonfigurationen zu bewerten. Diese repräsentieren die Summe der Kosten der eingesetzten Bestands-, Kapazitäts- und Zeitpuffer. Die Ermittlung dieser Kostengrößen hängt dabei stark von der Ausprägung der einzelnen Pufferarten, der Granularität der Kostenermittlung sowie der vom Anwender zu importierenden Kostengrößen ab. Die Bestandspufferkosten können sich aus Flächen- bzw. Mietkosten, Abschreibungskosten für Lagereinrichtungen sowie dem gebundenen Kapital für das in diesen enthaltenen Material ergeben. Zur Bewertung des Kapazitätspuffers lassen sich hingegen die anfallenden Personalaufwände sowie die eingesetzten Betriebsmittel durch entsprechende Stundensätze monetär bewerten. Zur Ermittlung der Zeitpufferkosten ist es erforderlich, die Abweichung des bestätigten Liefertermins vom Kundenwunschtermin monetär zu bewerten. Monetäre Auswirkungen einer derartigen Abweichung kann ein Umsatzverlust, bspw. durch Kundenauftragsverlust oder Verzugsstrafen sein. Neben Betrachtung des Kosteneinflusses potentieller Pufferkonfigurationen ist es zudem erforderlich, Anforderungen hinsichtlich der logistischen Leistungsfähigkeit des Gesamtsystems zu definieren. Dabei sind in Abhängigkeit des zugrundeliegenden Anwendungsfalls verschiedene logistische Zielgrößen, wie Liefertermintreue, Durchlaufzeit oder Durchsatz, zu nutzen. Auf Basis dieser Größen sind anschließend Rahmenbedingungen der Puffergestaltung zu definieren.

Im Anschluss an die Zielgrößendefinition ist ein geeignetes Versuchsdesign auszuwählen, wobei im Rahmen der Puffergestaltung die Nutzung eines Latin Hypercube Designs (LHD) als geeignet zu bewerten ist. LHD besitzen bei der Planung und Analyse von Simulations- bzw. Computerexperimenten einen hohen Verbreitungsgrad und weisen ein geeignetes Aufwand-Nutzen-Verhältnis auf, weshalb sie sich daher sehr gut zur Untersuchung großer Lösungsräume eignen (Viana 2013; Kleijnen 2015). Im Rahmen praxisrelevanter Aufgabenstellungen, die im Anwendungsfokus der Methodik stehen, ist der entsprechende Lösungsraum selten stark begrenzt, weshalb der Einsatz von LHD in diesen daher sehr geeignet ist. Darüber hinaus stellen LHD geringste Anforderungen hinsichtlich erforderlichen Vorwissens über bestehende funktionale Zusammenhänge zwischen den einzelnen Einflussgrößen (Viana 2013). Aufgrund der Tatsache, dass bei konzeptionellen bzw. gestalterischen Aufgabenstellungen wie der Puffergestaltung i. d. R. die funktionalen Zusammenhänge der einzelnen Puffergrößen nicht bekannt sind, ist dies ein zusätzliches Argument für den Einsatz eines LHD. Nach Auswahl eines geeigneten Versuchsdesigns sind die zu betrachtenden, variablen Einflussgrößen, d. h. die durch den Anwender definierten Bestands-, Kapazitäts- und Zeitpuffergrößen, zu übergeben. Die erforderlichen Daten bzw. Informationen sind dazu dem datenbasierten Konzeptmodell zu entnehmen. Darüber hinaus liefert das Konzeptmodell Informationen über zu berücksichtigende Extremwerte sowie nicht zulässige Wertebereiche der jeweiligen Puffergrößen, welche die im realen Anwendungsfall vorliegenden Rahmenbedingungen abbilden. Die gemäß grundlegender Vorgehensweise sich anschließende Festlegung der Einstellstufen der Einflussfaktoren entfällt, da entsprechend der Definition des LHD die Stufenanzahl je Einstellfaktor der Versuchsanzahl entspricht. Die Einstellstufen der Einflussfaktoren sind dabei zwischen dem oberen und dem unteren Grenzwert gleichverteilt. Die im so definierten LHD enthaltenen Pufferkonfigurationen werden anschließend an den evaluativen Methodenbaustein übergeben und in diesem simulationsbasiert bewertet.

Die Auswertung der Analyseergebnisse erfolgt zur Sicherstellung einer hohen Praxistauglichkeit innerhalb der vorgestellten Methodik mehrstufig. Zuerst erfolgt basierend auf den Simulationsergebnissen die Ermittlung einer Pareto-Front, welche diejenigen Punkte des Lösungsraums identifiziert, in denen eine Pufferkostenreduktion lediglich bei gleichzeitiger Verschlechterung der zu betrachtenden logistischen Zielerreichung möglich ist. Dadurch ist der Anwender bereits sehr schnell in der Lage für verschiedene Vorgaben der Logistikleistung kostenoptimierte Pufferkonfigurationen zu identifizieren. Darauf aufbauend erfolgt im Rahmen der Methodik die Anpassung eines Modells auf die erzeugten Simulationsdaten. Hierbei werden die Effekteinflüsse der verschiedenen Pufferfaktoren analysiert, um zunächst diejenigen mit signifikantem Einfluss zu identifizieren und Informationen im Hinblick auf die Effekte der Anpassung einzelner Pufferfaktoren zu erhalten. Als letzter Schritt der simultanen Puffergestaltung ist abschließend eine Optimierung des angepassten Modells vorzunehmen, um eine optimale Einstellung der Pufferfaktoren zu identifizieren, die abschließend im Ist-Produktionssystem zu implementieren ist.

4 Anwendung der Methodik in der Elektronikfertigung

Um die Praxistauglichkeit der Methode zu demonstrieren und weitere Impulse für eine anwendungsnahe Weiterentwicklung zu erhalten, wurde die Methodik anhand eines Fallbeispiels aus der variantenreichen Elektronikfertigung angewendet und auf ihre Eignung im Evaluationsprozess für Pufferdesigns getestet. Das Fallbeispiel bildet eine klassische Prozesskette der Elektronikfertigung ab, die sich aus einer Flachbaugruppenfertigung (kurz: Vorfertigung) und einer daran anschließenden Gerätemontage (kurz: Montage) zusammensetzt (Abb. 2).

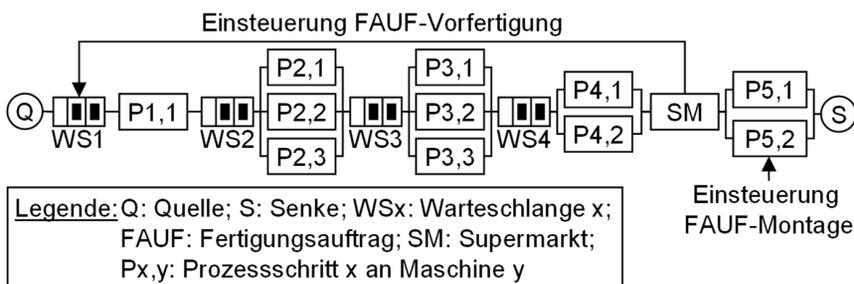


Abbildung 2: Prozesskette der variantenreichen Elektronikfertigung

An der Schnittstelle der Fertigungsstufen erfolgt die Entkopplung der Materialflüsse durch einen Flachbaugruppensupermarkt. Bei Einsteuerung eines Fertigungsauftrags in die Montage werden die erforderlichen Flachbaugruppen aus dem Supermarkt entnommen. Zur Wiederauffüllung des Supermarkts werden in Abhängigkeit dessen Füllstandes Fertigungsaufträge in die Vorfertigung eingesteuert. Innerhalb der Vorfertigung durchlaufen die Flachbaugruppen die Prozessschritte Surface-Mount Technology (Prozessschritt 1), In-Circuit-Test (Prozessschritt 2), Röntgeninspektion (Prozessschritt 3) und Through-Hole Technology (Prozessschritt 4), bevor sie in den

Supermarkt eingelagert werden. Die Verkettung der verschiedenen Prozessschritte erfolgt durch Warteschlangen. Die anschließende Gerätemontage (Prozessschritt 5) erfolgt auf zwei parallelen Montagelinien.

Die Variabilität im System wird durch stochastisch verteilte Bearbeitungszeiten sowie den zeitlich und mengenmäßig schwankenden Auftragseingang induziert. Innerhalb des Systems erfolgt die Pufferung vorherrschender Variabilitätseinflüsse durch eine Kombination aus Zeit-, Kapazitäts- und Bestandspuffern, die in Form der folgenden Pufferfaktoren ausgeprägt sind:

- Bestandspuffer: Supermarktkapazität (variantenspezifisch) (X1 bis X20) [Stück]
- Bestandspuffer: Warteschlangenkapazität Vorfertigung (X23 bis X26) [Stück]
- Kapazitätspuffer: Schichtanzahl Vorfertigung (X21) und Montage (X22) [Stück]
- Zeitpuffer: Differenz zwischen bestätigtem Liefertermin und Kundenwunschtermin (X27) [Tage]

Die zu optimierende logistische Zielgröße des Systems ist die Liefertermintreue. Diese ist als prozentualer Anteil der Aufträge definiert, deren tatsächlicher Liefertermin vor dem zugehörigen bestätigten Liefertermin liegt. Anhand dieser Zielgröße ist eine akzeptable logistische Leistung durch eine Liefertermintreue über 90 % definiert, bei dessen Erreichen die Pufferkosten zu minimieren sind. Basierend auf den oben genannten Pufferfaktoren wurde ein LHD (Versuchszahl:1000) in der Statistiksoftware JMP erzeugt und in die Simulationssoftware Plant Simulation integriert. Die Rückführung der Simulationsergebnisse und deren Analyse in JMP ermöglichen die nachfolgenden Ergebnisdiskussionen.

In der Analyse der Simulationsergebnisse zeigt sich als kostenoptimaler pufferbezogener Betriebspunkt die Faktorzusammenstellung [662; 566; 151; 240; 443; 946; 832; 595; 474; 247; 829; 875; 359; 679; 78; 474; 269; 66; 286; 518; 15; 21; 2091; 105; 734; 1105; 1] bei einer Liefertermintreue von 100 %. Die Generierung einer Pareto-Front bestätigt die starke Dominanz des gefundenen Punktes und zeigt auf, dass eine Pufferkostenreduktion lediglich unter teils erheblichen Leistungseinbußen (Liefertermintreue < 90 %) möglich ist. Eine Analyse der Effekteinflüsse zeigt, dass die Ergebnisse am stärksten positiv mit einer Erhöhung der Schichtanzahl in der Vorfertigung sowie insbesondere in der Endmontage korrelieren. Zudem hat eine Erhöhung der variantenspezifischen Supermarktkapazitäten, insbesondere für unregelmäßig und in großen Los nachgefragte Varianten des Auftragsprogramms, einen starken Einfluss. Die Nutzung des Zeitpuffers besitzt zwar einen hohen Einfluss hinsichtlich der Liefertermintreue, ist jedoch aus Kostengründen in den präferierten Lösungen nur schwach vertreten. Ein geringer Einfluss auf die Liefertermintreue und hohe Pufferkosten führten zu einer niedrigen Auslegung der Bestandspuffer in Form der Warteschlangen in der Vorfertigung. Hinsichtlich der Bewertung des gefundenen Ergebnisses ist darauf hinzuweisen, dass die gefundene Lösung keine universell einsetzbare Pufferkonfiguration ist. Vielmehr stellt diese eine spezifische, optimierte Pufferkonfiguration für das zu Grunde gelegte Produktionsprogramm dar. Bei einer deutlichen Abweichung von diesem Produktionsprogramm sieht die Methodik ein erneutes Durchlaufen des Pufferauslegungszyklus vor.



5 Zusammenfassung und Fazit

Im Rahmen des vorliegenden Beitrags wurde ein neuartiger Ansatz zur simultanen Allokation von Bestands-, Kapazitäts- und Zeitpuffern beschrieben. Dieser integriert die Methodik der ereignisdiskreten Simulation zur Evaluation potentieller Pufferkonfigurationen mit der statistischen Versuchsplanung zur Generierung optimierter Pufferkonfigurationen. Der evaluative Methodenbaustein umfasst dabei erste Ansätze der datengetriebenen Modellgenerierung, um den anwenderseitigen Bedarf an Expertenwissen und den Zeitaufwand der Modellgenerierung zu reduzieren. Die vorgestellte Methodik wurde prototypisch anhand eines Fallbeispiels angewendet, wobei die grundlegende Funktion der Methodik demonstriert werden konnte.

Weiterer Forschungsbedarf besteht hinsichtlich einer ganzheitlichen Validierung der Gesamtmethodik. Diesbezüglich ist die Anwendung in unterschiedlichen Fertigungsszenarien geplant, wodurch die Sensitivität der Ergebnisse hinsichtlich verschiedener Systemtypologien, Produktionsprogramme und Kostenfunktionen untersucht werden kann. Zudem ist die Ergebnisauswertung sowie die Optimierung des angepassten mathematischen Modells zur Identifikation einer optimalen Pufferkonfiguration im Rahmen weiterer Anwendungsfälle zu validieren. Im Rahmen der Modellgenerierung besteht weiterer Forschungsbedarf hinsichtlich einer automatischen Initialisierung und Validierung des Simulationsmodells.

Literatur

- Almomani, M. H.: A selection approach for solving buffer allocation problem. In: *International Journal of the Physical Sciences* 7 (2012) 3, S. 413-422
- Alon, G.; Kroese, D. P.; Raviv, T.; Rubinstein, R. Y.: Application of the cross-entropy method to the buffer allocation problem in a simulation-based environment. In: *Annals of Operations Research* 134 (2005) 1, S. 137-151
- Can, B.; Beham, A.; Heavey, C.: A comparative study of genetic algorithm components in simulation-based optimisation. In: Mason, S.; Hill, R.; Mönch, L.; Rose, O.; Jefferson, T.; Fowler, J. W. (Hrsg.): *Proceedings of the 2008 Winter Simulation Conference (WSC), Miami (USA), 7.-10. December 2008*, S. 1829-1837.
- Chaharsooghi, S. K.; Nahavandi, N.: Buffer allocation problem, a heuristic approach. In: *Scientia Iranica* 10 (2003) 4, S. 401-409
- Colledani, M.; Ekvall, M.; Lundholm, T.; Moriggi, P.; Polato, A.; Tolio, T.: Analytical methods to support continuous improvements at Scania. In: *International Journal of Production Research* 48 (2010) 7, S. 1913-1945.
- Costa, A.; Alfieri, A.; Matta, A.; Fichera, S.: A parallel tabu search for solving the primal buffer allocation problem in serial production systems. In: *Computers & Operations Research* 64 (2015), S. 97-112.
- Demir, L.; Tunali, S.; Eliiyi, D. T.: An adaptive tabu search approach for buffer allocation problem in unreliable non-homogenous production lines. In: *Computers & Operations Research* 39 (2012) 7, S. 1477-1486.
- Demir, L.; Tunali, S.; Eliiyi, D. T.: The state of the art on buffer allocation problem. A comprehensive survey. In: *Journal of Intelligent Manufacturing* 25 (2014) 3, S. 371-392.

- Deuse, J.; Heuser, C.; Konrad, B.; Lenze, D.; Maschek, T.; Wiegand, M.; Willats, P.: Pushing the limits of lean thinking–Design and management of complex production systems. In: Viles, E.; Ormazábal, M.; Lleó, A. (Hrsg.): Closing the gap between practice and research in industrial engineering, Cham: Springer International Publishing 2018, S. 335-342.
- Eckardt F.: Ein Beitrag zu Theorie und Praxis datengetriebener Modellgeneratoren zur Simulation von Produktionssystemen. Aachen: Shaker 2002.
- Hopp, W. J.; Spearman, M. L.: To pull or not to pull: What is the question? In: Manufacturing & Service Operations Management 6 (2004) 2, S. 133-148.
- Hopp, W. J.; Spearman, M. L.: Factory physics, 3. Edition. Boston: McGraw-Hill 2009.
- Kleijnen, J. P. C.: Design and Analysis of Simulation Experiments, 2. Edition. Cham: Springer 2015.
- Lorentzen, K.; Maschek, T.; Richter, R.; Deuse, J.: Entwicklung einer Variabilitätstypologie. In: Zeitschrift für wirtschaftlichen Fabrikbetrieb 106 (2011) 4, S. 214-218.
- MacGregor Smith, J.; Cruz, F.R.B: The buffer allocation problem for general finite buffer queueing networks. In: IIE Transactions 37 (2005) 4, S. 343-365.
- Maschek, T.; Heuser, C.; Hasselmann, Veit-Robert; Deuse, Jochen; Willats, Peter: Variabilitätsbasierte Klassifizierung von Produktionssystemen. In: Zeitschrift für wirtschaftlichen Fabrikbetrieb 109 (2014) 9, S. 591-594.
- Narasimhamu, K. L.; Reddy, V. Venugopal; Rao, C.S.P.: Optimal buffer allocation in tandem closed queuing network with single server using PSO. In: Procedia Materials Science 5 (2014), S. 2084-2089.
- Palominos, P.; Quezada, L.; Moncada, G.: Modeling the response capability of a production system. In: International Journal of Production Economics 122 (2009) 1, S. 458-468.
- Papadopoulos, H.; Vidalis, M.: Optimal buffer storage allocation in balanced reliable production lines. In: International Transactions in Operational Research 5 (1998) 4, S. 325-339.
- Park, T.: A two-phase heuristic algorithm for determining buffer sizes of production lines. In: International Journal of Production Research 31 (1993) 3, S. 613-631.
- Pound, E. S.; Spearman, M. L.; Bell, J. H.: Factory physics for managers. How leaders improve performance in a post-lean six sigma world. Europe: McGraw-Hill Education Ltd 2014.
- Siebertz, K.; van Bebber, D.; Hochkirchen, T.: Statistische Versuchsplanung. Design of Experiments, 2. Auflage. Berlin, Heidelberg: Springer Vieweg 2017.
- Spinellis, D.; Papadopoulos, C. T.: A simulated annealing approach for buffer allocation in reliable production lines. In: Annals of Operations Research 93 (2000) 1, S. 373-384.
- Viana F. A.: Things you wanted to know about the latin hypercube design and were afraid to ask. In: 10th World congress on structural and multidisciplinary optimization, Orlando (USA) 19.-24. May 2013, S. 1-9.
- Wiegand, M.; Konrad, B.; Jalali Sousanabady, R.; Willats, P.; Deuse, J.: Quantifizierung von Variabilität durch Big Data Technologien. In: PRODUCTIVITY Management 19 (2014) 5, S. 35-38.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Simulationsbasierte Validierung eines automatisierten Produktions- und Materialflusssteuerungssystems

Simulation-based validation of an automated production and material flow control system

Isabella Lichtenstern, Stefanie Wucherer, Florian Kerber,
Technologietransferzentrums für flexible Automation (TTZ), Hochschule Augsburg,
Augsburg (Germany), isabella.lichtenstern@hs-augsburg.de, stefanie.wucherer@hs-augsburg.de, florian.kerber@hs-augsburg.de

Steffen Klarmann, Valeo Schalter und Sensoren GmbH, Wemding (Germany),
steffen.klarmann@valeo.com

Abstract: In recent years, more and more simulation tools have been used by manufacturing companies to virtually commission, update and optimize increasingly complex production systems. This paper presents a simulation-based approach and first results for the validation of a flexible production and material flow control system called *ProLogCloud*. Integrated at the MES level, *ProLogCloud*, which was developed as part of a publicly funded research project, schedules assembly orders thus controlling a set of manufacturing islands using automated guided vehicles. To validate the scheduling algorithms and to evaluate the overall performance of the system, a simulation model of the assembly area has been set up. A relational database constitutes the central interface to exchange production data and commands with the system. The simulation setup is used to analyze and optimize the production and material flow control system in different scenarios. To reduce the effort required to instantiate the model, a parametric model generator is used to create different test scenarios. The simulation model also contains an integrated fleet management system that is used to coordinate the automated guided vehicles and assign them transport orders.

1 Einleitung und Motivation

Die Automobilindustrie ist durch einen stetigen technologischen Wandel und einen hohen Wettbewerbsdruck gekennzeichnet. Insbesondere Innovationen auf dem Gebiet der Elektromobilität und des autonomen Fahrens beeinflussen neben den klassischen OEMs auch die Zulieferindustrie. Um die Marktanforderungen zu erfüllen, müssen diese Unternehmen neben kontinuierlicher Verbesserung der Produkte auch die

erforderlichen Fertigungsverfahren optimieren. Die gestiegenen Kundenanforderungen an Individuallösungen erhöhen die Komplexität der Produktionsplanung und des Materialflusses und führen zu verkürzten Planungszyklen mit häufigem Umrüsten. Dies stellt Produktionsplaner vor enorme Herausforderungen. Um die Flexibilität der Materialflusssteuerung zu erhöhen, setzen außerdem immer mehr Firmen auf einen automatisierten Materialtransport mit fahrerlosen Transportfahrzeugen (FTF), deren Einsatz ebenfalls koordiniert werden muss (L. Sabattini u. a. 2013).

Am Beispiel einer vollautomatischen Produktion von Ultraschallparkensensoren soll deshalb im Rahmen eines öffentlichen Forschungsprojekts ein zusätzliches Steuerungssystem namens *ProLogCloud* als Produktionsleitsystem zur automatisierten Materialflusssteuerung mit FTF entwickelt werden. Dabei wird die bisher realisierte starre Verkettung der einzelnen Fertigungsschritte aufgelöst und ein vollflexibles Hybrid Flow Shop Prinzip mit parallelen Produktionsinseln umgesetzt. Neben der Prozesslogistik werden auch die einzelnen Fertigungsverfahren optimiert. Die Absicherung der zu entwickelnden Systemlösung kann dabei nicht in der realen Produktionsumgebung erfolgen, sondern muss durch digitale Methoden vorbereitet werden. Anstatt eines teuren, risikoreichen und fehleranfälligen Validierungsansatzes bei der Realinbetriebnahme im Feld, kann durch den gezielten Einsatz einer Materialflusssimulation die Leistungsfähigkeit und Robustheit des Steuerungssystems anhand von Worst-Case-Szenarien, Maschinenausfällen und anderen Risiken bereits während der Entwicklungsphase evaluiert werden (Versteegt/Verbraeck 2002).

Im vorliegenden Beitrag wird die Verknüpfung des flexiblen Produktions- und Materialflusssteuerungssystems *ProLogCloud* auf MES- Ebene mit einem Simulationsmodell vorgestellt. Darin übernimmt die *ProLogCloud* die Planung und Steuerung aller beteiligten Produktionskomponenten – dazu zählen Fertigungseinheiten, fahrerlose Transportsysteme, Warenhäuser und Puffer. Diese werden im Simulationswerkzeug *Plant Simulation* digital abgebildet und über definierte Schnittstellen mit dem Produktionssteuerungssystem gekoppelt. Neben grundlegenden Funktionstests werden insbesondere die Integrationsaufgaben für die FTF sowie Performanceanalysen der implementierten Planungsalgorithmen durchgeführt.

2 Literaturübersicht

In den vergangenen Jahren sind immer mehr Simulationssysteme bei produzierenden Firmen zum Einsatz gekommen, um die zunehmend komplexer werdenden Produktionsprozesse abzusichern (Daniel 2020). Im Folgenden wird eine Auswahl an Forschungsergebnissen zur simulationsbasierten Produktionsplanung und –steuerung zusammengefasst. In (Golmohammadi 2015) werden eine Reihe von Simulationsszenarien für einen Master-Produktionsplan und die Drum-Buffer-Rope-Terminierungsmethode vorgestellt. Die Optimierungstechniken werden verwendet, um optimale Werte für die Eingabevariablen im Simulationsexperiment zu finden. In (Volling/Spengler 2011) wird ein Framework aus separaten, miteinander verknüpften quantitativen Modellen für die Planung der auftragsgesteuerten Automobilproduktion aufgebaut und deren Potenzial mit Hilfe der Simulation evaluiert. (Yang/Arndt/Lanza 2016) stellt ein Simulationssystem vor, das Produktionspläne und prognostizierte

Kundenaufträge integriert, um Produktionsunsicherheiten zu minimieren und die Produktionsleistung zu optimieren. (Gyulai/Kádár/Monosotori 2015) stellt eine Planungs- und Steuerungsmethodik vor, die auf adaptiven Berechnungen basiert. Dabei werden historische Daten als direkter Input für ereignisdiskrete Simulationen verwendet, um den Kapazitätsbedarf der künftigen Produktion zu bestimmen. Außerdem verbindet (Gyulai u. a. 2016) ein Simulationsmodell direkt mit der Steuerung von modularen Zellen und evaluiert verschiedene Produktionspläne mit Hilfe der Simulation. Als Ergänzung eines digitalen Zwillings wird in (Herlyn/Zadek 2020) ein digitaler Steuerungszwilling vorgestellt, der die Auftragsabwicklung dynamisch steuert. Er überwacht die virtuelle Abbildung des Materialflusses. Bei zu großen Abweichungen zwischen Soll- und Ist-Werten leitet der digitale Steuerungszwilling entsprechende Optimierungsmaßnahmen ein. (Pfeiffer/Kádár/Monostori 2003) analysiert die Steuerung einer Transportanlage und eines automatisierten Regalbediengerätes mittels einer Simulationsstudie. (Block/Lins/Kuhlenkötter 2018) stellt den Ansatz eines ereignisgesteuerten Assistenzsystems für Produktionsplaner vor, das auf einer ganzheitlichen Fertigungsmodellierung basiert. Diese Modelle werden in Simulationen wiederverwendet, die mit Optimierungsalgorithmen verknüpft sind.

Im vorliegenden Beitrag wird die Umstellung auf einen vollautomatisierten Materialfluss bei gleichzeitiger Entkopplung der einzelnen Fertigungsschritte mittels digitaler Methoden vorgestellt. Bisher überwiegend manuell ablaufende Produktionsplanungsprozesse übernimmt und automatisiert die *ProLogCloud*. Um deren Planungsergebnis zu validieren und auf Robustheit zu prüfen, werden die generierten Produktions- und Materialflussprozesse innerhalb einer Simulation ausgeführt und getestet. Als zusätzliches Merkmal soll der Materialfluss komplett durch automatisierte Transportsysteme übernommen werden.

3 Aufbau und Funktion des Simulationssystems

Die Produktionsplattform *ProLogCloud* wird auf MES-Ebene entwickelt, um individuell ansteuerbare Fertigungsinseln über eine intelligente Prozesslogistik digital miteinander zu vernetzen und den Produktionsbereich vollständig zu automatisieren. Durch ein ganzheitliches Integrationskonzept mit universellen Schnittstellen zu allen Ebenen der IT-Architektur soll die Übertragbarkeit des Konzepts auf unterschiedliche Branchen und Unternehmensstrukturen ermöglicht werden. Ziel ist es, trotz einer immer größeren Produktvarianz eine äußerst flexible und effiziente Fertigung zu ermöglichen.

Mit Hilfe einer computergestützten Simulation soll die cloudbasierte Produktionsplattform *ProLogCloud* bereits vor der realen Inbetriebnahme beim Kooperationspartner virtuell in einer Software-in-the-loop (SIL) -Simulation aufgebaut und validiert werden. Die SIL-Simulation ist Teil der X-in-the-loop (XIL) - Entwicklungs- und Testumgebungen. Das X ist dabei austauschbar und steht jeweils als Platzhalter für den zu untersuchenden Prüfling. Eine XIL- und damit auch die SIL-Simulation besteht grundsätzlich aus zwei Komponenten, dem zu testenden System, das sogenannte System-under-test (SUT) und einem Simulator, der auf die Signale und Zustandsänderungen der SUT reagiert. Software-in-the-loop Simulationen erfordern keine physikalische Hardware, sondern finden ausschließlich in einer simulierten Umgebung statt. Durch Anschluss der *ProLogCloud*-Software an ein

digitales Anlagenmodell können Entwickler den Quellcode ohne teurere Systeme, Prototypen oder Prüfstände direkt am PC verwenden. Ausschließlich auf Basis des vorliegenden Quellcodes werden die implementierten Funktionen auf Korrektheit geprüft. Bereits vor Beginn eines realen Komponententests können so Softwaretests erfolgen (Albers u. a. 2010). Wie in Abbildung 1 dargestellt, wird das Simulationsmodell im Rahmen der SIL-Technologie mit der *ProLogCloud*, dem System-under-test, über eine relationale Datenbank verknüpft und ersetzt als Bestandteil eines digitalen Fabrikmodells somit die physikalischen Schnittstellen.

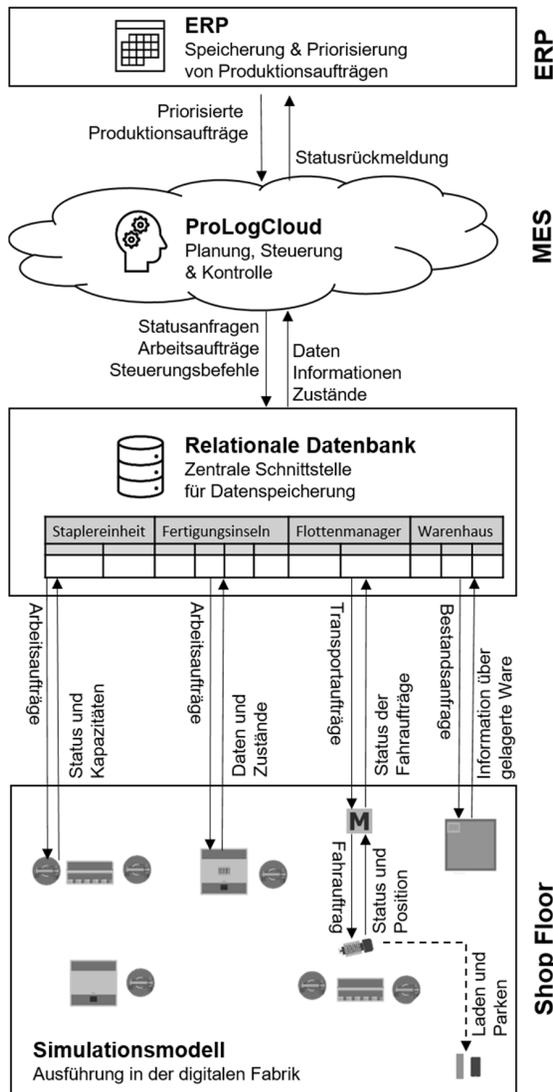


Abbildung 1: Funktionsweise und Ablauf der gesteuerten Simulation

Der Auftragseingang und die Priorisierung der Kundendaten erfolgt auf der hierarchisch höchsten Ebene der Unternehmensplanung. Das System *ProLogCloud* implementiert deshalb eine Schnittstelle zum ERP-System, um die Liste der priorisierten Auftragsdaten zyklisch abzufragen.

Die *ProLogCloud* übernimmt dann innerhalb der Produktionsprozesse die Planung und Steuerung des automatisierten Materialflusses und der Logistik mit Hilfe von FTF. Durch einen kontinuierlichen Datenaustausch zwischen den einzelnen Produktionskomponenten und der *ProLogCloud* optimiert sie mit Planungsalgorithmen in Echtzeit alle auftragsbezogenen Prozesse – von der Materialflusslogistik bis hin zum Einrichten der Steuerungsprogramme von einzelnen Produktionsinseln. Das Produktions- und Materialflusssteuerungssystem generiert außerdem Transportaufträge mit Start- und Zielpositionen inklusive Lieferzeiten für entsprechende Transportgüter.

Für die zentrale Datenspeicherung des Planungssystems *ProLogCloud* wurde eine relationale Datenbank gewählt. Zum einen wird dadurch die Rückverfolgbarkeit der Planungsentscheidungen gewährleistet, zum anderen die Echtzeitvisualisierung vereinfacht.

Die *ProLogCloud* gibt Steuerungsbefehle über die SQL-Datenbank an die entsprechenden Produktionskomponenten in der Simulationsumgebung weiter, die diese durch programmierte Steuerungsmethoden ausführen.

Die Instanziierung des produktionstechnischen Simulationsmodells erfolgt mittels eines parametrischen Modellgenerators (siehe Abbildung 2). Dieser liest unterschiedliche Datenformate zur Layout- und Hardwarekomponentendefinition wie CAD-, XLSX oder XML-Dateien ein und erweitert diese um relevante Prozessparameter, die die Ausführung von Testscenarien ermöglichen und das Modell für die SQL-Schnittstelle vorbereiten. Auf Basis dieser Parameter wird das Modell dann automatisch generiert und eine XML-Datei für die Initialisierung der *ProLogCloud* erzeugt. Mit Hilfe der XML-Datei werden anschließend die benötigten Steuerungen für die SQL-Schnittstelle instanziiert. So kann der

Implementierungsaufwand für unterschiedliche Szenarien deutlich reduziert werden.

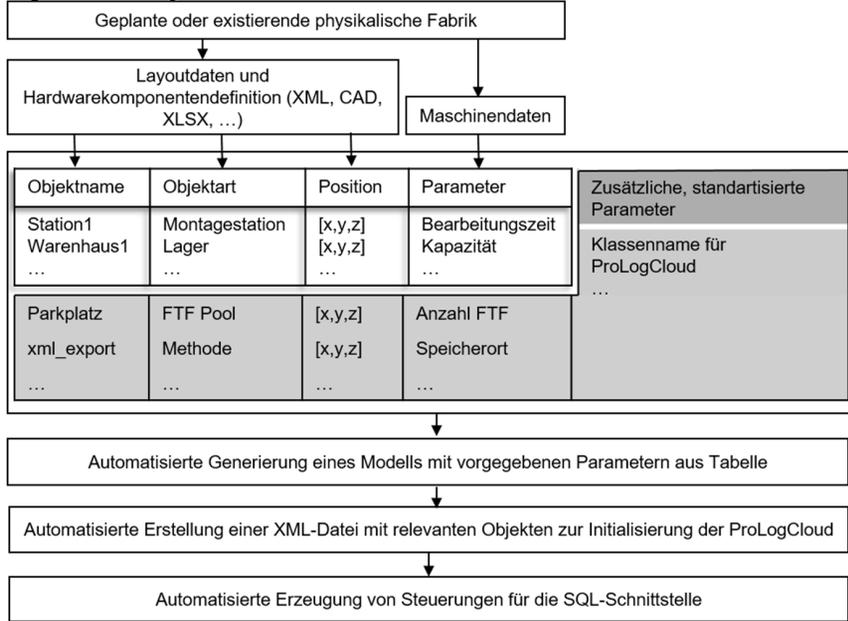


Abbildung 2: Parametrischer Modellgenerator

Wie im Simulationsmodell in Abbildung 1 dargestellt, bestehen die automatisch generierten Modelle aus folgenden Komponenten:

- FTF für den Warentransport mit Park- und Ladestation
- flexibel ansteuerbare Fertigungsinseln
- mit den Fertigungsinseln verbundene Palettiervorrichtungen, an denen zu bearbeitende Rohmaterialien und Halbfertigerzeugnisse von FTF angeliefert und von denen bearbeitete Halbfertigerzeugnisse wieder abtransportiert werden
- Warenhausstationen und Puffer als Schnittstellen zu vor- und nachgelagerten Fertigungsschritten, aus denen die FTF bedarfsorientiert die Inseinheiten beliefern

Der Materialfluss im betrachteten Produktionsbereich wird mit Hilfe einer proprietären Leitsteuerung simuliert, die die Zuweisung der Transportaufträge zu den einzelnen FTF übernimmt. Die priorisierten Transportaufträge empfängt sie von der *ProLogCloud*. Die Funktionsweise der Leitsteuerung wird in der Simulationsumgebung programmiert, um das reale Verhalten der Fahrzeuge so exakt wie möglich nachzubilden. Jedes FTF gibt seinen Status und die Position an die Leitsteuerung in Echtzeit zurück.

Aus den Experimenten gewonnene Daten werden an die *ProLogCloud* zurückgegeben, die kontinuierlich neu plant und wiederum entsprechende Steuerungsbefehle in die Datenbank einträgt.

4 Ziele des Simulationsmodells

Der Einsatz im Simulationsmodell bietet den Vorteil, das Verhalten der *ProLogCloud* an allen virtuellen Schnittstellen zu testen und dient bereits während der Entwicklungsphase als Absicherung und Validierung der Funktionalität. Da die Simulationsbausteine die realen Objekte in einer Testumgebung ersetzen, können Kosten eingespart werden – bei gleichzeitig hoher Flexibilität.

Die Simulation bietet zudem die Möglichkeit der Anwendung in unterschiedlichsten Fabriken, um so die universelle Einsatzbarkeit der *ProLogCloud* zu prüfen. Mit zunehmender Komplexität der Produktionsprozesse soll schrittweise das Optimierungspotenzial für die *ProLogCloud* evaluiert werden. Außerdem sollen unterschiedliche Planungsalgorithmen innerhalb des Steuerungssystems untersucht werden – in einer vollständig virtuellen Umgebung, deren dynamische Abläufe auch beschleunigt werden können.

Tabelle 1 zeigt eine Übersicht zweier unterschiedlich komplexer Szenarien mit den zugehörigen Validierungszielen:

Tabelle 1: Übersicht der Simulationsszenarien

Szenario	Validierungsziel	Anzahl FTF	Komplexität
Sequenziell ablaufendes Flow-Shop Prinzip	Funktionsnachweis und Integrationstest	N=1	Reduziert
Parallel ablaufendes Hybrid Flow Shop-Prinzip	Performanceanalyse und Optimierung des Gesamtszenarios	N>1	Hoch

4.1 Optimierung des Gesamtsystems

Innerhalb der einzelnen Szenarien können verschiedenste FTF- und Fertigungsabläufe aufgebaut und mit der *ProLogCloud* gesteuert werden. Es ist möglich, unterschiedliche FTF-Leitsteuerungen zu vergleichen und die optimale Anzahl an FTF zu prognostizieren. Insbesondere werden das Verhalten in Kollisionsszenarien mit dem Personal oder anderen Transportfahrzeugen sowie Lade- und Stillstandszeiten untersucht. Außerdem besteht die Möglichkeit, das Hallenlayout bereits vor dem physikalischen Aufbau optimal auszulegen.

4.2 Performanceanalyse

Die Datenanalyse im Simulationsmodell erzeugt während des gesamten Produktentstehungsprozesses KPIs, auf deren Grundlage die einzelnen Parameter in der Produktion optimiert werden können. Leistungskennzahlen wie Termintreue, Durchlauf- und Rüstzeiten sowie Auslastungen werden in der Simulation gemessen, visualisiert und anschließend analysiert. Außerdem werden Ursachen und Auswirkungen von Störungen, Engpässen und negative Einflüsse wie menschliches Fehlverhalten oder technische Störungen detektiert.

Hierfür werden u. a. folgende Störfälle evaluiert:

- Produktion von Ausschussmaterial an einzelnen Fertigungsinseln
- Schwankungen der Durchlaufzeit einzelner Fertigungsinseln
- Zu geringe Lagerkapazitäten an den Palettiervorrichtungen, Puffern oder im Warenhaus
- Durch Hindernisse blockierte Routen und Stillstand einzelner FTF

5 Implementierung

Abbildung 3 zeigt die Einbettung der einzelnen Elemente in die SIL-Simulationsumgebung. Für den Integrationstest wurde das SUT *ProLogCloud* auf einem separaten Steuerungsrechner installiert, das Simulationsmodell läuft auf einem eigenen Simulationsrechner. Beide Komponenten sind mit dem Datenbankserver verbunden.

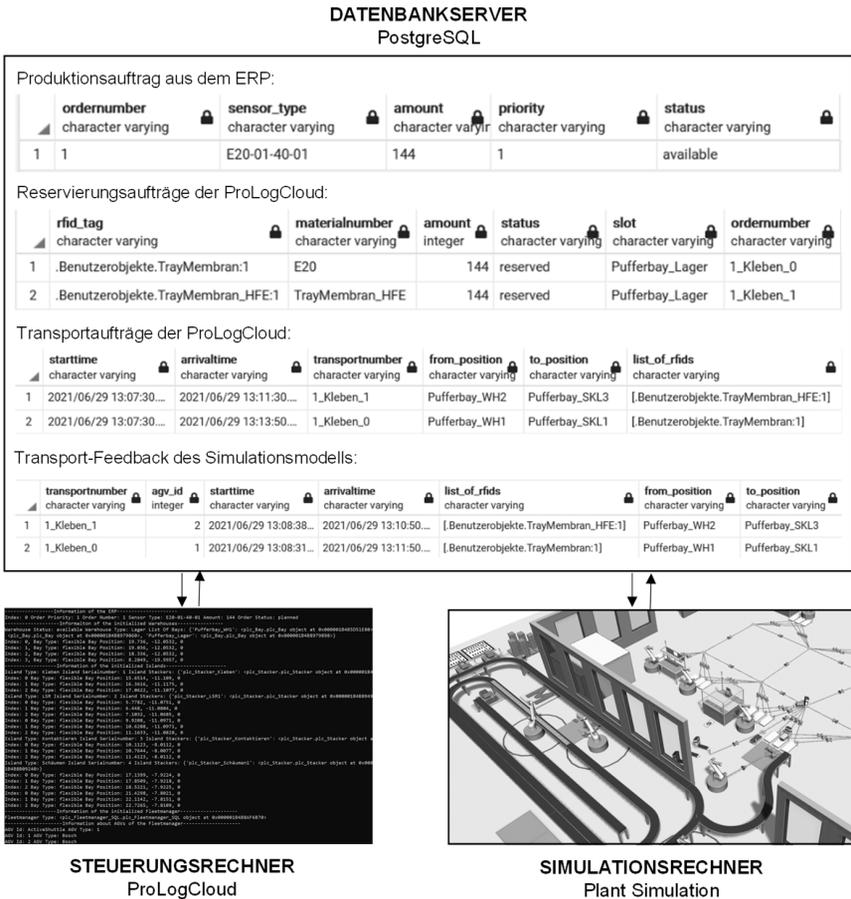


Abbildung 3: Implementierung des Simulationssystems

In Abbildung 3 wird dargestellt wie der Produktionsauftrag „1“ mit 144 Sensoren des Typs „E20-01-40-01“ abgearbeitet wird. Die *ProLogCloud* berechnet zunächst für den ersten Fertigungsschritt benötigte Rohmaterialien und Leergut-Objekte und prüft über den Datenbankserver, ob diese im Warenhaus des Simulationsmodells verfügbar sind. Nach erfolgreicher Suche reserviert die *ProLogCloud* Membrane und Leergut durch einen Reservierungsauftrag. Das Simulationsmodell palettiert im Anschluss die angeforderte Ware auf Fördermittel und stellt sie über Förderbänder den FTF zur Abholung bereit. Die *ProLogCloud* generiert dann Transportaufträge mit entsprechenden Transportzeiten, um Leergut und Membrane vom Warenhaus zur ersten Fertigungsinsel zu befördern. Der integrierte Flottenmanager im Simulationstool weist diese Aufträge den FTF mit der ID 1 und 2 zu. Nach erfolgreichem Transport gibt das Simulationsmodell ein Feedback. Die eingetroffene Ware wird an der Fertigungsinsel bearbeitet und anschließend auf das Leergut umgelagert. Die *ProLogCloud* berechnet und veranlasst währenddessen bereits die nächsten Produktions- und Logistikprozesse.

Damit konnten alle notwendigen Schnittstellen für die Koppelung des Simulationsmodells mit der *ProLogCloud* geschaffen werden. Der Integrationstest wurde anhand des Simulationsszenarios eines sequenziell ablaufenden Flow-Shop Prinzips durchgeführt. Das Simulationsmodell wird durch Reservierungs- und Bereitstellungsanweisungen im Warenhaus, Fahraufträge und Lieferzeiten für die FTF und Fertigungsaufträge für Inseln und Palettiervorrichtungen durch die *ProLogCloud* fremdgesteuert. Zustände, Zeiten und Informationen zum Fortschritt des Produkts werden in Echtzeit für eine erneute Planung erfolgreich zurückgegeben. Die im Modell programmierte Leitsteuerung koordiniert die Fahrzeuge und weist diesen Transportaufträge zu.

6 Zusammenfassung

In diesem Beitrag wurde die Kopplung eines Produktions- und Materialflusssteuerungssystems namens *ProLogCloud* auf MES-Ebene mit einem Simulationsmodell vorgestellt. Im Rahmen des Forschungsprojekts konnten alle Schnittstellen geschaffen werden, um die Validierung der *ProLogCloud* in einer Software-in-the-loop Simulation zu ermöglichen. Der parametrische Modellgenerator ermöglicht dabei eine flexible und schnelle Instanziierung von verschiedenen Szenarien. Der SIL-Aufbau wurde vollständig implementiert und der Integrationstest von Steuerungssoftware und Simulationsmodell erfolgreich durchgeführt.

Im nächsten Schritt soll dann in Anlehnung an den industriellen Anwendungsfall ein Hybrid Flow Shop-Szenario zur Validierung der Planungsalgorithmen der *ProLogCloud* untersucht und deren Performance anhand von KPIs bewertet werden. Dabei wird der Detaillierungsgrad der Modellkomponenten wie beispielsweise des im Modell integrierten Flottenmanagers zunehmend geschärft.

Danksagung

Diese Arbeit wurde gefördert durch das Bayerische Staatsministerium für Wirtschaft, Landesentwicklung und Energie im Rahmen des Bayerischen Förderprogramms für Forschung und Entwicklung „Elektronische Systeme“.

Literatur

- Albers, A. u. a.: X-in-the-Loop-Framework für Fahrzeuge, Steuergeräte und Kommunikationssysteme, in: ATZechnik, 5. Jg., 2010, H. 5, S. 60–65.
- Block, C.; Lins, D.; Kuhlenkötter, B.: Approach for a simulation-based and event-driven production planning and control in decentralized manufacturing execution systems, in: Procedia CIRP, 72. Jg., 2018, S. 1351–1356.
- Daniel, C.: Simulationsplattform für Automatisierungslösungen, in: ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb, 115. Jg., 2020, special, S. 32–35.
- Golmohammadi, D.: A study of scheduling under the theory of constraints, in: International Journal of Production Economics, 165. Jg., 2015, S. 38–50.
- Gyulai, D. u. a.: Simulation-based Production Planning and Execution Control for Reconfigurable Assembly Cells, in: Procedia CIRP, 57. Jg., 2016, S. 445–450.
- Gyulai, D.; Kádár, B.; Monostori, L.: Robust production planning and capacity control for flexible assembly lines, in: IFAC-PapersOnLine, 48. Jg., 2015, H. 3, S. 2312–2317.
- Herlyn, W. J.; Zadek, H.: Der Digitale Steuerungs-Zwilling, in: ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb, 115. Jg., 2020, special, S. 70–73.
- L. Sabattini; V. Digani; C. Secchi; G. Cotena; D. Ronzoni; M. Foppoli; F. Oleari: Technological roadmap to boost the introduction of AGVs in industrial applications, in: 2013 IEEE 9th International Conference on Intelligent Computer Communication and Processing (ICCP), 2013, S. 203–208.
- Pfeiffer, A.; Kádár, B.; Monostori, L.: Evaluating and improving production control systems by using emulation, in: undefined, 2003.
- Versteegt, C.; Verbraeck, A.: The extended use of simulation in evaluating real-time control systems of AGVs and automated material handling systems, in: Yücesan, E. (Hrsg.): Exploring new frontiers, 2002, S. 1659–1666.
- Volling, T.; Spengler, T. S.: Modeling and simulation of order-driven planning policies in build-to-order automobile production, in: International Journal of Production Economics, 131. Jg., 2011, H. 1, S. 183–193.
- Yang, S.; Arndt, T.; Lanza, G.: A Flexible Simulation Support for Production Planning and Control in Small and Medium Enterprises, in: Procedia CIRP, 56. Jg., 2016, S. 389–394.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Einfluss von Störungen auf den Umgang mit Deadlocks in einem fahrerlosen Transportsystem

Influence of disruptions on deadlock handling in an automated guided vehicle system

Marcel Müller, Jan Hendrik Ulrich, Tobias Reggelin, Hartmut Zadek,
Otto-von-Guericke-Universität Magdeburg, Magdeburg (Germany),
marcell.mueller@ovgu.de, jan.ulrich@st.ovgu.de, tobias.reggelin@ovgu.de,
hartmut.zadek@ovgu.de

Lorena Silvana Reyes Rubiano, University of La Sabana, Chía (Colombia),
lorena.reyes1@unisabana.edu.co

Abstract: Deadlock handling is an important task when planning automated guided vehicle systems. The selection of the right strategy approach has a major impact on logistical figures. This paper presents the results of a simulation model for a comparison of the three typical deadlock handling strategies prevention, avoidance, and detection & resolution for a transport system with disruptions in a warehouse, while varying the amount of AGVs. We investigated different disruption profiles with the variation of the mean time to repair (MTTR) and availability of the resources. The avoidance strategy receives the best results for different profiles of disruptions. Detection & resolution provides good results for non-frequent, small disruptions. The prevention strategy easily solves the deadlock problem but has the worst performance in terms of system throughput, average waiting time per transport order and average distance travelled per transport order.

1 Motivation und Problemstellung

Die Industrie 4.0 stellt als Kerndiskussionsthema in Produktion und Logistik Unternehmen vor ganz pragmatische Herausforderungen der Digitalisierung und Automatisierung. Für deutsche Unternehmen hat deshalb auch die Relevanz von konkreten Technologien wie Fahrerlose Transportsysteme (FTS) und Robotik in den letzten Jahren stark zugenommen (Kohl; Pfretzschner, 2018). Die Zunahme von automatisierten Systemen in Produktion und Logistik und das immer häufigere Ineinandergreifen von verschiedenen automatisierten Systemen in Unternehmen führen zu immer größeren Herausforderungen, bei der Beantwortung der Frage, wie mit Kollisionen und Deadlocks (zu Deutsch: Systemverklümmungen) in diesen komplexen Systemen am besten umgegangen werden kann. Besonders der Umgang

mit Deadlocks wird in der Logistikplanung nicht ausreichend berücksichtigt. Dabei kann das Auftreten eines Deadlocks zu einem kompletten Systemstillstand führen und der richtige Strategieansatz entscheidend für die Leistungsfähigkeit des Systems sein. Bei den Strategieansätzen wird grundsätzlich zwischen Prävention (engl.: „prevention“), Vermeidung (engl.: „avoidance“) und Erkennen und Auflösen (engl.: „detection & resolution“) unterschieden (Coffman et al., 1971). In der Literatur wird sich zum Umgang von Deadlocks ohne größere Untersuchungen für einen dieser drei Strategieansätze entschieden. In eigenen simulativen Experimenten konnte jedoch gezeigt werden, dass bereits kleinere Parameteränderungen zur Änderung des optimalen Strategieansatzes führen (Müller et al., 2020). Ein wichtiger Aspekt, der in diesen Experimenten jedoch noch nicht betrachtet wurde, ist das Auftreten von Störungen. Aus Gesprächen mit Experten aus der Industrie wurde festgestellt, dass im Betrieb von fahrerlosen Transportsystemen regelmäßig Störungen und ungeplante Haltemanöver auftreten. Welche Auswirkungen Störungen auf den optimalen Deadlock-Strategieansatz haben, ist bisher, auch in der Literatur, nicht bekannt und soll deshalb in dieser Veröffentlichung simulativ aufgezeigt werden. Durch die Berücksichtigung stochastischer Einflussgrößen und die Durchführung zahlreicher Zufallsexperimente kann eine möglichst induktive Antwort auf diese Fragestellung gefunden werden.

2 Wissenschaftliche Abgrenzung

Dieser Beitrag erweitert die Erkenntnisse von Müller et al. (2020) um Untersuchungen zum Einfluss von Störgrößen auf die Deadlock-Strategieansätze. Weitere eigene Vorarbeiten zum Thema Deadlocks umfassen einen Erkennen-&-Auflösen-Strategieansatz für ein schienenbasiertes Regalbediensystem (Müller et al., 2019). Grundsätzlich bedarf es bei dieser Untersuchung keiner großen Abgrenzung zur wissenschaftlichen Literatur, denn diese berücksichtigt Deadlocks in der Logistikplanung quasi nicht. In der Logistikplanung von Prozessen und Systemen spielen Deadlocks, wenn überhaupt, erst spät in der operativen Detailplanung von Steuerungsstrategien eine Rolle. Für einen Deadlock-Präventionsansatz ist es da meist zu spät. Grundsätzlich scheinen die klassischen Vorgehensweisen zur Planung in der Logistik, wie

- die prozesskettenbasierte Planung nach Kuhn und Bernemann (1995),
- die VDI-Richtlinie zur Systemfindung von Kommissioniersystemen (VDI-Richtlinie 3590),
- die Phasen der Planung und Realisierung von Logistiksystemen nach Gudehus (2011),
- und die vierstufige Vorgehensweise zur Planung von Transportsystemen nach van Bonn (2013),

Deadlocks nicht bei der Dimensionierung von Ressourcen oder der Infrastrukturplanung zu berücksichtigen. So erläutert beispielsweise Gudehus (2011, S. 869 ff.) detailliert wie die Ziele „Transportoptimierung“ und „Flächenminimierung“ bei der Auslegung von Logistikhallen mittels Toranordnung und weiteren Maßnahmen erreicht werden. Doch weder bei der Auslegung der Logistikhalle noch bei der Planung und Dimensionierung des Transportsystems geht Gudehus auf mögliche Deadlocksituationen ein, die bei einer vermeintlich günstigen

Auslegung auftreten könnten. Bei den anderen genannten Quellen wird die Problematik von Deadlocks ebenfalls in keiner Weise erwähnt und es kann nur gemutmaßt werden, welche Auswirkungen eine identifizierte Deadlocksituation auf die entsprechende Vorgehensweise hätte.

Im Bereich der Fabrikplanung, woran sich die Vorgehensweisen in der Logistikplanung häufig orientieren (Schenk et al., 2014; Tempelmeier, H. 2018), wird ebenfalls nicht klar, ob Deadlocks methodisch berücksichtigt werden. Im Vorgehen nach der VDI-Gesellschaft Produktionstechnik (VDI-Richtlinie 5200) könnte der Ansatz Deadlock-Prävention sowohl in der Phase 3 – Konzeptplanung bei der Dimensionierung der Betriebsmittel als auch anschließend Phase 4 – Detailplanung durchgeführt werden. Bei letzterem kommt es im Subprozess der Feinplanung zur „Planung der Materialflüsse, der Informationsflüsse und der Kommunikationsflüsse in Form von Prozessdarstellungen und Prozessbeschreibungen“ (VDI-Richtlinie 5200). Dabei erfolge auch „die Zuordnung von Produkten und Ressourcen zu den Prozessen, die Abfolge der Prozessschritte, deren organisatorische Eingliederung sowie die jeweils eingesetzten Arbeitshilfsmittel“ (VDI-Richtlinie 5200). Ob bei den Prozessdarstellungen oder der Zuordnung der Ressourcen Deadlocksituationen ersichtlich werden und wie der Konzept- oder Detailplan dann gegebenenfalls nochmal angepasst werden muss, berücksichtigt die VDI-Richtlinie nicht.

Zusammenfassend lässt sich sagen, dass es in der Literatur zwar Vorgehensweisen für die Logistikplanung gibt, wo sich eine Betrachtung und Analyse von möglichen Deadlocksituationen einordnen ließe, aber bisher keine Methodik für eine solche Betrachtung im Detail erläutert wird.

Dies führt dazu, dass Deadlocks häufig im Vorfeld der Planung nicht erkannt werden, sondern erst beim Testen der Steuerungs- und Entscheidungsregeln. Dies geschieht beispielsweise in einem Simulationsmodell, welches nach zahlreichen Simulationsdurchläufen selbst unwahrscheinliche Konstellationen im System, die zu einem Deadlock führen, hervorbringt.

Der Umgang in der Logistik mit den auftretenden Deadlock-Problematiken erfolgt dann häufig mit zusätzlichen problemspezifischen Steuerungsregeln (Mayer und Furmans, 2010) oder aber mit präventiven, deterministischen Maßnahmen durch Vorausplanung und Reservierung von Ressourcen (Kim et al., 2006; Lienert und Fottner, 2017), sofern die Informationslage dafür ausreichend ist. Beide Ansätze benötigen einen zentralen Steuerungsansatz und skalieren schlecht mit größer dimensionierten Systemen.

Dezentrale Steuerungsansätze in der Logistik, die eine solche Skalierbarkeit versprechen, setzen bisher ebenfalls auf problemspezifische Vermeidungsregeln durch Optimierungsalgorithmen für Deadlocks (Schönung et al., 2011; Seibold, 2016) und auf agentenbasierte Herangehensweisen (Forget et al., 2009; Mors, 2010; Yalcin, 2017; Lu et al., 2019).

Generell wird sich bei den problemspezifischen Lösungen bereits auf einen Strategieansatz festgelegt, so dass ein Vergleich zwischen den Strategien und unter Berücksichtigung von Störungen noch nicht durchgeführt wurde.

3 Vorgehensweise und Anwendbarkeit

Zur Untersuchung des Einflusses von Störungen auf den Umgang mit Deadlocks in fahrerlosen Transportsystemen wird auf ein bestehendes Simulationsmodell in der Simulationssoftware „Tecnomatix Plant Simulation“ aufgebaut (Müller et al., 2020). Es wurden Störungsprofile implementiert und Regeln für den Umgang bei Störungen definiert. Im Simulationsmodell wird ein flurgebundenes, fahrerloses Transportsystem in einem gassenbasierten Lager betrachtet. Zur Untersuchung von Störungen werden verschiedene Störungsprofile an unterschiedlichen Stellen im Simulationsmodell implementiert. Abbildung 1 zeigt eine Übersicht der verwendeten Eingangs-, Stell- und Ergebnisgrößen. Neben der Variation verschiedener Stellgrößen bezüglich Störungen, wurde auch untersucht, ob mit zunehmenden Verkehrsaufkommen Änderungen beim optimalen Strategieansatz auftauchen oder ob die Störungen einen so signifikanten Einfluss auf die Strategieauswahl haben können, so dass der bisherige Faktor der FTF-Anzahl in den Hintergrund rückt.

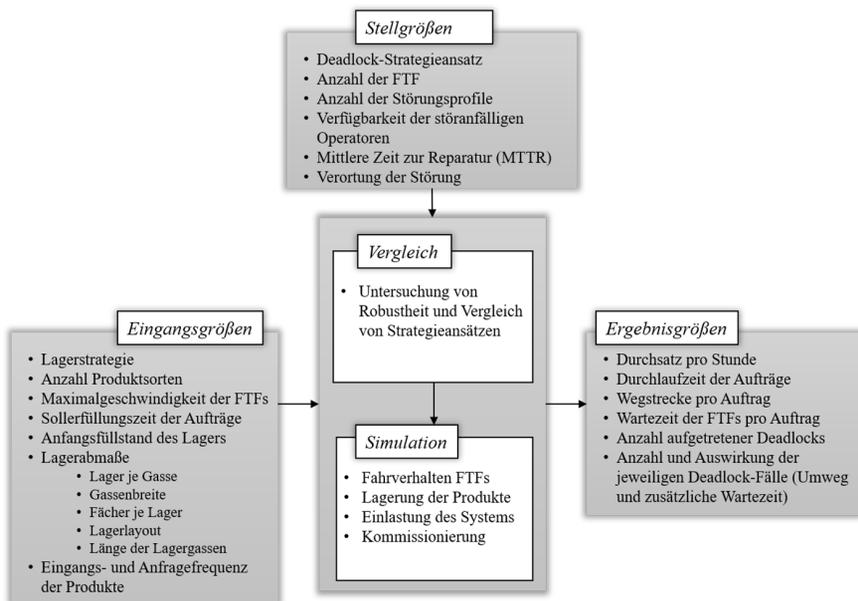


Abbildung 1: Übersicht der verwendeten Eingangs-, Stell- und Ergebnisgrößen, Darstellung in Anlehnung an März et al., 2011

Die Erkenntnisse aus den Simulationsergebnissen sind auch für die Industrie anwendbar, sofern in der Intralogistik ein fahrerloses Transportsystem verwendet wird. Durch die Variation der Störungsprofile soll eine Vergleichbarkeit mit realitätsnahen Anwendungsszenarien hergestellt werden können, so dass deutlich wird, bei welchen Störungen und Zielgrößen ein bestimmter Deadlock-Strategieansatz verfolgt werden kann.

Das Simulationsmodell kann mehrere Layouts abbilden (Müller et al., 2020). Für die Untersuchungen in diesem Artikel wurde jedoch nur das Layout 3 genutzt, da dort alle

drei Deadlock-Strategieansätze angewandt werden konnten. Die Präventionsstrategie wird durch unidirektionale Schleifen ermöglicht. Abbildung 2: zeigt die identifizierten Deadlock-Situationen für den Strategieansatz „Erkennen und Auflösen“.

	Deadlock-Situation:	Lösungsprinzip:
Fall 1: (Layout 1 & 3)		
Fall 2: (Layout 1 & 3)		
Fall 3: (Layout 3 & 4)		
Fall 4: (Layout 3)		

Abbildung 2: Mögliche Deadlock-Situationen und Lösungsansätze bei Erkennen und Auflösen als Strategieansatz (Müller et al., 2020)

4 Experimente und Ergebnisse

Je Experiment wurden zehn Simulationsdurchläufe durchgeführt. Der Simulationszeitraum beträgt acht Stunden. Tabelle 1 zeigt die konkrete Festlegung der Stellgrößen für die Experimentreihe. Im Summe ergeben sich aus dem Experimentplan 162 Experimente und folglich 1.620 Simulationsdurchläufe. Die Störungen wurden nur in den Fahrzeugen verortet. Eine Störung eines Fahrzeugs kam auch einer Blockierung/Störung des Weges gleich, so dass eine Betrachtung der unterschiedlichen Verortung der Störungen nicht weiter berücksichtigt wurde.

Tabelle 1: Übersicht der verwendeten Stellgrößen

Stellgröße	Wertebereich	Anzahl Experimente
Anzahl der FTF	5 - 10	6
Deadlock-Strategieansatz	Prävention (P), Erkennen und Auflösen (EA), Vermeidung (V)	3
Verfügbarkeit	91, 95, 99 [%]	3
MTTR	5, 30, 120 [min]	3

Durch die Kombination der Stellgrößen Verfügbarkeit und MTTR ergeben sich Störungsprofile mit beispielsweise wenigen, kurzen Störungen (Verfügbarkeit: 99 %, MTTR: 5 min) oder vielen, langen Störungen (Verfügbarkeit: 91 %, MTTR: 120 min).

Grundsätzlich bestätigt sich bei der Betrachtung der Simulationsergebnisse wieder die bereits bekannte Schlussfolgerung, dass der Durchsatz bei einem Präventionsansatz (P) deutlich geringer ausfällt (Müller et al., 2020). Abbildung 3 zeigt dies anhand eines Oberflächendiagramms. Zwar erhöht sich mit zunehmender FTF-Anzahl der Durchsatz, doch die Strategieansätze Erkennen & Auflösen (EA) und Vermeidung (V) erreichen den maximalen Durchsatz von 120 Stück pro Stunde schon bei sieben FTF, während ein Präventionsansatz selbst bei zehn FTF diesen Wert noch nicht erreicht. Bei diesem Störungsprofil treten zwischen Erkennen & Auflösen und einer Vermeidungsstrategie keine signifikanten Unterschiede hinsichtlich des Durchsatzes auf.

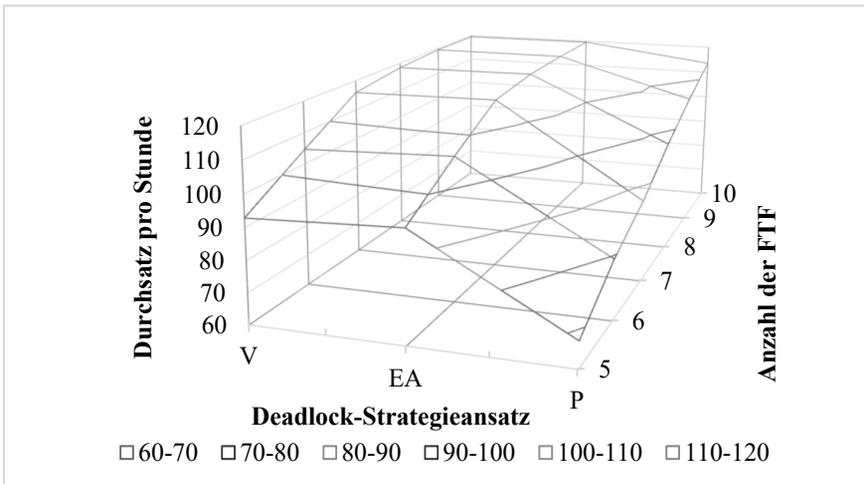


Abbildung 3: Oberflächendiagramm des Durchsatzes pro Stunde bei hoher Verfügbarkeit und geringer MTTR (Durchschnitt der 10 Simulationdurchläufe)

Treten die Störungen jedoch häufiger auf (geringere Verfügbarkeit), setzt sich ein Vermeidungsansatz gegenüber Erkennen & Auflösen durch. Während bei kurzen Störungen diese Dominanz nur bei einer mittleren Fahrzeuganzahl von sieben bis acht

FTF gegeben ist, zeigt sich bei schwereren Störungen ein noch deutlicheres Bild: Der Vermeidungsansatz kann hinsichtlich des Durchsatzes durchweg bessere Ergebnisse erzielen. In Abbildung 4 wird dies durch die konsequente Schräge der Oberfläche im Diagramm deutlich.

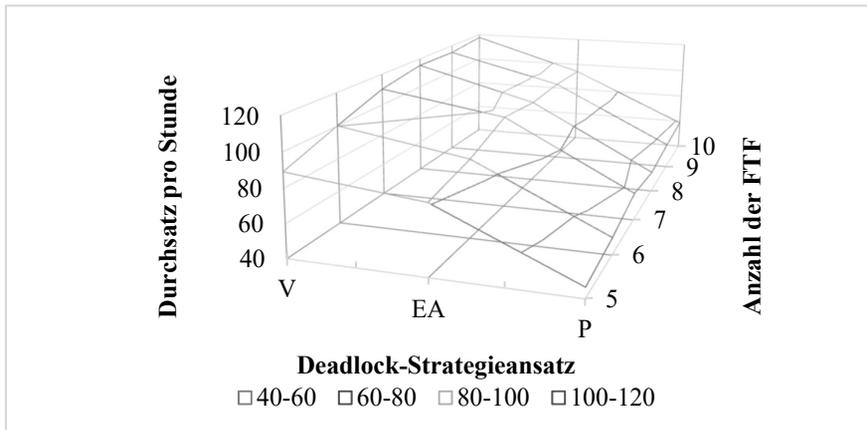


Abbildung 4: Oberflächendiagramm des Durchsatzes pro Stunde bei geringer Verfügbarkeit und hoher MTTR (Durchschnitt der 10 Simulationdurchläufe)

Bei wenigen, langen Störungen erzielen Erkennen & Auflösen und der Vermeidungsansatz jedoch wieder nahezu identische Durchsätze (eine Präventionsstrategie schneidet deutlich schlechter ab). In der durchgeführten Experimentreihe hat die Veränderung der Verfügbarkeit einen höheren Einfluss auf den Durchsatz als die Veränderung der MTTR. Dies hängt mit den gewählten Werten für die Stellgrößen zusammen und kann nicht verallgemeinert werden.

Hinsichtlich der Wegstrecke ergibt sich für die Strategieansätze ein ähnliches Bild. Die durchschnittliche Wegstrecke pro Transport ist bei einem Präventionsansatz für alle Störungsprofile länger. Zwischen Erkennen & Auflösen und Vermeidung liegen bei allen Störungsprofilen und variierter Fahrzeuganzahl keine signifikanten Unterschiede vor. Der wesentliche Einflussfaktor auf die durchschnittliche Wegstrecke pro Transportauftrag ist die Fahrzeuganzahl, nicht die Beschaffenheit der Störungen.

Bei der Betrachtung der durchschnittlichen Wartezeit pro Transportauftrag spielt das Störungsprofil dagegen eine wesentliche Rolle. Bei wenigen, kurzen Störungen weist Erkennen & Auflösen die kürzesten Wartezeiten auf, wie in Abbildung 5 zu sehen ist. Der Vermeidungsansatz erzeugt dagegen ab einer FTF-Anzahl von sechs die längste durchschnittliche Wartezeit pro Transportauftrag.

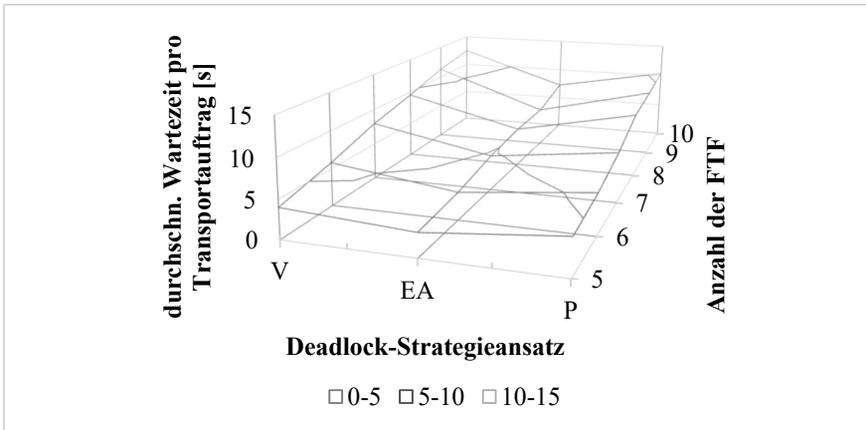


Abbildung 5: Oberflächendiagramm der durchschnittlichen Wartezeit pro Transportauftrag bei hoher Verfügbarkeit und geringer MTTR (Durchschnitt der 10 Simulationdurchläufe)

Nehmen die Störungen jedoch in Häufigkeit und Stärke zu, profitiert am meisten der Vermeidungsansatz. Erkennen & Auflösen und der Präventionsansatz erhöhen ihre mittlere Wartezeit stärker. Beim Präventionsansatz und hoher Fahrzeuganzahl entsteht ein erheblicher Anstieg. Abbildung 6 macht diesen Zusammenhang deutlich.

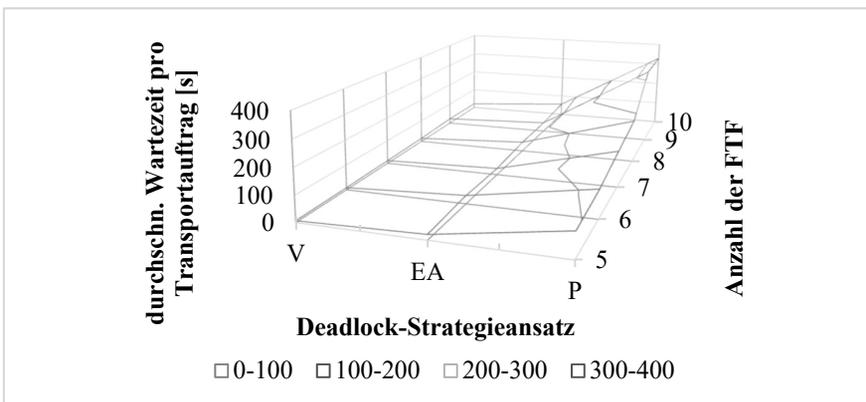


Abbildung 6: Oberflächendiagramm der durchschnittlichen Wartezeit pro Transportauftrag bei geringer Verfügbarkeit und hoher MTTR (Durchschnitt der 10 Simulationdurchläufe)

Der wesentliche Einflussfaktor auf die Wartezeit ist die Verfügbarkeit, nicht die MTTR. Das heißt in anderen Worten, dass die Häufigkeit der Störungen entscheidend ist für die Entstehung von Wartezeiten bei den Fahrzeugen.

Lässt man Deadlocks zu, indem ein Erkennen-und-Auflösen-Ansatz gewählt wird, ist der häufigste Deadlock-Fall 3. Dessen Anteil sinkt jedoch mit höherer Verfügbarkeit.

Die Gesamtanzahl der Deadlocks steigt sowohl mit der Anzahl der Fahrzeuge als auch mit einer höheren Verfügbarkeit. Die MTTR hat nur bei höherer Fahrzeuganzahl einen sichtbaren Einfluss auf die Anzahl der Deadlocks.

5 Fazit

Störungen können einen wesentlichen Einfluss auf logistische Kennzahlen haben, die entscheidend sind, um den optimalen Strategieansatz zum Umgang mit Deadlocks zu wählen. Grundsätzlich ließ sich erneut zeigen, dass mit einem Präventionsansatz die Deadlock-Problematik leicht umgangen werden kann, aber auch bei störanfälligen Systemen geht dies deutlich zu Lasten der Performance (Durchsatz, durchschnittliche Wege und Zeit pro Transportauftrag). Bei einem geringen Störaufkommen kann zwischen einem Ansatz der Vermeidung und Erkennen & Auflösen keine dominante Strategie festgestellt werden. Handelt es sich jedoch um ein sehr störanfälliges System (Verfügbarkeit $\leq 95\%$), zeigt die Experimentreihe, dass der Vermeidungsansatz sich hinsichtlich der betrachteten Kennzahlen am besten durchsetzt.

Dennoch ist diese Induktion kritisch zu betrachten. Zum einen liegt bei der Umsetzung der Strategien, gerade bei Vermeidung sowie Erkennen und Auflösen, eine spezifische technische Implementierung zugrunde, die in Detailfragen unterschiedlich umgesetzt werden kann und dadurch auch zu anderen Ergebnissen führen könnte. Zum anderen wurden die Störungen nur an einer Layout-Variante untersucht, die alle drei Deadlock-Strategieansätze erlaubt. Eine andere Systemstruktur könnte hierbei ebenfalls andere Ergebnisse liefern. Weitere Experimente, insbesondere mit anderen Simulationsmodellen für abgewandelte Anwendungsfälle in deadlockfähigen und störungsanfälligen Systemen, könnten die Aussagekraft und die Verallgemeinerung der Erkenntnisse dieser Publikation erhöhen.

Literaturverzeichnis

- Coffman, E.G.; Elphick, M.; Shoshani, A.: System Deadlocks. ACM Computing Surveys 3 (1971) 2, S. 67–78.
- Forget, P.; D'Amours, S.; Frayret, J.-M.; Gaudreault, J.: Study of the performance of multi-behaviour agents for supply chain planning. Computers in Industry 60 (2009) 9, S. 698–708.
- Gudehus, T.: Logistik: Grundlagen - Strategien - Anwendungen. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg 2011.
- Kim, K.H.; Jeon, S.M.; Ryu, K.R.: Deadlock prevention for automated guided vehicles in automated container terminals. OR Spectrum 28 (2006) 4, S. 659–679.
- Kohl, A.-K.; Pfrezschner, F.: Logistikmonitor 2018: Ergebnisse einer Expertenbefragung von Statista und der Bundesvereinigung Logistik (BVL) e.V., 2018, zuletzt geprüft am 23.10.2019.
- Kuhn, A.; Bernemann, S. (Hrsg.): Prozessketten in der Logistik: Entwicklungstrends und Umsetzungsstrategien. Dortmund: Verl. Praxiswissen 1995.
- Lienert, T.; Fottner, J.: No More Deadlocks – Applying The Time Window Routing Method To Shuttle Systems. In: Paprika, Z.Zoltay; Horák, P.; Váradi, K.;

- Zwierzcyk, P.Tamás; Vidovics-Dancs, Á.; Rádics, J.Péter (Hrsg.): ECMS 2017 Proceedings edited by Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierzcyk, Ágnes Vidovics-Dancs, János Péter Rádics, May 23 - May 26, 2017, 2017, S. 169–175.
- Lu, Z.; Zhuang, Z.; Huang, Z.; Qin, W.: A Framework of Multi-Agent Based Intelligent Production Logistics System. *Procedia CIRP* 83 (2019), S. 557–562.
- März, L.; Krug, W.; Rose, O.; Weigert, G.: *Simulation und Optimierung in Produktion und Logistik*. Berlin, Heidelberg: Springer Berlin Heidelberg 2011.
- Mayer, S.; Furmans, K.: Deadlock prevention in a completely decentralized controlled materials flow systems. *Logistics Research* 2 (2010) 3–4, S. 147–158.
- Mors, A.W.: *The world according to MARP: Multi-Agent Route Planning*. Delft: Technische Universiteit Delft 2010.
- Müller, M.; Schmidt, S.; Reggelin, T.: Deadlock and Collision Handling for Automated Rail-Based Storage and Retrieval Units. In: Mustafee, N.; Bae, K.-H.G.; Lazarova-Molnar, S.; Rabe, M.; Szabo, C.; Haas, P.; Son, Y.-J. (Hrsg.): *Proceedings of the 2019 Winter Simulation Conference*, National Harbor, MD, USA, 08.12.2019 - 11.12.2019, 2019, S. 1591–1601.
- Müller, M.; Ulrich, J.H.; Reyes-Rubiano, L.S.; Reggelin, T.; Lang, S.: Comparison of Deadlock Handling Strategies for different Warehouse Layouts with an AGVS. In: Bae, K.-H.G.; Feng, B.; Kim, S.; Lazarova-Molnar, S.; Zheng, Z.; Roeder, T.; Thiesing, R. (Hrsg.): *Proceedings of the 2020 Winter Simulation Conference*, 14.12.2020 - 18.12.2020, 2020, S. 1300–1311.
- Schenk, M.; Wirth, S.; Müller, E.: *Fabrikplanung und Fabrikbetrieb*. Berlin, Heidelberg: Springer Berlin Heidelberg 2014.
- Schönung, F.; Mayer, S.; Berbig, D.: Dezentrale, deadlockfreie Steuerung eines Stetigförderers. *at - Automatisierungstechnik* 59 (2011) 4, S. 73.
- Seibold, Z.: *Logical time for decentralized control of material handling systems*. Karlsruhe, Karlsruher Institut für Technologie (KIT), Dissertation, 2016.
- Tempelmeier, H. (Hg.): *Planung logistischer Systeme*. Berlin: Springer Vieweg 2018.
- van Bonn, B.: Basisdaten der Logistikplanung. In: Clausen, U.; Geiger, C. (Hrsg.): *Verkehrs- und Transportlogistik*. Berlin, Heidelberg: Springer Berlin Heidelberg 2013, S. 291–298.
- VDI-Gesellschaft Fördertechnik Materialfluss Logistik VDI-Richtlinie 3590: *Kommissioniersysteme*. Berlin: Beuth Verlag GmbH, 2002.
- VDI-Gesellschaft Produktionstechnik (ADB) VDI-Richtlinie 5200: *Fabrikplanung*. Berlin: Beuth Verlag GmbH, 2009.
- Yalcin, A.: *Multi-Agent Route Planning in Grid-Based Storage Systems*. Frankfurt (Oder), Europa-Universität Viadrina Frankfurt (Oder), Wirtschaftswissenschaftliche Fakultät, Dissertation, 2017.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Simulationsbasierte Untersuchung der Grenzproduktivität von Robotern in einem AutoStore-Lagersystem

***Simulation-based analysis of the marginal productivity
of robots in an AutoStore storage system.***

Stefan Galka, OTH Regensburg, Regensburg (Germany),
stefan.galka@oth-regensburg.de

Christoph Scherbarth, ehem. OTH Regensburg, Regensburg (Germany),
cscherbarth@gmx.de

Abstract: AutoStore is a storage system for automated warehousing of small parts with low space requirements and picking according to the “goods-to-person” principle. The performance of the system depends on factors such as the number of robots and ports as well as the order structure. Due to the numerous interactions, the influence of the individual factors on the system performance can only be investigated using simulation studies. In this paper, the effects of selected factors on the system performance are investigated using a discrete-event simulation model. At the beginning of the paper, the system elements and processes are described. Afterwards, the correlation between the elements and the simulation model will be explained more in detail. Finally, the results of the simulation study are discussed.

1 Einführung

Die Firma Hatteland aus Norwegen entwickelte in den 90-er Jahren des letzten Jahrhunderts ein neues Lagersystem, um es im eigenen Unternehmen einzusetzen. Regale und Gassen wurden im Lager eliminiert und stattdessen die Ladungsträger aufeinandergestellt. Die automatisierte Ein- und Auslagerung übernahmen kleine Roboter. Inzwischen wird das AutoStore-System am Markt angeboten und verzeichnete in den letzten Jahren eine wachsende Beliebtheit. Bisher wurden weltweit mehr als 500 Systeme realisiert (Handling.de 2020). In einer Umfrage unter AutoStore-Nutzern wurden die vergleichsweise hohe Lagerdichte und die einfache Skalierbarkeit des Systems als Vorteile hervorgehoben (Galka et al. 2020). Allerdings ist das AutoStore-System nicht für jeden Anwendungsfall geeignet. Es handelt sich um ein Lagersystem für Kleinteile, was das zu lagernde Sortiment hinsichtlich des Gewichts und der Abmessungen reglementiert.



1.1 Bestandteile des AutoStore-Systems

Ein AutoStore-System setzt sich im Wesentlichen aus fünf Komponenten zusammen. Dazu gehören das „Grid“, welches mit den „Bins“ befüllt wird und die „Roboter“, welche die Bins aus dem Grid holen und an die „Ports“ für die Ein- und Auslagerung befördern. Das Ganze wird durch den „Controller“ orchestriert und gesteuert. In der Abbildung 1 ist ein AutoStore-System dargestellt.

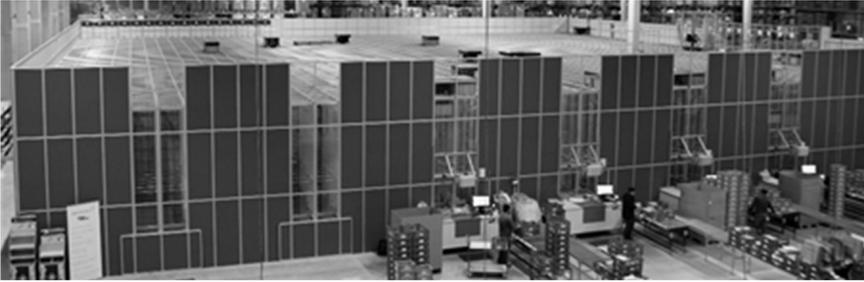


Abbildung 1: AutoStore-System bei einem Logistikdienstleister (Foto: Autor)

Das Grid wird mit den Bins befüllt, die als Lagerhilfsmittel für die zu lagernden Artikel genutzt werden. Die Grundfläche der Bins ist immer gleich ($B=449$ mm/ $L=649$ mm). Hinsichtlich der Höhe gibt es unterschiedliche Bins. Da die gesamte Höhe des Autostore-Systems auf 5,40 Meter beschränkt ist, definiert die Höhe der Bins gleichzeitig die maximale Anzahl an Bins, die übereinandergestellt werden können. Die Bins werden von den Robotern gehandhabt. Aktuell werden zwei unterschiedliche Versionen des Roboters angeboten. Die sogenannte „Redline“ (R5) stellt die Basisversion der Roboter dar. Die zweite Version des Roboters wird als „Blackline“ (B1) bezeichnet, die etwas höhere Geschwindigkeiten ermöglicht.

Die vierte wesentliche Komponente eines AutoStore-Systems ist der Port. Die Ports sind die Schnittstelle zu Mitarbeitern, die die Artikel aus den Bins entnehmen oder Artikel in die Bins legen. Aktuell gibt es vier unterschiedliche Porttypen. Diese unterscheiden sich hinsichtlich des maximal möglichen Durchsatzes.

1.2 Funktionsweise des AutoStore-Systems

Das AutoStore-System wird in der Regel für die Kommissionierung verwendet. Es handelt sich dabei um eine Ware-zur-Person Kommissionierung (WzP). Wesentlich für die WzP-Kommissionierung ist die Bereitstellung von Artikeln am Arbeitsplatz des Kommissionierers (Port). Dieser Vorgang wird auch als Behälterpräsentation bezeichnet. Nach der Entnahme durch den Kommissionierer muss der Bin wieder rückgelagert werden. Dies erfolgt auch, wenn der Bin leer ist. Leere Bins verbleiben im System und werden bei Bedarf für einen Wareneingang am Port bereitgestellt. Damit muss hinsichtlich des Ablaufs nicht zwischen Wareneingang und Kommissionierung (Warenausgang) unterschieden werden. In beiden Fällen muss der Roboter einen Bin am Port bereitstellen (Auslagerung) und diesen dann später wieder abholen, wie auch zurück ins Grid stellen (Einlagerung).

Im Folgenden wird kurz und vereinfacht der Prozess beschrieben. Eine grafische Darstellung des Vorgangs findet sich in Abb. 2. Die Erläuterungen zur Abbildung sind im Text kursiv dargestellt. Der Controller erhält aus dem überlagerten System den

Auftrag, einen bestimmten Bin auszulagern. Der Controller ermittelt die Position des Bins im Grid (*Im Beispiel muss der grüne Bin ausgelagert werden*). Im besten Fall kann ein Roboter direkt auf den Bin zugreifen, falls sich keine anderen Bins im Stapel über dem benötigten Bin befinden. Ist dies nicht der Fall, erzeugt der Controller Umlageraufträge für die Bins über dem gewünschten Bin. Dies wird auch als „Ausgraben“ bezeichnet. Die „störenden“ Bins werden von einem oder mehreren Robotern umgelagert. Dazu werden die Bins aus dem Schacht entnommen und in einem anderen Schacht abgestellt. (*siehe Abb. 2; „Ausgraben“ beendet*.) Ist der Zugriff auf den benötigten Bin möglich, wird ein Roboter mit dessen Abholung beauftragt. Der Roboter transportiert den Bin dann direkt zum Port oder wenn der Bin dort noch nicht benötigt wird, wird dieser wie die umgelagerten Bins auf einen anderen Stapel (oberste Position) gestellt und später zum Port gebracht (dies wird „Vorbereitung“ genannt). Die umgelagerten Bins verbleiben nicht an der neuen Position. Der Controller beauftragt einen Roboter mit der Rücklagerung dieser Bins. Dabei wird auf die Reihenfolge geachtet. Die Bins werden in der gleichen Reihenfolge in den Schacht wieder eingelagert. (*siehe Abb.2; „Endzustand“*)

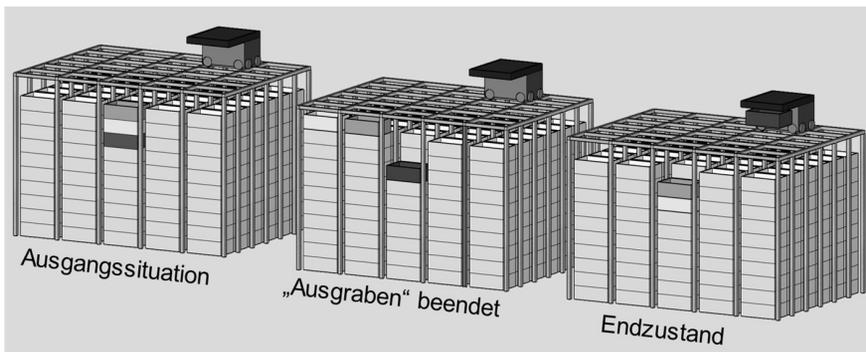


Abbildung 2: Ablauf der Auslagerung eines Bins (grün)

Dies ist eine der Grundphilosophien des AutoStore-Systems. Nur der Zugriff an einem Port verändert die Position eines Bins im Schacht. Damit wird automatisch eine ABC-Zonierung erreicht, da Artikel mit einer hohen Zugriffshäufigkeit tendenziell weiter oben im Schacht gelagert werden und damit der Aufwand für die Auslagerung von A-Artikel gering ist. In Abhängigkeit der Zugriffsverteilung über das Sortiment liegt die mittlere Grabtiefe oft nur bei 1,5-2,5 Bins. Die mittlere Grabtiefe gibt an, wie viele Bins im Mittel für eine Auslagerung umgelagert werden müssen, damit der Ziel-Bin durch einen Roboter erreicht werden kann.

1.3 Planung und Dimensionierung eines AutoStore-Systems

Die Dimensionierung eines AutoStore-Systems ist eine komplexe Fragestellung, da viele unterschiedliche Aspekte die Leistungsfähigkeit der gesamten Anlage beeinflussen. In der Abb. 3 sind die Gestaltungsparameter aufgeführt. Ausgewählte Gesichtspunkte werden im Anschluss diskutiert.

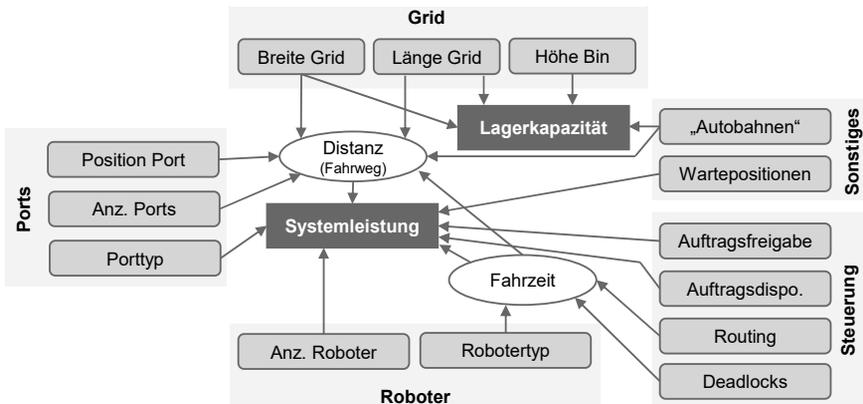


Abbildung 3: Gestaltungsmöglichkeiten und deren Wechselwirkungen

Die Kubatur des Grids kann flexibel gestaltet und dem Gebäude angepasst werden. In der Regel entspricht die Kubatur einem Quader und es wird die maximale Höhe des Grids ausgenutzt. Damit sind die Länge und die Breite des Grids die wesentlichen Gestaltungsparameter, die gemeinsam mit den eingesetzten Bins (Höhe) die Lagerkapazität definieren. Im Zusammenspiel mit der Position und Anzahl der Ports beeinflussen die Gridabmessungen die Distanz, die die Roboter für die Aus- und Einlagerungen zurückzulegen haben.

Die Distanz beeinflusst nicht alleine die Fahrzeit. Neben dem Robotertyp mit seinen spezifischen kinematischen Eigenschaften ist die gewählte Route entscheidend. Da die Roboter bei jeder Richtungsänderung zum Stillstand kommen und anschließend wieder beschleunigen müssen, ist ein häufiger Richtungswechsel nicht vorteilhaft. Zusätzlich müssen bei der Wahl der Route auch der Fahrweg und die Arbeitsposition von anderen Robotern beachtet werden. Um das Routing zu vereinfachen und Behinderungen zwischen den Robotern zu vermeiden, werden häufig Bereiche des Grids für die Lagerung von Bins gesperrt. Solche Bereiche werden als „Autobahn“ bezeichnet, da hier die Roboter ungestört fahren können. Die Nutzung von Autobahnen steigert die Leistungsfähigkeit des Systems, reduziert aber die nutzbare Lagerkapazität des Grids.

Die Auswahl des eingesetzten Porttyps wird im Wesentlichen von den Durchsatzanforderungen am Port beeinflusst. Ports mit einem höheren Durchsatz müssen von mehr Robotern pro Zeiteinheit versorgt werden, was zu einer höheren Verkehrsdichte um den Port führt. Roboter können rechts und links neben der Abgabeposition(en) warten. Liegen die Ports zu eng beieinander, dann sind nicht ausreichend Wartepositionen verfügbar, was leistungsmindernd wirken kann. Die Prozesszeit für die Abgabe der Bins am Port ist abhängig vom Hub- bzw. Senkweg. Ports müssen nicht wie in Abb. 1 dargestellt am Boden positioniert sein. Durch den Einsatz von Bühnen kann der Port angehoben werden, was den Hub- und Senkweg des Roboters reduziert und damit die Abgabe-/Aufnahmezeiten für die Bins am Port mindert.

Neben den dargestellten Parametern, die im Rahmen der Planung definiert werden können, gibt es weitere Aspekte, die die Systemleistung beeinflussen, aber nicht direkt verändert werden können. Wesentlich dabei ist die Verteilung der Zugriffshäufigkeit

über das Sortiment. AutoStore ist geeignet, wenn die Zugriffsverteilung dem Paretoprinzip folgt, da dies zu einer geringen mittleren Grabtiefe führt.

Aspekte wie die Anzahl an Positionen je Auftrag im Zusammenspiel mit der Kommissionierstrategie haben auch Einfluss auf die Leistung, werden an dieser Stelle aber nicht diskutiert.

2 Stand der Wissenschaft

In wissenschaftlichen Veröffentlichungen spielt das AutoStore-System bisher eine untergeordnete Rolle. Bipan Zoua et al. untersuchen in ihrem Beitrag den Einfluss von Lagerbelegungsstrategien auf die Leistung eines AutoStore-Systems. Dabei entwickeln die Autoren neue Ansätze, die sich von den in der Praxis genutzten Strategien unterscheiden. Das System wird mithilfe der Warteschlangentheorie modelliert und untersucht. Hierzu wurden starke Vereinfachungen vorgenommen. So werden das Routing der Roboter und das Deadlock-Handling nicht berücksichtigt. (Zoua et al. 2016)

In der Veröffentlichung „Redesign of the AutoStore order processing line: A multi-scenario discrete-event simulation study“ von Tjeerdsma (Tjeerdsma 2019) wird die Auftragsabwicklung in einem Fulfillment Center der niederländischen Post betrachtet. Dabei geht es nicht direkt um die Leistung des AutoStore-Systems. Vielmehr untersucht Tjeerdsma Ansätze, wie das Sortiment zwischen AutoStore und anderen Lagerbereichen aufgeteilt werden sollte und welchen Einfluss die Aufteilung auf die Konsolidierungsaufwände hat.

In einer weiteren Veröffentlichung untersucht Odland Håvås die technische Verfügbarkeit von AutoStore-Systemen und mögliche Vorteile eines Predictive Maintenance Ansatzes für ein AutoStore-System. Die eigentliche Systemleistung spielt bei dieser Untersuchung keine Rolle. (Odland 2018)

Lienert (2018) beschäftigt sich in seinem Beitrag „Simulation-based Performance Analysis in Robotic Mobile Fulfillment Systems“ mit einem Ware-zur-Person-Kommissioniersystem, bei welchem durch mobile Roboter (Fahrerlose-Transportfahrzeuge) Ladungsträger zu Pickstationen transportiert werden. Bestimmte Aspekte lassen sich auf AutoStore-System übertragen. Allerdings unterscheidet sich das Routing in einem gassenbasierten System deutlich von den Möglichkeiten auf einem „Grid“, bei dem zu jeder Zeit die Möglichkeit eines Richtungswechsels gegeben ist. Gleiches gilt auch für das Deadlock-Handling. Allerdings lassen sich teilweise Ansätze von Lienert et al. auf die Modellierung eines AutoStore-Systems übertragen. (Lienert 2021)

Andere Beiträge wie von Deatcu et al. (2018) behandeln Shuttle-basierte Lagersysteme und nutzen die Simulation für die Untersuchung dieser Technik. Auch diese Systeme unterscheiden sich im Verhalten und den Betriebsstrategien vom AutoStore-System, was sich wieder mit den Möglichkeiten des Routings begründen lässt.

Zusammenfassend lässt sich feststellen, dass sich bisher noch keine Veröffentlichung mit der simulationsbasierten Abbildung eines AutoStore-Systems befasst hat. Durch die Entwicklung eines herstellerneutralen Simulationsmodells für ein AutoStore-System und die Durchführung von Parameterstudien sollen Gestaltungshinweise für

Planer und Nutzer erarbeitet werden. Damit soll das asymmetrisch verteilte Wissen zwischen Planern/Nutzern und Herstellern/Integratoren angeglichen werden.

3 Abbildung des Systems in einem Simulationsmodell

Für die Untersuchung wurde ein parametrierbares Simulationsmodell in Plant Simulation eingesetzt. Neben den in Abb. 3 aufgezeigten Aspekten lassen sich weitere Parameter wie die Prozesszeiten am Port und Vorgaben für die Auftragsgenerierung festlegen. Im Folgenden werden ausgewählte Aspekte der Simulation beschrieben.

Bei der Initialisierung des Modells wird der Füllgrad der Schächte bestimmt, da nicht alle Schächte mit der gleichen Anzahl an Bins befüllt werden (entsprechend des vorgegebenen Füllgrades). Eine Bestandsverwaltung wurde nicht implementiert, vielmehr wird durch den Auftragsgenerator zufällig ein Schacht für den auszulagernden Bin ausgewählt. Dies entspricht einer Gleichverteilung der Zugriffshäufigkeit über alle Schächte. Die Position des Bins innerhalb des Schachtes ist abhängig von der festgelegten Höhenverteilung und wird durch den Auftragsgenerator unter Berücksichtigung der Anzahl von Bins im ausgewählten Schacht bestimmt. Bei der Festlegung des Einlagerschachtes wird der Füllstand des Schachtes beachtet. Das Auftragsmanagement bevorzugt Schächte mit einer geringeren Anzahl an Bins.

Nach dem Abschluss eines Auftrages durch den Roboter erfolgt die Zuweisung eines neuen Auftrages. Für die Auslageraufträge ist neben dem Ausgangspunkt (Schacht und Position im Schacht) auch das Ziel (Port) bekannt. Zuerst wird geprüft, wie viele Aufträge für den Ziel-Port bereits von anderen Robotern bearbeitet werden. Ist die maximale Anzahl bereits erreicht, besteht die Möglichkeit, den Auftrag vorzubereiten. Dabei wird der Bin ausgegraben und in der Nähe des Ports abgestellt. Auch die Anzahl der vorzubereitenden Aufträge ist begrenzt. Wenn auch die maximale Anzahl der vorzubereitenden Aufträge erreicht wurde, wird der Auftrag zurückgestellt und ein Auslagerauftrag für einen anderen Port überprüft. Sollte kein Auftrag verfügbar sein, wird der Roboter auf einer Ladestation geparkt.

Muss der Bin ausgegraben werden, werden durch das Auftragsmanagement Umlageraufträge erzeugt. Der eigentliche Auslagerauftrag wird zurückgestellt und es werden zuerst die Umlageraufträge dem Roboter zugewiesen. Nachdem der Roboter die Auslagerung vorgenommen hat, werden Umlageraufträge für die Rücklagerung der umgelagerten Bins erstellt, die einem freien Roboter zugewiesen werden.

Die Roboter bewegen sich auf der obersten Ebene des Grids und können an jeder Position des Grids die Richtung wechseln. Dazu müssen die Roboter zum Stillstand kommen und einen Umsetzungsvorgang ausführen. Weiterhin können Bins im Grid so abgestellt werden, dass diese zwar vom Roboter aufgenommen werden können, aber der Roboter an dieser Stelle nicht mehr über das Grid fahren kann. Im Rahmen der Wegeplanung wird mithilfe eines angepassten A*-Algorithmus der Weg mit der geringsten erwarteten Fahrzeit bestimmt. Bei jedem Aufruf der Wegefindung wird ein gerichteter Graph erstellt (je Schacht ein Knoten), dessen Kantenbewertung abhängig vom aktuellen Systemzustand ist. Abb. 4 zeigt ein Beispiel und die Kennwerte, die zur Bestimmung der Kantenbewertung herangezogen werden.

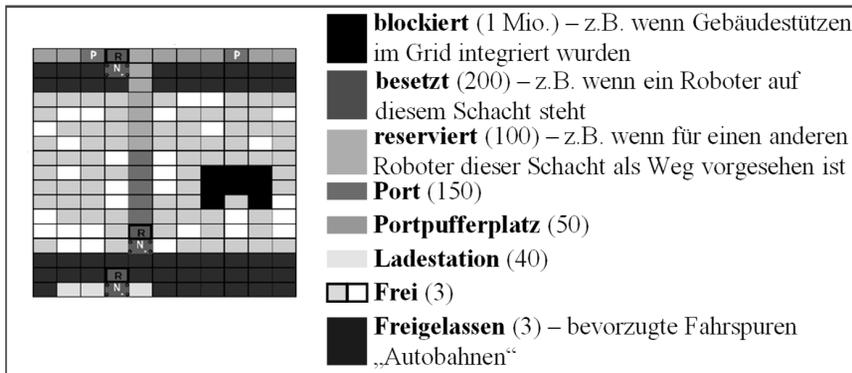


Abbildung 4: Visualisierung im Modell (links) und Kantenbewertungen in Abhängigkeit vom aktuellen Systemzustand (Bewertungszahl in Klammern)

Aufgrund der Berechnungszeit für die Wegfindung bei großen Grid-Abmessungen wurde diese Funktion als externes Modul in der Programmiersprache C++ implementiert.

Die Steuerungsstrategien wurden aufgrund von Beobachtungen realer AutoStore-Systeme entwickelt, da keine Informationen von AutoStore zu den verwendeten Algorithmen verfügbar sind. Das Modell wurde anhand von Bewegungsdaten einer realen Anlage validiert. Die Simulationsergebnisse haben sich um ca. 12 % von den Kennwerten der realen Anlage unterschieden.

4 Untersuchung der Grenzproduktivität

Mithilfe der Simulationsstudie soll untersucht werden, welchen Einfluss die Anzahl der eingesetzten Roboter auf die Systemleistung hat. Es ist davon auszugehen, dass bei zunehmender Anzahl an Robotern auf gleicher Fläche (Grid) mehr Behinderungen zwischen den Robotern auftreten und dies zu einem abnehmenden „Grenznutzen“ von zusätzlichen Robotern im System führt. Im Abschnitt 4.1 wird aufgezeigt, welche Experimente für diese Studie durchgeführt wurden. Die Ergebnisse und daraus resultierende Schlussfolgerungen werden im Abschnitt 4.2 dargestellt.

4.1 Experimentendesign

Im Rahmen der Simulationsstudie wurden 225 Experimente durchgeführt, wobei im Normalfall 5 Replikationen pro Experiment simuliert wurden. Die Experimente unterscheiden sich durch die Anzahl der eingesetzten Ports und Roboter. Da bekannt ist, dass die Verteilung der Zugriffe im Stapel einen wesentlichen Einfluss auf die Leistungsfähigkeit des Systems hat, wurden zusätzlich vier unterschiedliche Verteilungen der Zugriffe über den Stapel untersucht. In der folgenden Abb. 5 werden die variierten Parameter (dunkelgrau) und die konstanten Parameter (helles grau) zusammengefasst. Weiterhin sind die genutzten Zugriffsverteilungen in einem Diagramm abgebildet.

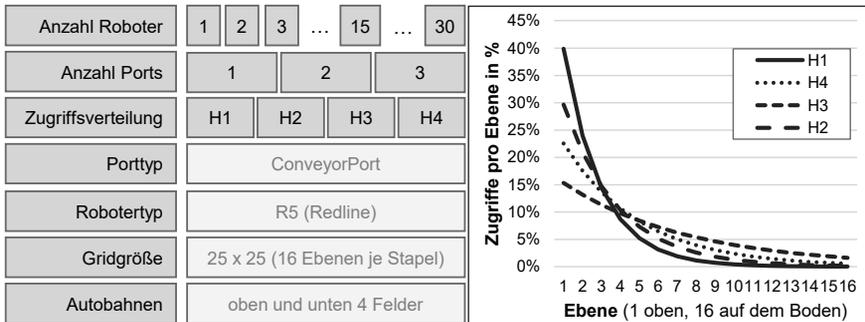


Abbildung 5: Übersicht über die variierten Parameter in den Experimenten

Die Prozesszeit für die Entnahme bzw. das Befüllen der Bins am Port beträgt im Mittel 10 Sekunden mit einer Standardabweichung von 2 Sekunden. Für die Experimente stehen ausreichend Aufträge zur Verfügung, sodass die Roboter nie auf einen neuen Auftrag warten müssen.

Für alle Höhenverteilungen wurde eine Roboteranzahl von eins bis 15 untersucht. Ergänzend wurden für die Höhenverteilung H1 weitere Experimente mit einer Roboteranzahl bis 30 durchgeführt.

4.2 Interpretation der Simulationsergebnisse

Die dargestellten Simulationsergebnisse basieren auf Mittelwerten, die aus den Kennwerten der einzelnen Replikationen berechnet wurden.

In Tabelle 1 ist die Systemleistung in Behälterpräsentationen pro Stunde dargestellt. Es ist zu erkennen, dass die Systemleistung höher ist, wenn die mittlere zu grabende Tiefe kleiner ist, da pro Auslagervorgang weniger Bins umgelagert werden müssen.

Tabelle 1: Systemleistung in Aufträgen (Bins)/Stunde und mittlere Grabetiefe

Anzahl Ports	Höhenverteilung H1			Höhenverteilung H2			Höhenverteilung H3			Höhenverteilung H4		
	1	2	3	1	2	3	1	2	3	1	2	3
1 Roboter	39	41	42	29	32	34	23	24	26	17	18	29
2 Roboter	73	81	85	56	61	65	45	47	49	36	35	56
5 Roboter	147	171	186	113	132	139	88	102	107	75	80	113
10 Roboter	198	281	306	160	214	232	136	159	178	119	133	160
15 Roboter	189	345	397	177	263	297	163	204	226	159	172	177
Grabetiefe	2,50			3,31			4,24			5,55		

In Abhängigkeit der Anzahl Ports und der Höhenverteilung sinkt die Systemleistung ab einer bestimmten Anzahl von Robotern im System. Dieser Effekt lässt sich auf zwei Ursachen zurückführen. Zum einen nehmen die Behinderungen und Wartezeiten am Port bei zunehmender Anzahl an Robotern zu, zum anderen werden keine weiteren Aufträge freigegeben, wenn die maximale Anzahl an Aufträgen für die Auslagerung bzw. Vorbereitung erreicht wurde. Ohne diese Vorgabe steigt der Anteil der Blockierungen bzw. Wegzeiten (durch mehr Richtungswechsel) an. Im Diagramm in Abb. 6 sind die Zeitanteile für eine Systemkonfiguration mit einem Port und der Höhenverteilung H1 dargestellt. Der Anteil der Fahrzeiten und Handhabungszeiten

nimmt mit steigender Anzahl an Robotern sukzessive ab, hingegen steigt der Zeitanteil für das Warten und Laden (wenn kein Auftrag verfügbar ist, wird der Roboter zum Laden geschickt). Ähnliche Verläufe zeigen sich auch bei den anderen Experimenten.

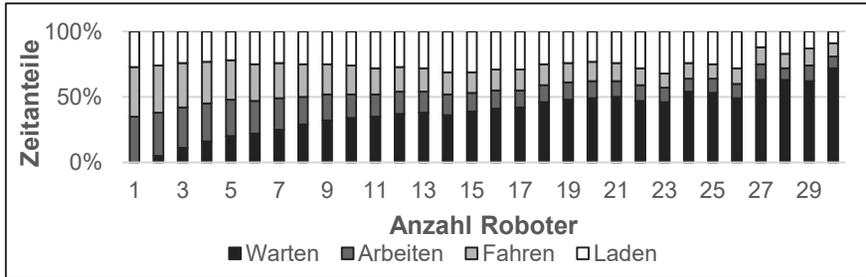


Abbildung 6: Zeitanteile für die Experimente mit 1 Port (H1)

Abschließend soll die Grenzproduktivität aufgezeigt werden. Darunter wird die zusätzliche Systemleistung verstanden, die durch den Einsatz eines weiteren Roboters erzielt werden kann. In der Abb. 7 ist der Verlauf der Grenzproduktivität für die Experimente mit der Zugriffsverteilung H1 visualisiert.

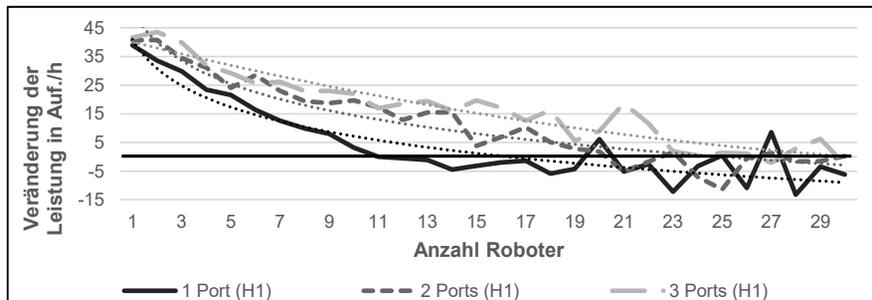


Abbildung 7: Grenzproduktivität für die Experimentreihe der Höhenverteilung H1

Die Simulationsergebnisse haben gezeigt, dass die Zugriffsverteilung, die Anzahl der Ports und die Anzahl der Roboter maßgeblich die Leistung eines AutoStore-Systems beeinflussen. Beispielsweise sollten im untersuchten Beispiel (H1, 1 Port) nicht mehr als 11 Roboter eingesetzt werden, da durch den Einsatz weiterer Roboter die Systemleistung reduziert wird. Diese Kennzahl ist für jedes untersuchte Beispiel unterschiedlich und eine Verallgemeinerung dieser Kennzahl ist nicht möglich. Neben der Zugriffsverteilung wird dieser Kennwert auch durch die Gridgröße beeinflusst, die nicht variiert wurde. Auch wird die Binhöhe (Anzahl Bins pro Stapel) Einfluss auf diesen Kennwert haben.

5 Zusammenfassung und Ausblick

Die Leistungsfähigkeit eines AutoStore-Systems wird von vielen Faktoren beeinflusst. Einige dieser Faktoren sind durch den Anwendungsfall vorgegeben und können durch die Systemgestaltung nicht verändert werden. Grundsätzlich sollte im Rahmen der Planung die Frage beantwortet werden, ob das AutoStore-Lagersystem eine geeignete Technologie ist. Neben den Arteikeigenschaften (Abmessung und Gewicht) ist vor allem die Zugriffsverteilung über das Sortiment entscheidend für die Eignung des AutoStore-Systems. Andere Aspekte wie die Anzahl an Ports und Roboter können im Rahmen der Planung definiert werden. Analytisch lässt sich die Leistungsfähigkeit des Systems nur abschätzen, fundierte Aussagen können mit Simulationsstudien abgeleitet werden. Die durchgeführten Experimente haben gezeigt, dass die Anzahl der Roboter auf die Anzahl der Ports und die Gridgröße abgestimmt sein muss. Vor allem bei einer geringen Anzahl an Ports kann eine hohe Anzahl an Robotern die Gesamtsystemleistung verschlechtern. Es wird empfohlen, ein AutoStore-System im Rahmen der Planung zu simulieren, um die Wechselwirkungen der Systemelemente im ausreichenden Maße zu berücksichtigen.

Eine durchgeführte Befragung von AutoStore-Nutzern hat darüber hinaus ergeben, dass die Prozesszeiten am Port Beachtung finden sollten. Weiterhin haben viele Nutzer berichtet, dass sich Anforderungen (Aufträge) über den Tag verändern und dezidiert untersucht werden sollten. (Galka 2020)

Danksagung

Das Simulationsmodell wurde im Rahmen einer Abschlussarbeit in Kooperation mit der Miebach Consulting GmbH entwickelt. An dieser Stelle sei Dr. Jens Zimmermann und Markus Nadj für die fachliche Betreuung und die eingebrachte Expertise gedankt.

Literatur

- Deatcu C., Schramm T., Zobel K.: Simulationsmodell eines Shuttle-Systems zur Ermittlung des Systemverhaltens, Tagungsband ASIM 2018, Hamburg
- Galka S., Troesch L., Scherbarth C.: AutoStore- Was Nutzer über das System berichten können. OTH Regensburg, 2020
- Lienert T., Staab T., Ludwig C., Fottner J.: Simulation-based Performance Analysis in Robotic Mobile Fulfilment Systems: Analyzing the Throughput of Different Layout Configurations. München, 2018.
- Lienert, T.: Methodik zur simulationsbasierten Durchsatzanalyse FTF-basierter Kommissioniersysteme. Dissertation, TU München, München. Vorab. überl.2021.
- Odlan A.-C.: Model to develop cost effective preventive maintenance program for material handling robot in intelligent warehousing system: A case study in AutoStore AS. Master's Thesis, University of Stavanger, 01.06.2018.
- Tjeerdsma S.: Redesign of the AutoStore order processing line: A multi-scenario discrete-event simulation study. Master thesis, University of Twente, 01.06.2019.
- Website Handling.de, <https://www.handling.de/lager-und-logistik/autostore-faehrt-2019-solides-finanzergebnis-ein.htm>. Aufruf am: 12.08.2020.
- Zoua, B., De Koster R., Xua X.: Evaluating dedicated and shared storage policies in robot-based compact storage and retrieval systems. PhD thesis, Huazhong University of Science and Technology. 21.12.2016



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Simulationsmodell mit 3D-Animation zur schnellen Bewertung von Ablaufplänen in der Produktion

***Simulation model with 3D animation for quick evaluation of production
schedules***

Benjamin Rolf, Tobias Reggelin, Otto-von-Guericke-Universität Magdeburg,
Magdeburg (Germany), benjamin.rolf@ovgu.de, tobias.reggelin@ovgu.de

Sebastian Lang, Fraunhofer-Institut für Fabrikbetrieb und -automatisierung IFF,
Magdeburg (Germany), sebastian.lang@iff.fraunhofer.de

Marcel Müller, Johann Prehm, Otto-von-Guericke-Universität Magdeburg,
Magdeburg (Germany), marcell.mueller@ovgu.de, johann.prehm@ovgu.de

Abstract: The paper presents a simulation case study dealing with a medium-sized company in the steel manufacturing industry. The production system of the company processes steel panels and profiles on two machines that are supplied with multiple roller conveyors and portal cranes. A special challenge is the partial return of raw material after processing that requires the reversing of the conveyors. We develop a discrete-event simulation model that is capable of testing sequences and evaluating the makespan of user-generated schedules quickly. The case study pays special attention to 3D animation of the material flow, flexible settings and ease of use even for users without simulation knowledge. The model returns a full schedule and provides key logistics indicators such as makespan and stock level. In the end we evaluate the functionality of the model with real job data provided by the company.

1 Motivation

In der stahlverarbeitenden Industrie sind die Ablauf- und Kapazitätsplanung entscheidende Faktoren, um die Auslastung der Maschinen zu maximieren und Kosten zu minimieren. Durch ereignisdiskrete Simulation soll die Planung einer neuen Fertigungslinie eines mittelständischen Unternehmens unterstützt werden. Die Planung sieht die Anschaffung von neuen Maschinen und Förderern vor und stellt für das Unternehmen eine hohe Investition dar.

Ein Simulationsmodell soll als Entscheidungshilfe in der operativen Planung eingesetzt werden und zusätzlich mögliche Engpässe und die zu erwartende Auslastung des geplanten Produktionssystems ermitteln. Die derzeitige Ablauf- und Kapazitätsplanung des Unternehmens basiert ausschließlich auf dem Erfahrungswissen der Mitarbeiter und nutzt kaum Hilfsmittel. Zukünftig soll das

Simulationsmodell dazu dienen, die von der Arbeitsvorbereitung erstellten Ablaufpläne zu visualisieren und zu validieren. Dadurch soll eine exaktere und transparentere Auftragsterminierung an die Werker übergeben werden. Außerdem können dadurch freie Kapazitäten für die Bearbeitung von Fremdaufträgen besser eingeschätzt werden.

2 Literaturanalyse

Konventionell wird ereignisdiskrete Simulation zur Entscheidungsunterstützung bei strategischen, langfristigen Entscheidungen wie der Layoutplanung oder der Prozessplanung eingesetzt. Das Simulationsmodell wird in diesem Fall für eine einmalige Anwendung im Planungsprozess entwickelt und hat danach keinen Nutzen mehr. Die Laufzeit spielt nur eine untergeordnete Rolle (Müller et al. 2019; Kasie et al. 2017; Smith 2003).

Durch stetige Verbesserung der Hard- und Software in den letzten Jahren werden Simulationsmodelle auch zunehmend in der operativen Planung in Bereichen wie Ablaufplanung, Kapazitätsplanung oder Transportplanung eingesetzt (Smith 2003). Eine geringe Laufzeit ist in diesem Fall ein entscheidender Faktor, um Entscheidungsunterstützung in Echtzeit zu ermöglichen (Müller et al. 2019; Kasie et al. 2017; Smith 2003). Laut Harmonosky und Robohn (1991) ist eine Echtzeitfähigkeit abhängig von der Zeit, die für eine Entscheidung zur Verfügung steht. Ein Modell ist echtzeitfähig, wenn die Laufzeit wesentlich geringer ist als die Zeit, in der die Entscheidung benötigt wird.

Das Hauptziel eines Echtzeit-Simulationsmodells in der Entscheidungsunterstützung ist es nicht, optimale Lösungen zu finden, sondern ein Hilfsmittel anzubieten, mit dem der Entscheider die bestmögliche Entscheidung treffen kann. Dadurch können in komplexen Systemen die Stärken von Mensch und Maschine kombiniert werden, nämlich die schnelle Datenanalyse der Simulation und die analytischen Fähigkeiten des Entscheiders (Heilala et al. 2010). Besonders wichtig ist es auch, die relevanten Informationen visuell anschaulich und übersichtlich darzustellen, um auch Anwendern mit geringen Simulationskenntnissen die Nutzung zu ermöglichen (Heilala et al. 2010; Andersson und Olsson 1998).

In der Literatur existieren bereits einige Anwendungen von ereignisdiskreten Simulationsmodellen als Entscheidungsunterstützung in Echtzeit. Kasie et al. (2017), Negahban und Smith (2014) und Smith (2003) geben einen Überblick über verschiedene Anwendungen in der Literatur. Giribone et al. (2003) präsentieren ein Simulationsmodell, mit dem wöchentliche Produktionspläne überprüft und validiert werden können. Müller et al. (2019) entwickeln ein echtzeitfähiges Simulationsmodell zur Sequenzoptimierung, Losgrößenoptimierung und Raumzuweisung für eine Industriewäscherei. Der Planer kann die Ergebnisse zur verbesserten Entscheidungsfindung für die tägliche Ablaufplanung verwenden. Andersson und Olsson (1998) entwickeln ein ereignisdiskretes Simulationsmodell für ein flexibles Fertigungssystem mit kundenauftragsorientierter Produktion. Durch das Modell kann abhängig vom aktuellen Systemzustand und dem Produktionsplan die Auslastung auf Tagesbasis ermittelt werden. Die Verwendung einer graphischen Benutzeroberfläche ermöglicht es, die Simulation ohne Simulationsvorkenntnisse des Anwenders durchzuführen. Die Lieferfähigkeit des Produktionssystems konnte durch die eingeführte Entscheidungshilfe erhöht werden.

3 Systembeschreibung

Die Produktion des Unternehmens aus der stahlverarbeitenden Industrie ist vollständig kundenauftragsorientiert und vor allem auf Großaufträge mit vielen Teilen ausgerichtet. Die Bearbeitung eines Auftrags kann daher mehrere Wochen dauern. Das Produktionssystem kann zwei verschiedene Arten von Rohstählen bearbeiten, nämlich Stahlbleche und Stahlprofile. Ein Auftrag besteht aus mehreren sogenannten Schachtelplänen, die jeweils geometrisch definieren, welche Teile an welcher Stelle aus einer Blechtafel bzw. aus einem Stahlprofil geschnitten werden sollen.

In diesem Beitrag wird der Material- und Produktionsfluss vom Wareneingang bis zur temporären Lagerung der Halbfabrikate betrachtet. In diesem Teilsystem befinden sich außerdem zwei Maschinen: Eine Brennerei für die Bearbeitung von Blechen und eine Säge für die Bearbeitung von Profilen. Die Komponenten des Produktionssystems sowie der Materialfluss sind in Abbildung 1 dargestellt.

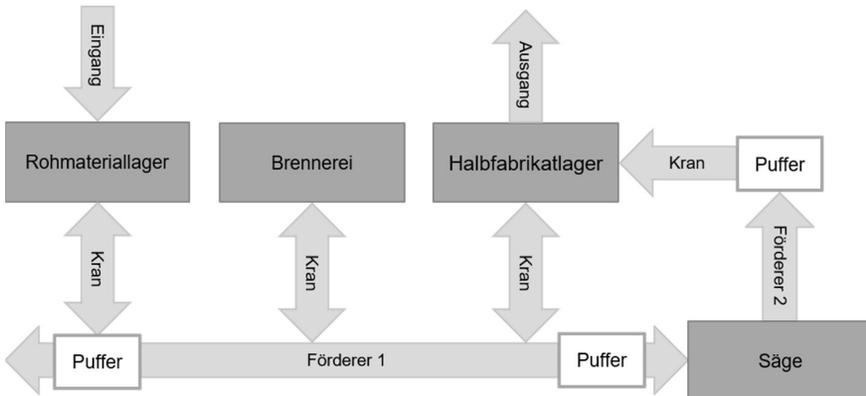


Abbildung 1: Materialfluss im Produktionssystem des Unternehmens

Über den langen Rollenförderer, der vom Außenlager bis zur Säge reicht, werden sowohl Bleche als auch Profile zu den jeweiligen Maschinen gefördert. Der Materialfluss vom Rohmateriallager, der Brennerei und dem Halbfabrikatlager zum Rollenförderer kann nur mithilfe von Portalkränen durchgeführt werden. Dazu kommen insgesamt vier Portalkräne zum Einsatz. Drei der Kräne laufen auf einer Schiene parallel zum langen Förderer, der Vierte läuft auf einer separaten Schiene im Anlieferungsbereich.

Nach der Bearbeitung durch die Brennerei bzw. die Säge entstehen Teile mit sehr unterschiedlichen Größen, die differenziert betrachtet werden müssen (siehe Tab. 1).

Tabelle 1: Teiletypen

Bleche	
Restbleche	Restbleche sind übriges Rohmaterial und müssen in das Rohmateriallager zurückgeführt werden.

Kleinteile	Kleinteile benötigen für den weiteren Transport Roste als Förderhilfsmittel, um nicht durch den Rollenförderer zu fallen.
Normalteile	Normalteile können ohne Einschränkungen transportiert werden.
Großteile	Großteile müssen ebenfalls zurückgeführt werden, da sie für den Transport auf den folgenden Rollenförderern zu breit sind.
Profile	
Restprofile	Restprofile sind übriges Rohmaterial und müssen in das Rohmateriallager zurückgeführt werden.
Teilprofile	Teilprofile können ohne Einschränkungen transportiert werden.

Eine Besonderheit in diesem Produktionssystem ist die Rückführung, bei der übrige Rohmaterialien in das Rohmateriallager zurückgeführt werden müssen. Die Rückführung findet ebenfalls über den langen Rollenförderer statt und erfordert dementsprechend eine temporäre Umkehrung der Förderrichtung. Alle Förderer und Portalkräne, die in beide Richtungen fördern können, sind als Doppelpfeile dargestellt. Die Brennerei hat keinen direkten Puffer, daher blockieren die erzeugten Halbfabrikate und die restliche Blechtafel die Brennerei solange, bis der Rollenförderer die Teile aufnehmen kann. Übrige Profile können zunächst im Halbfabrikatlager gelagert werden, um sie zu einem späteren Zeitpunkt zurückzuführen.

In diesem Produktionssystem ist vor allem eine effektive Steuerung für die Belegung des langen Rollenförderers notwendig, um Brennerei und Säge möglichst kontinuierlich mit Rohmaterial zu versorgen und gleichzeitig eine Blockierung der Maschinen durch rückzuführendes Rohmaterial zu vermeiden.

4 Konzeptionelles Modell

Das konzeptionelle Modell stellt die Grundlage für die Implementierung dar. Nach der Vorgehensweise von Robinson (2015) besteht die konzeptionelle Modellierung aus den folgenden fünf Schritten Problemverständnis, Festlegung der Projektziele, Identifikation der Eingabe- und Ausgabeparameter sowie die Definition des Modellinhalts mit Scope und Level of Detail.

Das Produktionssystem mit seinen Besonderheiten wurde bereits in Abschnitt 3 beschrieben. Hier muss vor allem der Rollenförderer als potenzieller Engpass durch Förderung von Teilen in beide Richtungen berücksichtigt werden. Das Ziel besteht darin, ein ereignisdiskretes Simulationsmodell als Echtzeit-Entscheidungsunterstützung für die Ablauf- und Kapazitätsplanung zu entwickeln. Zusätzlich wurden die folgenden Anforderungen durch das Unternehmen für die Entwicklung des Modells definiert:

1. Realitätsnahe 3D-Animation des Materialflusses
2. Einfache Bedienbarkeit für Anwender ohne Simulationskenntnisse
3. Flexible Einstellungsmöglichkeiten für das Systemverhalten

Eine realitätsnahe 3D-Animation und eine grafische Benutzeroberfläche mit anwenderfreundlichen Dialogen sind notwendig, um die Verwendung auch für Anwender ohne Simulationskenntnisse zu ermöglichen. Die Entscheidungen werden weiterhin durch den Planer getroffen und das Simulationsmodell hat eine reine Bewertungsfunktion (März und Krug 2011).

Die Eingabe- und Ausgabeparameter sind in Abbildung 2 dargestellt. Zu den Eingabeparametern zählen vor allem die Auftragsdaten mit Informationen zur Priorität, zum Material und zu den Bearbeitungszeiten. Die Priorität legt die Reihenfolge der Aufträge fest, in der diese aus dem Rohmateriallager entnommen werden. Außerdem können viele Einstellungen und Daten flexibel durch den Anwender festgelegt werden. Dazu zählen beispielsweise die Geschwindigkeiten und Aufnahmezeiten der Kräne, die Rüstzeiten der Maschinen, die Kapazitäten der Puffer, die Anlieferungszeiten im Rohmateriallager oder der Simulationszeitraum.

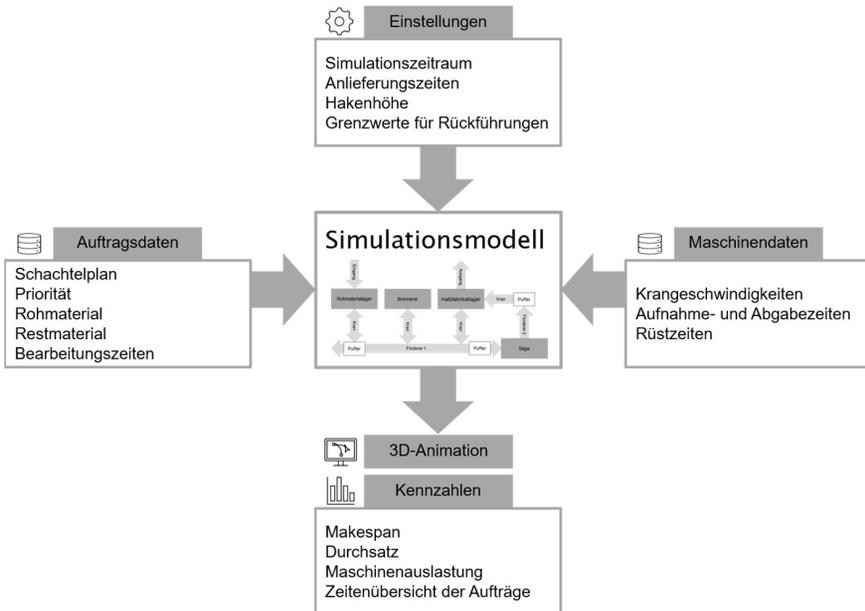


Abbildung 2: Eingaben und Ausgaben des Simulationsmodells

Die Ausgabe des Simulationsmodells besteht aus Kennzahlen zu den Maschinen und Aufträgen und der 3D-Animation. Hier liegt der Fokus auch auf dem Systemverhalten während des Simulationslaufs, um mögliche Probleme im Materialfluss durch das Erfahrungswissen des Planers frühzeitig zu erkennen. Wenn der Planer eine zufriedenstellende Auftragsreihenfolge gefunden hat, kann eine Übersicht ausgegeben werden, die für alle Aufträge das Eintreffen an bestimmten Punkten im System speichert. So kann an die Werker beispielsweise eine Übersicht übergeben werden, wann welches Rohmaterial aus dem Rohmateriallager entnommen werden muss. Dies ist wichtig, da der Transport vom Rohmateriallager bis zur Maschine durch hohe Transport- und Aufnahmezeiten des Krans bis zu eine Stunde dauern kann und im Voraus geplant werden muss.

Der Umfang des Modells umfasst den gesamten in Abbildung 1 dargestellten Bereich vom Eingang bis zum Ausgang. Beim Detailgrad wird zwischen den verschiedenen Elementen im System differenziert. Maschinen, Förderer, Portalkrane und bewegliche Elemente werden sehr detailreich dargestellt, da diese ausschlaggebend für die Einplanungsentscheidung sind. Die beiden Lager werden vereinfacht dargestellt, weil keine Informationen zu exakten Lagerplätzen des Rohmaterials bekannt sind und die Kapazitäten der Lager flexibel sind. Personalverhalten wird generell nicht berücksichtigt.

5 Implementierung

Zur Implementierung des Simulationsmodells wurde die ereignisdiskrete Simulationssoftware Tecnomatix Plant Simulation Version 15 verwendet. Das Modell ist mit einer Ausnahme ausschließlich deterministisch. Lediglich die tägliche Anlieferung im Rohmateriallager wird bei Bedarf durch eine normalverteilte Ankunft in einem definierten Zeitfenster modelliert. Das Modell benötigt für einen Simulationslauf nur wenige Sekunden und erfüllt daher die Anforderungen für den operativen Einsatz. Abbildung 3 zeigt einen Ausschnitt des Simulationsmodells.

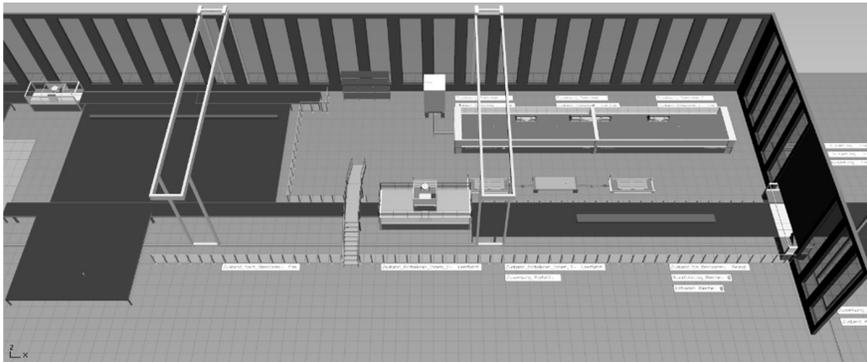


Abbildung 3: Ausschnitt des Simulationsmodells

Der Fokus der Implementierung liegt auf den in Abschnitt 4 formulierten Anforderungen. Um das Simulationsmodell für den Anwender möglichst einfach und übersichtlich zu gestalten, sind alle relevanten Daten in globalen Variablen und Tabellen gespeichert. Hier können beispielsweise Rüstzeiten, die Hakenhöhe der Kräne, Grenzwerte für Rückführungen oder der Simulationszeitraum festgelegt werden. Für die Portalkrane können die Geschwindigkeit für Last- und Leerfahrt sowie die Aufnahme- und Abladezeiten für die verschiedenen Teiletypen eingegeben werden. Die eingegebenen Werte werden automatisch in das Modell übertragen und bestimmen die Attribute der beweglichen Elemente (BEs) und Materialflussobjekte.

Außerdem wurde ein Dialog zur Eingabe von Schachtelplänen konzipiert. Der Anwender kann die Auftragsdaten aus seinem ERP-System direkt in den Dialog eingeben und im Nachhinein alle eingegebenen Schachtelpläne in einer Tabelle überprüfen. Für jeden Schachtelplan wird beim Starten eines Simulationslaufs ein BE angelegt, dem die Daten des Schachtelplans wie Auftragsnummer, Abmaße des Rohmaterials oder Bearbeitungszeit als Attribute übergeben werden. Nach dem

Simulationslauf können Kennzahlen und Ergebnisse direkt in Excel exportiert werden, um diese anschließend an die Werker im realen Produktionssystem zu übergeben. Der Dialog zur Eingabe von Schachtelplänen ist in Abbildung 4 dargestellt.

Index	Auftragsnummer	Auftragstyp	Priorität	Ausgangsmaterial Länge (mm)	Ausgangsmaterial Breite (mm)
39	112	Profil	1	18000	300
1	111	Profil	1	12000	800
2	111	Profil	1	12000	800
3	111	Profil	1	12000	800
5	112	Profil	1	18000	300
6	112	Profil	1	18000	300
7	112	Profil	1	18000	300
8	112	Profil	1	18000	300
9	112	Profil	1	18000	300
10	112	Profil	1	18000	300
11	112	Profil	1	18000	300
12	112	Profil	1	18000	300
13	112	Profil	1	18000	300
14	112	Profil	1	18000	300
15	112	Profil	1	18000	300
43	21166	Blech	2	6600	1600
44	21168	Blech	2	16100	2650
45	21172	Blech	2	16100	2650
46	21173	Blech	2	12100	2500
47	21178	Blech	2	12100	2500
48	21179	Blech	2	12100	2500
49	21180	Blech	2	12100	2500
50	21181	Blech	2	12100	2500

Abbildung 4: Dialog zur Eingabe von Schachtelplänen

Im gesamten Modell befinden sich Variablen auf der Oberfläche, die für wichtige Systemelemente Zustände zur aktuellen Simulationszeit anzeigen. Für die Brennerei als zentrales Objekt des Produktionssystems werden für beide Bereiche der aktuelle Zustand und der aktuell zugewiesene Schachtelplan angezeigt. Zusätzlich sorgt ein Farbcode für eine bessere Übersicht der Zustände in der Animation.

5.1 Verifikation und Validierung

Nach Rabe et al. (2008) wurden in dieser Fallstudie die Techniken Animation, Monitoring und Validierung im Dialog verwendet. Während des gesamten Simulationsprojektes wurden alle zwei Wochen Treffen mit einem Mitarbeiter des Unternehmens durchgeführt, der mit dem realen Produktionssystem vertraut ist. In diesem regelmäßigen Treffen wurden sowohl das konzeptionelle Modell, als auch das ausführbare Simulationsmodell im aktuellen Arbeitsstand verifiziert und validiert.

Da die realitätsnahe 3D-Animation ohnehin eine Anforderung des Unternehmens ist, kann diese auch zur Verifikation und Validierung während der Implementierung verwendet werden. Dadurch konnten die zeitlichen Abläufe im System kontinuierlich beobachtet und überprüft werden. Zur Übersichtlichkeit des Modells werden auch die Zustände der Maschinen, Puffer und Portalkräne zum aktuellen Zeitpunkt auf der Oberfläche abgebildet.

5.2 Evaluation

Das Simulationsmodell wurde beispielhaft mit einem aktuellen Großauftrag des Unternehmens getestet. Dieser Auftrag besteht aus 36 Blech-Schachtelplänen. Zusätzlich wurden 42 Profil-Schachtelpläne eingestreut, um eine realistische

Arbeitslast im untersuchten Zeitraum zu erreichen. In Tabelle 2 und Tabelle 3 sind die Eckdaten des Auftrags dargestellt.

Tabelle 2: Auftragsdaten der Bleche (36 Schachtelpläne)

Output [Stk.]		Bearbeitungszeiten [dd:hh:mm:ss]	
Restbleche	1	Minimum	00:00:10:07
Kleinteile	112	Maximum	00:10:48:18
Normalteile	164	Durchschnitt	00:03:09:06
Großteile	42	Summe	04:17:27:40

Tabelle 3: Auftragsdaten der Profile (42 Schachtelpläne)

Output [Stk.]		Bearbeitungszeiten [dd:hh:mm:ss]	
Restprofile	6	Minimum	00:00:10:00
Profilteile	96	Maximum	00:01:00:00
		Durchschnitt	00:00:14:45
		Summe	00:10:20:00

Der Auftrag enthält alle Teiltypen, darunter auch 1 Restblech, 42 Großteile und 6 Restprofile, die zurückgeführt werden müssen. Die Bearbeitungszeiten für Bleche sind sehr heterogen und reichen von wenigen Minuten bis zu mehreren Stunden. Die Schachtelpläne der Profile sind häufig standardisiert und haben kürzere Bearbeitungszeiten. In der Summe liegen die Bearbeitungszeiten für alle Schachtelpläne bei ca. fünf Tagen.

Alle Schachtelpläne des Auftrags müssen anfangs vom Planer manuell in das Simulationsmodell eingegeben werden, da derzeit keine Schnittstelle vorhanden ist. In Relation zur realen Produktionszeit ist dieser Aufwand jedoch gerechtfertigt, da durch einige Minuten Eingabezeit mehrere Wochen reale Produktionszeit abgebildet werden können. Sobald die Schachtelpläne einmalig eingegeben wurden, kann der Planer die Reihenfolge der Schachtelpläne durch einfaches Verändern der Auftragsindizes verändern. So lassen sich schnell unterschiedliche Szenarien simulieren. Die Erzeugung von unterschiedlichen Szenarien wurde bewusst nicht automatisiert, da durch die Komplexität des Produktionssystems ein ausgeprägtes Expertenwissen benötigt wird, um sinnvolle Reihenfolgen festzulegen. Die Planer des Unternehmens besitzen dieses Expertenwissen und nutzen das vorgestellte Simulationsmodell hauptsächlich, um Reihenfolgen zu validieren und zu visualisieren.

Außerdem erzeugt das Simulationsmodell für eingegebene Reihenfolgen einen detaillierten Ablaufplan mit Start- und Endzeiten an verschiedenen Kontrollpunkten im System. Abbildung 5 zeigt einen Ausschnitt aus dem durch das Simulationsmodell

generierten Ablaufplan für den Beispielauftrag. Der Ausschnitt enthält verschiedene Zeitstempel für die Bearbeitung auf der Brennerei. Dieser Ablaufplan kann von den Planern des Unternehmens an die verantwortlichen Mitarbeiter im Produktionssystem weitergegeben werden. Dadurch haben die Mitarbeiter genaue Informationen darüber, wann die Bearbeitung eines Schachtelplanes begonnen werden muss, um die Fertigstellungstermine einzuhalten und den Ablaufplan zu erfüllen. Zusätzlich gibt das Simulationsmodell Kennzahlen wie die Durchlaufzeit für jeden Auftrag oder den Makespan aus. Das Simulationsmodell gibt für den Auftrag unter Berücksichtigung aller Schichten und Transportzeiten einen Makespan von 14:02:17:00 (dd:hh:mm:ss) aus. Der Makespan ist hierbei der Zeitraum vom Ausgang des ersten BEs aus dem Rohmateriallager bis zur Ankunft des letzten BEs im Halbfabrikatlager.

Index	Auftragsnummer	Auftragstyp	Länge (mm)	Breite (mm)	Dicke (mm)	Rückführung	Ausgang Puffer	Eingang Brennerei	Beginn Bearbeitung	Ende Bearbeitung	Ende Entladen
43	1	Bleche	6600	1600	20	false	26.10.2020 08:06:03.3523	26.10.2020 08:29:56.8296	26.10.2020 08:29:56.8296	26.10.2020 10:40:45.8296	26.10.2020 11:45:22.8833
44	2	Bleche	16100	2650	10	false	26.10.2020 11:45:22.8833	26.10.2020 12:06:59.2133	26.10.2020 12:06:59.2133	26.10.2020 23:10:17.2133	27.10.2020 07:43:15.8952
45	3	Bleche	16100	2650	10	false	27.10.2020 07:43:15.8952	27.10.2020 08:04:52.6002	27.10.2020 08:04:52.6002	27.10.2020 18:23:46.6002	27.10.2020 20:34:26.7194
46	4	Bleche	12100	2500	10	false	27.10.2020 20:34:26.7194	27.10.2020 20:56:11.5583	27.10.2020 20:56:11.5583	28.10.2020 04:52:40.5583	28.10.2020 07:37:42.5922
47	5	Bleche	12100	2500	10	false	28.10.2020 07:37:42.5922	28.10.2020 07:59:27.4312	28.10.2020 07:59:27.4312	28.10.2020 10:46:30.4312	28.10.2020 12:35:12.5322
48	6	Bleche	12100	2500	10	false	28.10.2020 12:35:12.5322	28.10.2020 12:54:57.7711	28.10.2020 12:54:57.7711	28.10.2020 15:25:09.7711	28.10.2020 16:45:07.6655
49	7	Bleche	12100	2500	10	false	28.10.2020 16:45:07.6655	28.10.2020 17:06:52.5045	28.10.2020 17:06:52.5045	28.10.2020 19:30:16.5045	28.10.2020 20:47:31.2719
50	8	Bleche	12100	2500	10	false	28.10.2020 20:47:31.2719	28.10.2020 21:09:17.7358	28.10.2020 21:09:17.7358	28.10.2020 23:37:29.7358	29.10.2020 06:46:50.7674
51	9	Bleche	5050	1500	10	false	29.10.2020 06:46:50.7674	29.10.2020 07:08:49.0360	29.10.2020 07:08:49.0360	29.10.2020 07:55:15.0360	29.10.2020 08:12:39.2606
52	10	Bleche	6100	2300	25	false	29.10.2020 07:02:30.7765	29.10.2020 07:26:09.8216	29.10.2020 07:55:15.0360	29.10.2020 10:14:54.0360	29.10.2020 10:48:59.0785

Abbildung 5: Ausschnitt des Ablaufplans für den Beispielauftrag

6 Fazit und Ausblick

In dieser Fallstudie konnte das Produktionssystem eines mittelständischen Unternehmens erfolgreich in ein ereignisdiskretes Simulationsmodell übertragen werden. Das entwickelte Modell kann in der operativen Ablaufplanung als Hilfsmittel für die Entscheidungsfindung eingesetzt werden, um verschiedene Szenarien schnell und übersichtlich zu überprüfen. Durch eine realitätsnahe 3D-Animation und eine einfach gestaltete Benutzungsoberfläche profitieren auch Anwender ohne Simulationskenntnisse von diesem Modell.

Um realistische Ergebnisse zu erreichen, müssen derzeit die Daten eines jeden Schachtelplans manuell eingegeben werden. Dadurch entsteht ein Aufwand für die Auftragseingabe. In Zukunft könnte eine Schnittstelle zum ERP-System des Unternehmens oder eine automatische Erkennung von Schachtelplänen diesen Aufwand erheblich verringern. Außerdem kann der Einsatz von simulationsbasierter Optimierung mit Methoden des maschinellen Lernens oder Metaheuristiken weiterführend untersucht werden. Durch eine Automatisierung der zurzeit größtenteils manuellen Prozesse könnte die Anwendung von Optimierungsmethoden deutlich erleichtert werden.

Literatur

Andersson, M.; Olsson, G.: A Simulation Based Decision Support Approach for Operational Capacity Planning in a Customer Order Driven Assembly Line. In: Medeiros, D.J.; Watson, E.F.; Carson, J.S.; Manivannan, M.S. (Hrsg.): Proceedings of the 1998 Winter Simulation Conference, Washington, D.C., 13.-16.12.1998, 1998, S. 935–941.

- Giribone, P.; Mosca, R.; Queirolo, F.: Validating the Production Weekly Plan by Scheduling Simulation. In: Verbraeck, A.; Hlupic, V. (Hrsg.): Proceedings of the 15 European Simulation Symposium, 2003, S. 332–336.
- Harmonosky, C.M.; Robohn, S.F.: Real-time scheduling in computer integrated manufacturing: a review of recent research. *International Journal of Computer Integrated Manufacturing* 4 (1991) 6, S. 331–340.
- Heilala, J.; Montonen, J.; Jarvinen, P.; Kivikunnas, S.; Maantila, M.; Sillanpaa, J.; Jokinen, T.: Developing simulation-based Decision Support Systems for customer-driven manufacturing operation planning. In: Johansson, B.; Jain, S.; Montoya-Torres, J.; Hagan, J.; Yücesann, E. (Hrsg.): Proceedings of the 2010 Winter Simulation Conference, Baltimore, 5.-8.12.2010, 2010, S. 3363–3375.
- Kasie, F.M.; Bright, G.; Walker, A.: Decision support systems in manufacturing: a survey and future trends. *Journal of Modelling in Management* 12 (2017) 3, S. 432–454.
- März, L.; Krug, W.: Kopplung von Simulation und Optimierung. In: März, L.; Krug, W.; Rose, O.; Weigert, G. (Hrsg.): *Simulation und Optimierung in Produktion und Logistik*. Berlin, Heidelberg: Springer Berlin Heidelberg 2011, S. 41–45.
- Müller, M.; Reggelin, T.; Schmidt, S.: Operational Simulation-Based Decision Support in Intralogistics Using Short-Term Forecasts. In: Kabashkin, I.; Yatskiv, I.; Prentkovskis, O. (Hrsg.): *Reliability and Statistics in Transportation and Communication*. Cham: Springer International Publishing 2019, S. 345–352.
- Negahban, A.; Smith, J.S.: Simulation for manufacturing system design and operation: Literature review and analysis. *Journal of Manufacturing Systems* 33 (2014) 2, S. 241–261.
- Rabe, M.; Spieckermann, S.; Wenzel, S.: *Verifikation und Validierung für die Simulation in Produktion und Logistik: Vorgehensmodelle und Techniken*. Berlin, Heidelberg: Springer Berlin Heidelberg 2008.
- Robinson, S.: A tutorial on conceptual modeling for simulation. In: Yilmaz, L.; Chan, W.Kin; Moon, I.-C.; Roeder, T.M. (Hrsg.): Proceedings of the 2015 Winter Simulation Conference, 6.-9.12.2015, 2015, S. 1820–1834.
- Smith, J.S.: Survey on the use of simulation for manufacturing system design and operation. *Journal of Manufacturing Systems* 23 (2003) 4, S. 157–171.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Demand-Driven Resupply of Offshore Components by Cascading Simulation and Linear Optimization

Bedarfsgerechte Bereitstellung von Offshore-Komponenten durch die Kombination kaskadierender Simulation mit Linearer Optimierung

Daniel Rippel, Michael Lütjen, Michael Freitag, BIBA – Bremer Institut für
Produktion und Logistik GmbH, Bremen (Germany), rip@biba.uni-bremen.de,
ltj@biba.uni-bremen.de, fre@biba.uni-bremen.de

Helena Szczerbicka, Leibniz Universität Hannover, Hannover (Germany),
hsz@sim.uni-hannover.de

Abstract: The installation of offshore wind farms constitutes a highly weather-dependent process. Despite this dependency, practice and research generally apply a single, rigid resupply cycle to transport components from their production sites to the installation's base port, resulting in high storage costs. This article proposes combining mathematical optimizations with a cascading discrete-event simulation framework to select a viable resupply cycle online from a previously optimized pool of cycles during the project execution. This combination brings the best of both methods together by allowing high flexibility while reducing the possible search space drastically and guaranteeing that each route remains optimal by itself. The evaluation shows that dynamically selecting cycles using current weather measurements and forecasts reduces the required base port storage capacity by approximately 17% while still maintaining full installation efficiency.

1 Introduction

Over the last decade, wind energy has evolved into one of the primary sources of sustainable, green energy. It has witnessed an exponential increase in installed capacity, e.g., shown by REN21 (2020). Offshore wind farms provide advantages over their onshore counterparts. The open sea offers larger areas to install wind farms without interfering with local populations, allowing more capable wind farms. Additionally, higher wind speeds and availability result in higher energy yields. Nevertheless, the same advantages result in additional challenges for installing such wind farms, as, e.g., summarized by Rippel et al. (2019a).

Especially the high wind speeds at sea interfere with necessary installation operations. These installations require crane operations in approximately 100 meters height and highly depend on current weather conditions. When planning and

executing installation projects, planners and operatives need to rely on historical data or forecasts introducing a high amount of uncertainty into the planning process. Consequently, literature attributes 15% to 30% of a wind farm's overall lifetime cost to logistics during the installation (Dewan et al., 2015; Muhabie et al., 2018). These costs result from vessels' charter rates, heavy-duty handling equipment, and port spaces to store and process turbine components. Consequently, such projects include large safety margins for charter times or spaces.

While most literature focuses on efficient vessel schedules or fleet-mixes, only very few articles consider port-side resources and storage spaces. Nevertheless, current studies show that trends towards larger and heavier components (Wiser; Bolinger, 2018) and an increasing number of concurrent installation, refurbishing, and decommissioning projects, e.g., indicated in Beinke et al. (2020), might result in bottlenecks considering port-side storage areas (Oelker et al., 2020).

The article at hand focuses on the resupply of components to the base port. It proposes a framework to combine (offline) mathematical optimization with a cascading, simulation-based online approach to reduce the required capacity at base ports during the installation phase. Instead of applying metaheuristics to determine an optimized resupply cycle, this approach uses linear optimization to generate a small set of resupply cycles consisting of optimal routes and pickup amounts. The simulation model tracks the installation progress as online simulation until the real-world system requests a new resupply cycle. It then creates and evaluates additional instances using historical weather recordings. Each of these sub-simulations might create further instances, resulting in a cascading hierarchy of nested simulations.

1.1 Installation Process for Offshore Wind Farms

While several concepts exist for the installation of offshore wind farms, research and practice mainly apply the so-called conventional installation concept, e.g., described in Oelker et al. (2017). This concept includes three supply-chain stages. First, the production and (re-)supply of components to the base port. Second, the base port itself as a decoupling point between the first and the last stage, and, third, the actual installation (Fig. 1). In practice, companies conduct the installation sequentially, i.e., they first install all foundations and then the top-structures as both phases require the same resources but with different equipped tools for the vessels.

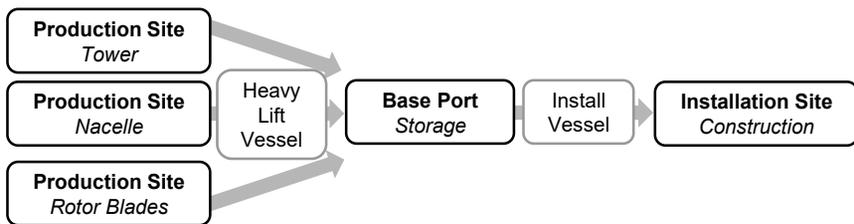


Figure 1: Conventional Installation Concept (Rippel et al., 2019b)

The base port stores delivered components and provides these for installation. On the one hand, its capacity needs to accommodate incoming components even when current weather conditions prohibit installations as planned. On the other hand, the

base port needs to provide sufficient inventory to continue installations even if the weather turns out better than expected. During the installation, a jack-up vessel travels between the base port and the installation site. It has to be noted that the jack-up vessel needs to finish installing a turbine once started. This requirement results from jack-up operations puncturing the seabed, rendering additional jack-up operations a risk to the vessel and already installed foundations and components.

The resupply applies one (or rarely more) heavy-lift vessel, which fetches components from the production ports and transports those to the base port. In practice and science, these vessels usually follow a predefined resupply cycle determined by project planners before the project starts. These cycles include a defined number of round-trips between the base port and one or more production ports. Each round-trip fetches a given number of components. As described above, the installation vessel always requires full sets of components, consisting of a tower, one nacelle, one hub, and three blades to perform an installation. Therefore, this article omits partial deliveries.

1.2 Approaches to Support Offshore Wind Farm Installations

Compared to other areas, like the maintenance of offshore installations, only a few works focus on installing offshore wind farms (Vis and Ursavas, 2016). Thereby, most authors focus on optimizing or evaluating the installation part of the supply chain (Rippel et al., 2019a), e.g., in terms of different weather assumptions, e.g., (Muhabie et al., 2018), installation concepts, e.g., (Vis and Ursavas, 2016), or fleet mixes, e.g., (Ait Alla et al., 2013). Other articles provide models to schedule, e.g., the commissioning of vessels (Kerkhove and Vanhoucke, 2017) or operations in various resolutions, e.g., (Ursavas, 2017; Irawan et al., 2019)

Even fewer articles explicitly include evaluations of port-side resources or the resupply part of the supply chain. For example, Beinke et al. (2017) evaluated the impact of sharing heavy-lift vessels between several installation projects to reduce downtime. In their current work, the authors describe a simulation study that demonstrates an increasing demand for jack-up vessels, and in consequence port-side resources, when the first wind farms need refurbishing or decommissioning over the following years (Beinke et al., 2020). Oelker et al. (2020) describe a simulation study that evaluates available heavy-duty storage areas at the base port in Eemshaven. The study demonstrates that the port's capacity will be exceeded soon if current trends to larger turbines and the increasing number of concurrent projects continue. Rippel et al. (2020a) describe a mathematical model to determine optimal resupply cycles for varying round-trips numbers. Further, that article selects a fixed, optimal resupply cycle for a practical use-case using additional optimization models.

The models denoted above primarily provide support for the installation stage of the supply chain. They assume that the base port offers sufficient components. A few models assume that a fixed number of components become available periodically.

1.3 Optimization of Resupply Routes

This section summarizes the optimization model proposed in Rippel et al. (2020a), as this article applies it to generate alternative resupply cycles. The model uses a combination of traveling salesmen problems with a customized multi-periodic knapsack formulation to determine a sequence of N^{rt} round-trips, that maximize the

number of delivered component sets while minimizing the overall cycle duration. The formulation exploits the small size of the transportation network to calculate optimal routes for each combination of visited ports. Assuming symmetrical traveling times, each round-trip uses one of eight possible routes (Fig. 2).

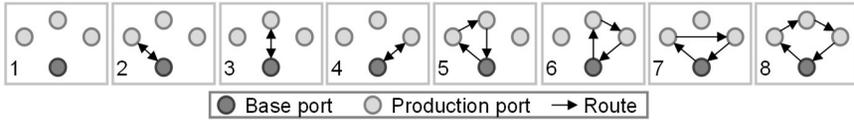


Figure 2: Possible routes for one base-port and three production ports

The model first determines the traveling times for each possible route by solving a traveling salesman problem. Afterward, it solves a multi-periodic knapsack problem to determine the optimal number of components of each type to fetch in each round-trip. The model uses a binary encoding to map each combination of visited ports to the precalculated traveling times of the respective route.

Consequently, the model obtains an optimized sequence of routes for a given number of round-trips N^{rt} , considering the supply network's geographical locations and the heavy-lift vessel's characteristics, e.g., its speed, deck area, and maximum payload. Finally, the article discusses the characteristics of different cycles. Finally, it shows that allowing four round-trips yields the most efficient cycle and results in the evaluated scenario's lowest capacity requirement.

2 Cascading Simulation using Optimized Routes

While the model presented above produces optimal cycles for a given supply network, it still assumes that the project repeats a single cycle. Nevertheless, the demand for components depends on current weather conditions, which cannot be known during planning. Consequently, this article extends that work by using nested simulations, as, e.g., introduced by Kindler (2004), at different cascades to choose the best cycle for the current state and expected weather conditions.

In contrast to widespread meta-heuristics or search-based approaches, the proposed method applies the optimization model to reduce the search space drastically. For example, allowing the heavy-lift vessel to decide between cycles with one to ten round-trips only requires evaluating a maximum of ten optimized routes. In contrast, evaluating the same space, e.g., using a simulation-based exhaustive search or a meta-heuristic, would require checking the complete or at least large parts of the search space. Therefore, a cycle with a single round-trip offers eight alternatives, as shown in Figure 2. Adding a second round-trip increases the search space multiplicatively, as each of the previous routes can combine with each one in the new trip. Thus, the total search space is given as $\sum_{n=1}^{N^{rt}} 8^n$, resulting in over 1.2 billion combinations if evaluating all options for $N^{rt} = 10$. Even assuming a single second to generate and either discard or simulate each alternative requires over 13.888 days for all alternatives. Additionally, these alternatives do not include the decision on how many components to pick up at each port.

In terms of a mathematical approach, the model would need to incorporate the interdependence between resupply cycles, time progression, weather-dependent installation times, and the influence of historical weather data. Such a high level of detail probably requires full-time-indexing of the project or a vast number of constraints. Consequently, this model would require excessive time to solve.

2.1 General Concept

The general concept consists of four major components. First, the real-world system, i.e., a current installation process that delivers current measurements and information to the framework and requests decisions on how to proceed. Second, the cascading simulation framework. This framework contains and manages a series of simulations on different levels. The underlying simulation model constitutes the third component. Thereby, this model takes on two tasks. On the one hand, it acts as an online simulation or digital twin, tracking the current real-world system's progress. The framework updates this model each time it receives updates from the real world, e.g., if an operation starts or ends or if new weather measurements or forecasts become available. On the other hand, the framework uses the simulation model to decide between alternatives. While the online simulation follows the real-world in terms of time progression and weather measurement, these nested simulations rely on forecasts and historical data. After all simulation on a cascade finish, the framework evaluates their results and updates the requesting simulation. Finally, the offline optimization model constitutes the fourth component to provide a set of optimized resupply cycles that act as inputs for the cascades.

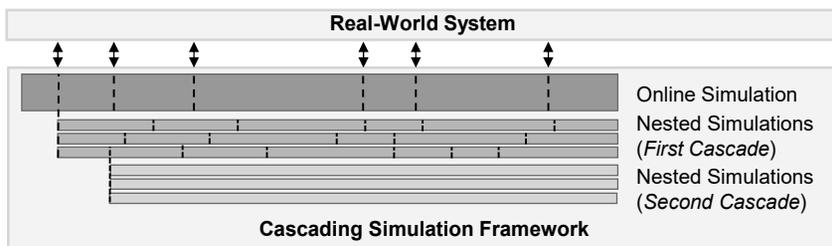


Figure 3: General Concept of the framework. The real-world system updates and request decisions from the framework (arrows). Internally, each such decision point (dotted line) spawns new nested simulations to evaluate three alternatives.

Figure 3 schematically shows this process for the first decision on each level. In this figure, dotted lines depict decision points, where, e.g., the heavy-lift vessel needs to decide on the next resupply cycle. The framework then instantiates a series of nested simulations (only three in the figure) based on the number of alternatives delivered from the offline simulation. If the framework allows for more cascades, each of these simulations may request the same decision, resulting in a new cascade for the decision points on this level too. In addition to what the figure shows, each decision point (dotted line) generates a set of simulations, e.g., resulting in six complete simulation sets in the first cascade.

2.2 Simulation Model and Weather Data

This article's simulation model has been implemented in AnyLogic 8.7.2, while the self-developed BIBA Cascading Simulation Framework (BCSF) has been implemented in Java. The framework consists of manager-classes that each reference one AnyLogic-Simulation and register themselves as listeners on this simulation. This structure allows for two-way access. On the one hand, the manager can update the simulation and control its execution. On the other hand, the simulation can request decisions or actively notify if it finishes. For each decision request, the manager creates more instances of itself. It then monitors these child-instances for completion and selects the best result for its own simulation. This article uses the following decision rules: First, select the child-instances with the lowest project duration. Second, select the instances which added the lowest value to the required base port capacity. Third, select the instances with the highest minimum inventory. Last, select the instance with the longest resupply cycle as a tie-breaker. While this article only considers resupply cycles as decision alternatives, the manager can influence basically all parameters and variables present within the simulation model. Its implementation uses Java-Reflections to modify variables "by name" which allows a high degree of freedom when formulating the request. Basically, when requesting a decision from the manager, the simulation provides a list of parameter names and values for each alternative, in this case resulting from the optimization model described earlier.

The installation uses a decision strategy that obtains current measurements and forecasts and applies the algorithms described in Rippel et al. (2019b) to estimate operations' durations. It then constructs different installation cycles, consisting of the number of loading and installation operations, to install between one and four (vessel capacity) turbines. It finally selects the alternative that incurs the lowest offshore waiting time, as these times inflict comparably high project costs. For more information on the simulation model, please refer to Rippel et al. (2020b).

Finally, this article uses a database of hourly historical weather recordings from Germany's Northern Sea between 1959 and 2006. The online simulation uses this database directly to provide "current" measurements and generates forecast data. The nested simulations cannot use this data directly but use aggregates of "historical" data as they would have been available at that time. For example, if the online simulation uses data from the year 2000, the nested simulations use hourly mean values from 1979 to 1999. The simulations slowly transition from current forecasts (2000) to the historical data (1979-1999) to accommodate for the forecasts' limited scope. According to the homepage of Deutscher Wetterdienst (2020), the forecasting uncertainty $u(t)$ of current models starts at 0.0 for the first hour (measurement), increases to approximately 0.25 at one week and 0.65 at two weeks. Afterward, it increases quickly to 0.95 at three weeks. The nested simulations use these values to mix the forecasts and current measurements with the mean aggregates as $f(t, d_c, d_h) = (1 - u(t)) \cdot d_c + u(t) \cdot d_h$ with t being the difference between the nested simulation time and the one of its parent, d_c being the vector of current measurements (2000), and d_h containing the historical aggregates (1979-1999). This procedure allows simulating an installation project with realistic data. The online simulation uses current measurements (of that time) while the child-simulations can only use historical data that would have been available at that time.

3 Experimental Setup

The evaluation applies the same parametrization and use-case as the original experiment in Rippel et al. (2020a), as it uses the same algorithm to generate optimized resupply cycles. This experiment follows a practical use-case originally described by Beinke et al. (2017). Moreover, it applies averaged data from different studies and websites to estimate the characteristics of turbine components (size and weight) and heavy-lift vessels (deck area, payload, speed). Table 1 summarizes the most relevant parameters for this evaluation. In contrast to the original experiment, the article at hand applies different decision strategies for the installation vessel. Moreover, this article selects the project start date as the first of August 2000 to have more volatile weather conditions, i.e., more time windows where no installation can occur. Both slightly change the results considering project durations and capacity requirements between the original results and those presented later.

Table 1: Summary of relevant parameters

Parameter	Tower	Blade	Nacelle
Base Port		Eemshaven	
Installation Site		Northern Sea	
HLV – Deck Area / Payload		2646 m ² / 8900 t	
HLV – Agv. Speed		9.5 knots	
Production Prot	Cuxhaven	Bremerhaven	Bremerhaven
Loading/ Unloading / Setup Time [h]	2 / 1.2 / 0	8 / 4.8 / 0	10 / 6 / 0
Weight	600 t	240 t	500 t
Required Space	650 m ²	300 m ²	263 m ²

4 Results and Discussion

The presented experiment aims to verify if the cascading framework can reduce base port capacity requirements by selecting viable resupply cycles online. Therefore, the simulations assume an infinite capacity and measure the maximum inventory level observed during the simulation. The original experiment concluded that cycles with three or four round-trips constitute the optimal choice. Therefore, these settings serve as the baseline for the evaluation by simulating the installation with fixed, repeated cycles. This article only depicts the results for four round-trips, as both simulation runs show equal results. Furthermore, the evaluation provides two sets of experiments allowing a single cascade (only the online simulation may create child-instances) and two cascades (instances on the first cascade can instantiate a second cascade). Each experiment includes ten simulation runs, each assuming an increasing initial inventory level, assessing this level's influence on the project duration and capacity requirement. Figure 4 depicts the results for fixed and cascading experiments. The graphs show the project duration, capacity requirement, and mean inventory levels observed during the simulation run for initial inventory levels between zero and twenty.

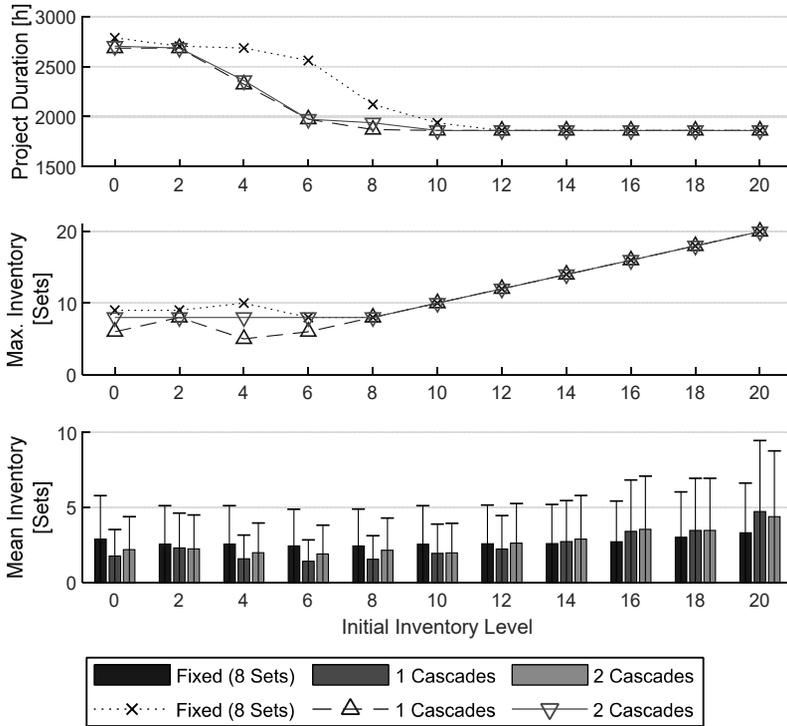


Figure 4: Results for all three experiments, showing the project duration, capacity requirement, and mean inventory level for different initial inventory levels.

The results show that the cascading approach outperforms the fixed cycle scenario by achieving an uninterrupted installation process at an initial inventory level of ten compared to twelve in the fixed case. Delays in the installation process result from missing components, also shown by the higher project durations at lower inventory levels. While expected for the fixed cycle experiment, the cascading approach also keeps the required capacity equal to the initial inventory. This stability and the low standard deviation of the mean inventory level show that the cascading approach reduces the impact of varying weather conditions quite well. Finally, the results show no notable differences between allowing one or two cascades for this example. Consequently, the reduced variance in inventory levels results from the framework's ability to adapt the resupply to the currently estimated demand. As the classical approach obtains the same number of component sets, good and bad weather periods have a stronger influence on the inventory level, even resulting in an undersupply. In contrast, the framework anticipates changing conditions and supplies components as needed, i.e., faster for good weather periods or slower for bad weather periods.

5 Conclusions and Future Work

This article presents a framework that combines mathematically optimized routing alternatives with a cascading simulation strategy to decide between them in an

online simulation context. Compared to a purely simulation-based optimization, mathematical optimization guarantees optimal route alternatives and drastically reduces the search space. The approach further uses current weather data and forecasts for the online simulation but relies on historical data aggregates to estimate the project's progression in its nested simulations. The experiment shows that applying the framework to dynamically choose the most promising resupply cycle reduces the required capacity by approximately 17% compared to repeating the same cycle periodically as it is the current state of practice. Moreover, the results show that this application only requires a single cascade. Future work will focus on applying the proposed approach to other decisions, e.g., to decide between different installation cycles. It will also investigate if the current listener-based connection between the simulation and the framework can be replaced using web services. This change might increase flexibility and facilitate integrating additional decision points.

Acknowledgment

The authors gratefully acknowledge the financial support by the DFG (German Research Foundation) for the Project "OffshorePlan", grant number (LU 2049/1-1 | SZ 51/33-1).

References

- Ait Alla, A.; Quandt, M.; Lütjen, M.: Simulation-based Aggregate Installation Planning of Offshore Wind Farms. *International Journal of Energy* 7 (2013) 2, pp. 23–30.
- Beinke, T.; Ait Alla, A.; Freitag, M.: Resource Sharing in the Logistics of the Offshore Wind Farm Installation Process Based on a Simulation Study. *International Journal of e-Navigation and Maritime Economy* 7 (2017) 6, pp. 42–54.
- Beinke, T.; Ait Alla, A.; Oelker, S.; Freitag, M.: Demand for Special Vessels for the Decommissioning of Offshore Wind Turbines in the German North Sea - a Simulation Study. In: Chung, J.S.; Akselsen, O.M.; Jin, H.; Kawai, H.; Lee, Y.; Matskevitch, D. (Hrsg.): *Proceedings of the 30th International Ocean and Polar Engineering Conference (ISOPE)*, October 12-16, Shanghai, China, 2020, pp. 5.
- Deutscher Wetterdienst, 2020: Verification of Results of Numerical Weather Prediction. Online verfügbar unter https://www.dwd.de/EN/research/weatherforecasting/num_modelling/05_verification/verification_node.html, zuletzt geprüft am 12.08.2020.
- Dewan, A.; Asgarpour, M.; Savenije, R.: Commercial Proof of Innovative Offshore Wind Installation Concepts Using Ecn Install Tool. Petten: ECN 2015.
- Irawan, C.A.; Wall, G.; Jones, D.: An Optimisation Model for Scheduling the Decommissioning of an Offshore Wind Farm. *OR Spectrum* 41 (2019) 3, 513–548.
- Kerkhove, L.-P.; Vanhoucke, M.: Optimised Scheduling for Weather Sensitive Offshore Construction Projects. *Omega* 66 (2017), pp. 58–78.

- Kindler, E.: SIMULA and Super-Object-Oriented Programming. In: Owe, O.; Krogdahl, S.; Lyche, T. (Hrsg.): *Lecture Notes in Computer Science*: Berlin: Springer 2004, pp. 165–182.
- Muhabic, Y.T.; Rigo, P.; Cepeda, M.; D'Agosto, M.A.: A Discrete-event Simulation Approach to Evaluate the Effect of Stochastic Parameters on Offshore Wind Farms Assembly Strategies. *Ocean Engineering* 149 (2018), pp. 279–290.
- Oelker, S.; Ait Alla, A.; Büsing, S.; Lütjen, M.; Freitag, M.: Simulative Approach for the Optimization of Logistic Processes in Offshore Ports. In: Chung, J.S.; Akselsen, O.M.; Jin, H.; Kawai, H.; Lee, Y.; Matskevitch, D. (Hrsg.): *Proceedings of the 30th International Ocean and Polar Engineering Conference (ISOPE)*, October 12-16, Shanghai, China, 2020, pp. 7.
- Oelker, S.; Lewandowski, M.; Ait Alla, A.; Ohlendorf, J.-H.; Haselsteiner, A.F.: Logistikszenerarien Für Die Errichtung Von Offshore-windparks - Herausforderungen Der Wirtschaftlichkeitsbetrachtung Neuer Logistikkonzepte. *Industrie 4.0 Management* 33 (2017) 1, pp. 24–28.
- REN21, 2020: *Renewables 2019 Global Status Report*, REN21. Online verfügbar unter <https://wedocs.unep.org/handle/20.500.11822/28496>, zuletzt aktualisiert am 2020, zuletzt geprüft am 14.01.2021.
- Rippel, D.; Jathe, N.; Becker, M.; Lütjen, M.; Szczerbicka, H.; Freitag, M.: A Review on the Planning Problem for the Installation of Offshore Wind Farms. *IFAC-PapersOnLine* 52 (2019a) 13, pp. 1337–1342.
- Rippel, D.; Jathe, N.; Lütjen, M.; Freitag, M.: Evaluation of Loading Bay Restrictions for the Installation of Offshore Wind Farms Using a Combination of Mixed-integer Linear Programming and Model Predictive Control. *Applied Sciences* 9 (2019b) 23, ArticleID: 5030.
- Rippel, D.; Jathe, N.; Lütjen, M.; Freitag, M.: A Mixed-Integer Formulation to Optimize the Resupply of Components for the Installation of Offshore Wind Farms. In: Bae, K.-H.; Feng, B.; Kim, S.; Lazarova-Molnar, S.; Zheng, Z.; Roeder, T.; Thiesing, R. (Hrsg.): *Proceedings of the 2020 Winter Simulation Conference*. New York: IEEE 2020a, pp. 1420–1431.
- Rippel, D.; Peng, S.; Lütjen, M.; Szczerbicka, H.; Freitag, M.: Model Transformation Framework for Scheduling Offshore Logistics. In: Ringle, C.M.; Kersten, W.; Jahn, C. (Hrsg.): *Data Science in Maritime and City Logistics*. Berlin: epubli GmbH 2020b, pp. 521–552.
- Ursavas, E.: A Benders Decomposition Approach for Solving the Offshore Wind Farm Installation Planning at the North Sea. *European Journal of Operational Research* 258 (2017), pp. 703–714.
- Vis, I.F.; Ursavas, E.: Assessment Approaches to Logistics for Offshore Wind Energy Installation. *Sustainable Energy Technologies and Assessments* 14 (2016), pp. 80–91.
- Wiser, R.; Bolinger, M., 2018: *2018 Wind Technologies Market Report*, U.S. Department of Energy. Online verfügbar unter <https://www.energy.gov/eere/wind/downloads/2018-wind-technologies-market-report>, zuletzt geprüft am 07.04.2020.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Utilising Relations between Actions to Improve the Performance of Optimisation Procedures for Distribution Networks

Nutzung von Aktionsbeziehungen zur Verbesserung der Performance von Optimierungsverfahren für Distributionsnetze

Markus Rabe, TU Dortmund, Dortmund (Germany), markus.rabe@tu-dortmund.de
Majsa Ammouriouva, TU Dortmund, Dortmund (Germany), Universitat Oberta de Catalunya, Barcelona (Spain), mammouriouva@uoc.edu

Abstract: Optimising distribution networks is challenging because of the complex relationships between entities in the network and their conflicting objectives. It is an NP-hard combinatorial optimisation problem, and a simheuristic approach has been proposed to solve these problems. Simheuristics combines simulation with metaheuristics to optimise a problem. However, in real-world networks, metaheuristics and simulation require long computation time. This paper describes correlation as domain-specific information to guide the search of a metaheuristic algorithm to construct promising solutions in order to reduce the number of simulation runs. The approach was tested on a distribution network of an international trading company. The utilisation of a correlation definition randomly biased the selection of solution elements to construct promising solutions.

1 Introduction

Managing distribution networks is a complex task, and the decision-makers responsible for this task have to select actions from a large number of potential actions to improve the networks' performance. Such actions could, e.g., increase the stock level of a stock keeping unit (SKU) at a site or centralise an SKU at a site in the network. These actions affect either a single entity or multiple entities in the network. Additionally, decision-makers have to consider conflicting goals, such as decreasing costs and increasing the service level (Rushton et al. 2017).

In distribution networks, selected actions are arranged in an action plan. Forming action plans is a combinatorial optimisation problem that is an NP-hard problem in large networks (Dross and Rabe 2014). Thus, decision-makers search for approaches to support them in selecting actions.

Because of distribution network's complexity, a close mathematical formulation of networks cannot be obtained. Instead, simulation is used (Law 2015). A simheuristic

approach combining simulation with metaheuristics could optimise these networks (Juan and Rabe 2013). In this approach, simulation evaluates the performance of the network, and metaheuristics optimises it. Even though this approach finds promising solutions in a finite time, its computation time increases with the size of the action space and the complexity of the simulation model. Researchers investigated approaches to improve optimisation approaches, such as simplifying optimisation problems (Ku and Arthanari 2016) and modifying the search mechanism (Grasas et al. 2017).

We defined domain-specific information to guide the search and exploration of actions and, hence, improve the selection of promising actions, such as success and type of changes (Rabe et al. 2017; Rabe et al. 2018a; Rabe et al. 2021). This information is associated with actions and is extracted from the problem. Additionally, correlation as domain-specific information is defined (Rabe et al. 2021). In this paper, we utilise this definition to modify the selection of actions in an evolutionary algorithm to form action plans. Additionally, we evaluate the effect of utilising the correlation definition to search for promising action plans.

This paper is organised as follows: Section 2 gives an overview of the optimisation methods and enhancement approaches, and Section 3 defines correlation. Modifying the action plan's construction in an evolutionary algorithm and its evaluation are presented in Sections 4 and 5, respectively. Section 6 closes with a conclusion.

2 Related Work

This section presents briefly related work in the optimisation of distribution networks and the improvement to optimisation methods.

2.1 Optimisation of Logistics Distribution Networks

Various optimisation methods could be used to optimise logistics distribution networks. Most of the problems in these networks are combinatorial optimisation problems, in which elements are selected from a finite set of elements (Korte and Vygen 2018), such as the travelling salesman problem and vehicle routing problem. These problems are NP-hard problems that are difficult to be solved.

Additionally, distribution networks are characterised by high uncertainty, a large number of entities, e.g., SKUs and sites, and conflicting performance measures (Rushton et al. 2017). A performance measure evaluates the performance of the network (Rushton et al. 2017). Decision-makers can use these measures to compare a network's performance to previous periods or analyse the impact of changes on the network. Examples of conflicting performance measures are costs to be decreased and service level to be increased. These complex networks are challenging to be formulated using mathematical equations and optimised using exact methods.

Researchers used several approaches to optimise these networks. For example, they used metaheuristics and simheuristics (Calvet et al. 2019). In a simheuristic approach, simulation is integrated with metaheuristics. A simulation evaluates changes applied on a network model, and metaheuristics explores the search space for solutions to optimise the network, e.g., the evolutionary algorithm (cf. Talbi (2009) for the description of evolutionary algorithms). These approaches are iterative and terminate if a pre-defined condition is met.

Additionally, researchers have built decision support systems based on optimisation methods. If the system is designed for a logistics network, it is called a logistics assistance system (LAS) (Liebler et al. 2013). Such a system was developed by Dross and Rabe (2014). Figure 1 demonstrates a simplified architecture of the developed LAS. The architecture presents a simheuristic framework that is the base of this LAS to recommend action plans. Discrete event simulation is integrated within a metaheuristic algorithm in the heuristic unit (cf. Fig. 1). This metaheuristic algorithm selects actions and forms action plans, while the simulation evaluates these action plans with respect to selected performance measures, such as costs and service level. The search space consists of actions derived from action types that present a generic description of actions, such as “centralise” without specifying the affected entities.

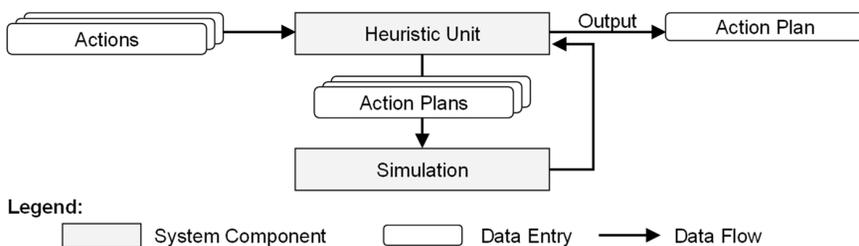


Figure 1: The simplified architecture of the LAS based on Rabe et al. (2017)

2.2 Improving the Performance of Optimisation Algorithms

Researchers investigated approaches to improve the performance of optimisation methods, such as screening solutions (Karimi et al. 2017; Bode et al. 2019) and simplifying optimisation problems (Ku and Arthanari 2016). Karimi et al. (2017) clustered solutions for a genetic algorithm at the beginning of the search and showed that their approach outperformed the ordinary genetic algorithm. Bode et al. (2019) defined criteria to screen solutions to reduce the size of the search space. Ku and Arthanari (2016) and Rabe et al. (2018b) replaced the search space of a problem with a smaller search space for easier exploration. Juan et al. (2015) presented a methodology to solve stochastic optimisation problems based on the deterministic version of the problem.

Other researchers focused on modifying the search of the optimisation algorithms. For example, Mane and Narsingrao (2021) used a chaotic sequence to modify the search of a metaheuristic algorithm to solve multi-objective optimisation problems. Alsheddy et al. (2018) integrated a penalty term in their objective function, which is updated during the search to escape local optima. Grasas et al. (2017) proposed a random biased selection of customers in vehicle routing problems to obtain promising solutions.

In addition, researchers investigated hybrid approaches in the optimisation of problems. For example, Li et al. (2020) utilised simulated annealing in modifying solutions in a genetic algorithm. They used their approach to reduce the distribution costs of fresh food logistics. Umetani (2017) used data mining techniques to reduce the size of a search space.

3 Correlation between Actions in the Logistics Assistance System

Correlation in this research is domain-specific information. In this context, we define it as a relation between two sequential actions, e.g., a_i and a_j , and their impact on the performance of a distribution network, $R([a_i, a_j])$. This performance could be measured based on any of the performance measures, such as costs. The relation is defined by comparing $R([a_i, a_j])$ to the estimated impact of the single actions. This estimation assumes that the actions' impact is independent. Hence, their impact together is calculated as $R(a_i) + R(a_j)$. The impact of the actions could be determined based on expert knowledge or a simulation study.

According to the comparison between $R([a_i, a_j])$ and $R(a_i) + R(a_j)$, a positive, a negative, or no recognizable relation could be found. A positive relation (+) exists if the actual impact of the sequential actions exceeds the estimated impact, $R([a_i, a_j]) > R(a_i) + R(a_j)$. In this relation, applying the sequential actions improves the performance of the network. For example, applying a_1 followed by a_2 , $[a_1, a_2]$, reduces the costs by 80 € in the example shown in Figure 2. The estimated reduction in costs would have been 65 €, resulting from summing up the reduction caused by a_1 (25 €) and a_2 (40 €).

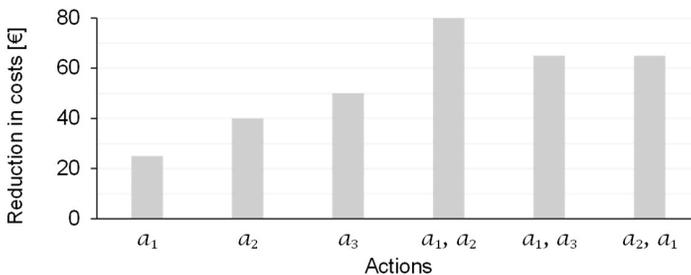


Figure 2: An impact of actions on the reduction of costs in a distribution network

In contrast, a negative relation (−) indicates that applying the sequential actions deteriorate the performance of the network, $R([a_i, a_j]) < R(a_i) + R(a_j)$. For example, applying $[a_1, a_3]$ reduces costs by 65 € compared to the estimated reduction of 75 € (Fig. 2). The last relation is defined if the actual impact is very near to the estimated impact of the sequential actions, $R([a_i, a_j]) \approx R(a_i) + R(a_j)$. In stochastic networks, several simulation runs are used to calculate the impact of the actions. Thus, we call this relation a weak relation (\sim), because no significant evidence is available to classify the relationship as positive or negative.

In the LAS, the actions are derived from action types (AT). Thus, relations between action types should be defined. These relation definitions are then assigned to the derived actions. In our work, we assume that actions that affect different entities in a network are independent; thus, the relation between these actions and their impact on the network is weak. Based on this assumption, relations between actions affecting the same entities should be investigated to define the relation between action types.

For example, to define the relation between applying actions from AT_1 followed by actions from AT_2 , we investigate $[a_i, a_j]$, where a_i and a_j affect the same entities and are derived from AT_1 and AT_2 , respectively. A significant relation defines the impact of applying actions from AT_1 followed by actions from AT_2 .

The relations are arranged in a matrix, as shown in Table 1. Rows represent the first applied actions, and columns represent the second applied actions. Thus, applying actions derived from AT_3 followed by actions derived from AT_1 has a negative impact if the actions affect the same entities. For example, action a_3 derived from AT_3 and action a_1 derived from AT_1 affect SKU 1 at site A. Then, the sequence $[a_3, a_1]$ has a negative impact on the performance of the network, but the sequence $[a_1, a_3]$ has a positive impact (cf. Tab. 1). If a_3 affects SKU 3 at site A, both sequences have a weak impact on the performance of the network.

Table 1: Correlation matrix of action types in the LAS

	AT_1	AT_2	AT_3	AT_4
AT_1	+	+	+	~
AT_2	+	+	~	-
AT_3	-	-	~	-
AT_4	-	-	~	-

4 Utilising Correlation in a Logistics Assistance System

Once the relations are defined, they are used to bias the selection of actions in constructing action plans in the LAS. The construction of action plans occurs at the initial generation of the evolutionary algorithm. Then, action plans are modified in subsequent generations in the evolutionary algorithm using crossover and mutation.

4.1 Constructing Action Plans Utilising Correlation

Action plans are constructed by selecting actions and adding them sequentially. For example, action plan $[a_3, a_5, a_{10}, a_1]$ is constructed by selecting a_3 as the first action. Then, a_5 is selected and added to the action plan followed by a_{10} and a_1 . To incorporate the correlation in constructing action plans, we defined several approaches to alter the selection of actions – three out of five approaches (Rabe et al. 2021) are described here and are utilised in experiments in Section 5. In these approaches, actions are classified into three classes: Positively (A^+), negatively (A^-), and weakly (A^\sim) correlated actions. Each class is assigned a probability to select actions from; the highest probability is assigned to A^+ and the lowest to A^- . After selecting a class, an action is selected randomly from the respective class.

In approach (1), the last-added action to an action plan classifies actions into the three classes. For example, after selecting the first action, a_{10} , actions are classified into $A_{a_{10}}^+$, $A_{a_{10}}^-$, and $A_{a_{10}}^\sim$ according to their correlation with a_{10} . Next, a class is selected, and an action is selected from the class. This classification and selection are repeated until the action plan is constructed.

The classification classes are associated with the action plan being constructed in approaches (2) and (3). Thus, after the selection of the first action, the classification classes become A_s^+ , A_s^- , and A_s^{\sim} and are updated with each selection of an action. In approach (2), actions are moved from their class to a higher related class if they are classified positively correlated with the newly selected action. Actions are moved from A_s^- to A_s^{\sim} or from A_s^+ to A_s^{\sim} if a newly selected action a^* classifies them positive, $A_{a^*}^+$. The actions are moved in the other direction if they are classified negatively by a^* . Figure 3 shows an example. On the one hand, actions a_5 and a_2 are classified positively correlated with a_4 and are moved from their classification class to A_s^{\sim} and A_s^+ , respectively. On the other hand, action a_7 was moved from A_s^{\sim} to A_s^- .

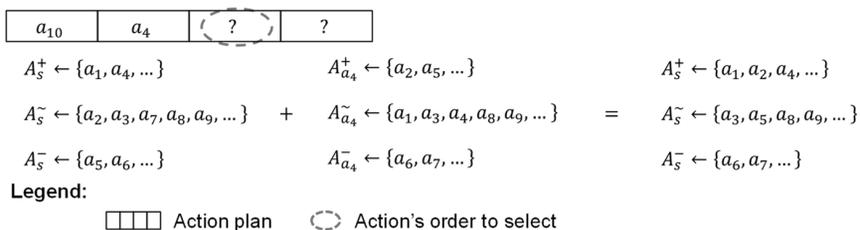


Figure 3: Updating classification classes based on approach (2)

In approach (3), if an action is classified as $A_{a^*}^-$ by a selected action a^* , it remains in A_s^- . Thus, once an action is classified with a negative relation with an action in the action plan, it stays negatively related to the whole action plan. An action that is classified positively by a^* is moved from A_s^{\sim} to A_s^+ . If approach (3) is applied in Figure 3, the resulted classification classes A_s^+ , A_s^- , and A_s^{\sim} become $\{a_1, a_2, a_4, \dots\}$, $\{a_3, a_8, a_9, \dots\}$, and $\{a_5, a_6, a_7, \dots\}$, respectively. Actions a_5 and a_7 are classified as A_s^- , because they have a negative relation with an action added to the action plan.

4.2 Modifying Action Plans Utilising Correlation

Variation operators modify the action plans in subsequent generations in the evolutionary algorithm, such as crossover and mutation operators. These operators cause destructive variations to the sequence of actions constructed in previous generations to produce new action plans (solutions).

Crossover could take several forms, such as one-point, two-point, and uniform crossover. Since approaches (2) and (3) classify actions according to all previously added actions in an action plan, two-point and uniform crossover do not cause destructive changes in the sequence as in the case of approach (1).

In mutation, one of the actions is replaced by another action from the available actions. The action to be replaced is selected randomly, and the replacement action is selected according to the approaches. For example, if approach (1) is utilised, the replacement action is selected according to the classification based on the preceding action. If approaches (2) or (3) are utilised, all actions before the replaced action should be considered to classify the actions. This replacement could take place for one or more actions in an action plan.

5 Experimental Results and Analysis

In order to evaluate the impact of correlation exploitation, we used a database of a distribution network supplied by an international material trading company. The supplied database stores the network data, such as SKUs, sites, and suppliers. This database is used to instantiate the simulation model in the LAS. The experiments and the correlation between action types in the LAS were defined (Section 5.1). Next, experiment results were collected and analysed (Section 5.2).

5.1 Experiment Setup

In the experiments, the evolutionary algorithm parameters were set to be used in the base experiment and the experiment utilising correlation (the biased experiment). The parameters were set as follows: Crossover probability 0.8, mutation probability 0.3, population size 50, and the maximum number of generations 100. Crossover forms varied in the experiment as one-point, two-point, and uniform crossover. The mutation was set as a one-action or multiple-actions mutation in an action plan. The performance of the evolutionary algorithm is measured by the quality of recommended action plans presented as costs (c) and service level, as well as the number of generations needed to stagnate (N_s). In the experiments, four action types were considered, and each experiment was repeated ten times.

Simulation experiments were used to define the correlation of actions and, accordingly, their action types. To determine the correlation between two action types, pairs of two actions derived from them and affect the same entities in the network were considered. The two actions were applied sequentially on a network model, and their effect on the performance was recorded as actual impact, $R([a_i, a_j])$. This impact was compared to the expected impact of single actions, $R(a_i) + R(a_j)$. For each experiment of a sequential pair of action types, the number of improvements and deterioration in the performance of the network was recorded. Table 2 shows that duplicating an action derived from AT_1 reduced costs in 53% of the pairs. Applying action derived from AT_1 followed by an action derived from AT_2 reduced costs in 75% of the actions pairs that affect the same entities.

Table 2: Sample of changes in costs after applying two sequential actions affecting the same entity

Action type of first action	Action type of second action	Percentage of action pairs that have		
		Positive impact	No impact	Negative impact
AT_1	AT_1	53	27	20
AT_1	AT_2	75	17	8
AT_1	AT_3	50	17	33

In order to define the impact of the sequential actions, the sign test (cf. Sheskin 2011) was used. In the test, the median of the changes presents this relation; a positive median indicates a positive relation. The test conclusions are presented in the correlation matrix in Table 1. The positive relation is defined for actions derived from AT_1 followed by actions derived from AT_2 with a p -value of 0.0000. The relation was

neither significantly positive nor negative for actions derived from AT_4 followed by actions derived from AT_3 ; hence, the relation is weak (cf. Tab. 1).

5.2 Results and Discussion

First, correlation approaches for selecting actions were compared. The approaches differ significantly with respect to the quality of found solutions (cf. Fig 4). The CD in Figure 4 is the critical difference calculated in the Fishers test (cf. Sheskin 2011) to identify significant factors after an ANOVA analysis. A similar analysis was performed considering the service level. These approaches do not differ significantly with respect to N_s . From these comparisons, we selected approach (3) for further experiments. In this approach, an action classified negative by an action in the action plan stays negative for the action plan.

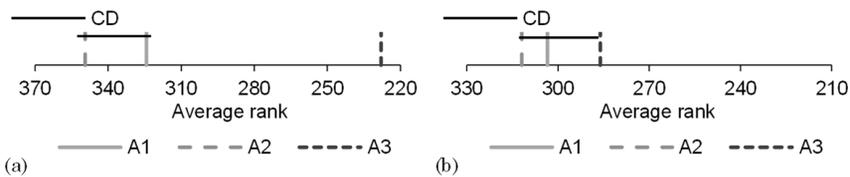


Figure 4: Critical difference plots for the comparison of correlation approaches: (a) c and (b) N_s

Next, we analysed the effect of utilising correlation (the biased experiment) compared to the base experiment. In this comparison, approach (3) was utilised in constructing action plans in the initial generation, initial solution. Table 3 summarises this comparison. According to the experiments, utilising correlation led to better solutions with a p -value of 0.000 in the Mann-Whitney U test (cf. Sheskin 2011). However, the base experiments converged earlier.

Table 3: Comparison between the base experiment and biased experiments for approach (3)

Crossover	Mutation	Base experiment		Biased experiment		Gap analysis	
		c [€] (1)	N_s (2)	c [€] (3)	N_s (4)	(3) - (1)	(4) - (2)
One-point	One-action	90,420	76	90,835	76	415	0
One-point	Multi-actions	91,414	41	90,802	81	-612	40
Two-point	One-action	90,630	73	90,106	69	-524	-4
Two-point	Multi-actions	90,673	65	89,883	83	-790	18
Uniform	One-action	90,086	89	89,765	73	-321	-17
Uniform	Multi-actions	91,699	63	90,820	74	-879	11

In the second part of the experiment, the effect of utilising correlation in subsequent generations was analysed. Utilising correlation in these generations led to finding

promising solutions with a p -value less than 0.1000, in general. To reduce N_s , specific combinations of crossover and mutation were needed.

We defined relations between actions and their impact on the performance of the network. These relations guided the exploration of actions and constructing action plans. As a result, promising solutions were found because actions with a negative impact on the network with respect to other actions get a lower selection probability. Utilising specific knowledge does not support early convergence with all variation operators' combinations. However, finding promising solutions earlier enables us to terminate the search in a fewer number of generations; thus, in a lower number of simulation runs.

6 Conclusion and Outlook

In this paper, we have presented an evaluation of domain-specific information, correlation. Correlation defines the relationship between sequential actions and their impact on the performance of a distribution network. This relation guides the exploration of actions and the construction of action plans to recommend promising action plans that improve the network. Recommending promising solutions in early generations could result in decreasing the number of simulation runs. For further research, correlation definition could be combined with other defined domain-specific information, such as success and type of changes.

References

- Alsheddy, A.; Voudouris, C.; Tsang, E. P. K.; Alhindi, A.: Guided Local Search. In: Martí, R.; Pardalos, P. M.; Resende, M. G. (Eds.): *Handbook of Heuristics*. Cham, Switzerland: Springer, 2018, pp. 261–297.
- Bode, F.; Reed, P.; Reuschen, S.; Nowak, W.: Search Space Representation and Reduction Methods to Enhance Multiobjective Water Supply Monitoring Design. *Water Resources Research* 55 (2019) 3, pp. 2257–2278.
- Calvet, L.; Wang, D.; Juan, A. A.; Bové, L.: Solving the Multidepot Vehicle Routing Problem with Limited Depot Capacity and Stochastic Demands. *International Transactions in Operational Research* 26 (2019) 2, pp. 458–484.
- Dross, F.; Rabe, M.: A SimHeuristic Framework as a Decision Support System for Large Logistics Networks with Complex KPIs. In: Wittmann, J.; Deatcu, C. (Eds.): *Proceedings of the 22nd Symposium Simulationstechnik*, Berlin, Germany, 3–5 September. 2014, pp. 247–254.
- Grasas, A.; Juan, A. A.; Faulin, J.; de Armas, J.; Ramalhinho, H.: Biased Randomization of Heuristics Using Skewed Probability Distributions: A Survey and some Applications. *Computers and Industrial Engineering* 110 (2017) August, pp. 216–228.
- Juan, A. A.; Rabe, M.: Combining Simulation with Heuristics to Solve Stochastic Routing and Scheduling Problems. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Eds.): *Simulation in Production and Logistics – Entscheidungsunterstützung von der Planung bis zur Steuerung*. Paderborn, Germany: HNI-Verlagsschriftenreihe, 2013, pp. 641–649.

- Juan, A. A.; Faulin, J.; Grasman, S. E.; Rabe, M.; Figueira, G.: A Review of Simheuristics: Extending Metaheuristics to Deal with Stochastic Combinatorial Optimization Problems. *Operations Research Perspectives* 2 (2015) 1, pp. 62–72.
- Karimi, M. B.; Isazadeh, A.; Rahmani, A. M.: QoS-aware Service Composition in Cloud Computing Using Data Mining Techniques and Genetic Algorithm. *The Journal of Supercomputing* 73 (2017) 4, pp. 1387–1415.
- Korte, B.; Vygen, J.: *Combinatorial Optimization: Theory and Algorithms*. 6th ed. Berlin, Germany: Springer, 2018.
- Ku, D.; Arthanari, T. S.: On the Abstraction Method for the Container Relocation Problem. *Computers and Operations Research* 68 (2016), pp. 110–122.
- Law, A. M.: *Simulation Modeling and Analysis*. 5th ed. New York City, New York: McGraw-Hill, 2015.
- Li, D.; Cao, Q.; Zuo, M.; Xu, F.: Optimization of Green Fresh Food Logistics with Heterogeneous Fleet Vehicle Route Problem by Improved Genetic Algorithm. *Sustainability* 12 (2020) 5, article 1946.
- Liebler, K.; Beissert, U.; Motta, M.; Wagenitz, A.: Introduction OTD-NET and LAS: Order-to-Delivery Network Simulation and Decision Support Systems in Complex Production and Logistics Networks. In: Pasupathy, R.; Kim, S.-H.; Tolk, A.; Hill, R. R.; Kuhl, M. E. (Eds.): *Proceedings of the 2013 Winter Simulation Conference*. Piscataway, New Jersey: IEEE, 2013, pp. 439–451.
- Mane, S. U.; Narsingrao, M. R.: A Chaotic-based Improved Many-objective Jaya Algorithm for Many-objective Optimization Problems. *International Journal of Industrial Engineering Computations* 12 (2021) 1, pp. 49–62.
- Rabe, M.; Dross, F.; Schmitt, D.; Ammouriova, M.; Ipsen, C.: Decision Support for Logistics Networks in Materials Trading Using a Simheuristic Framework and User-generated Action Types. In: Wenzel, S.; Peter, T. (Eds.): *Simulation in Production and Logistics 2017*. Kassel, Germany: Kassel University Press, 2017, pp. 109–118.
- Rabe, M.; Ammouriova, M.; Schmitt, D.: Utilizing Domain-specific Information for the Optimization of Logistics Networks. In: Rabe, M.; Juan, A. A.; Mustafee, N.; Skoogh, A.; Jain, S.; Johansson, B. (Eds.): *Proceedings of the 2018 Winter Simulation Conference*. Piscataway, New Jersey: IEEE, 2018a, pp. 2873–2884.
- Rabe, M.; Schmitt, D.; Ammouriova, M.: Improving the Performance of a Logistics Assistance System for Materials Trading Networks by Grouping Similar Actions. In: Rabe, M.; Juan, A. A.; Mustafee, N.; Skoogh, A.; Jain, S.; Johansson, B. (Eds.): *Proceedings of the 2018 Winter Simulation Conference*. Piscataway, New Jersey: IEEE, 2018b, pp. 2861–2872.
- Rabe, M.; Ammouriova, M.; Schmitt, D.; Dross, F.: Simheuristics Approaches for Efficient Decision-Making Support in Materials Trading Networks. *Algorithms* 14 (2021) 1, article 23.
- Rushton, A.; Croucher, P.; Baker, P.: *The Handbook of Logistics and Distribution Management*. 6th ed. London, United Kingdom: Kogan Page, 2017.
- Sheskin, D.: *Handbook of Parametric and Nonparametric Statistical Procedures*. 5th ed. Boca Raton, Florida: CRC Press, 2011.
- Talbi, E.-G.: *Metaheuristics: From Design to Implementation*. Hoboken, New Jersey: John Wiley and Sons, 2009.
- Umetani, S.: Exploiting Variable Associations to Configure Efficient Local Search Algorithms in Large-scale Binary Integer Programs. *European Journal of Operational Research* 263 (2017) 1, pp. 72–81.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Analytical and Empirical Study of Proper Parameters for Theory of Constraints under Uncertainty

Analytische und empirische Studie geeigneter Parameter für die Engpasstheorie bei Unschärfe

Andreas Rükbauer, Fachhochschule Würzburg-Schweinfurt (FHWS), Würzburg
(Germany), andreas.ruekgauer@fhws.de

Abstract: This paper discusses stochastic analysis and defining parameters of drum-buffer-rope (DBR) by way of a reorder point system, in particular, the relationship between reorder point and cycle service level (CSL), as well as fill rate (FR). Analytical solutions are verified and validated experimentally through simulation. Literature typically discusses batch-oriented consumption, in reality however piece wise consumption with processing time per piece is more typically for production. There is no simple conversion between the two, but either requires a different experiment setup altogether. The difference between CSL, and FR is discussed next. The relationship between FR and reorder point proves complex, in particular in light of distribution functions other than normal distribution. In addition to variation in consumption, the model is then extended to supply variation.

1 Introduction

The Theory of Constraints (TOC) (Goldratt and Cox, 2004) is extensively discussed and analyzed in literature (Cox, J. and Schleier, J. 2010; Lödging, 2016) This paper extends the discussion and tries to understand TOC better from a stochastic perspective.

Originally, TOC has been conceived as a means for robust production control. The core problem of robustness in production control results from variation in almost all variables involved: demand, execution time, yield in execution and supply, capacity and availability, and so forth. Therefore, parameter selection in TOC must take the major sources of variability into account.

The aim of this paper is to study and understand these stochastic effects better and to provide mathematical solutions for optimal setup, in particular with regard to the selection of a proper reorder-point.

It turns out that not all effects are easy to describe and some even require empirical analysis based on simulation results, where mathematic rigor for the lack of proper models at hand is substituted by empirical outcomes.

2 Problem Description

Dimensioning of inventory management systems always poses a dilemma between availability and cost. Any parameter selection will always require a decision within this dilemma (Lödding, 2016). Literature discusses this extensively, usually though limited to steady state condition. TOC as an example of a reorder-point system with respect to stochastic variation deserves detailed discussion.

2.1 TOC and DBR

The general approach for TOC is shown in Figure 1: A chain of process steps is lined up. The process step with the lowest output rate is the bottleneck. By definition, there will always be such bottleneck. Throughput from the upstream process steps relative to the bottleneck is loading up any buffer in front of it, and the bottleneck throughput limitation is starving the downstream process steps. So the output from the complete process chain is limited by the throughput of the bottleneck. Then it is prudent to synchronize the upstream processes with the bottleneck, otherwise the before mentioned buffer will quickly overflow.

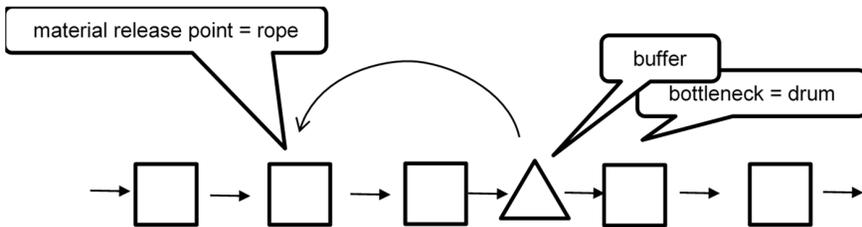


Figure 1: General TOC Approach with DBR

Such control is suggested by TOC through DBR: the bottleneck acts as pacemaker (“drum”). For the sake of optimal overall throughput it must always run. In order to prevent it from material starving, a “buffer” is placed right in front of it. The buffer controls material release upstream (“rope”).

2.2 DBR as Reorder Point Problem

Literature discusses DBR related control issues quite extensively (Cox, J. and Schleier, J. 2010; Lödding, 2016) The DBR system can be reduced to a reorder-point system, see Figure 2: the bottleneck consumes at a rate V/R , material release is issued when the buffer stock level surpasses the reorder point B , then an amount Q is released. The system must be set up such that the accumulated consumption during lead-time DZ remains manageable. For all practical purposes, $Q \geq B$ must always hold, otherwise the system would constantly be starving.

Dimensioning of such systems, in particular, with respect to B have been analyzed and discussed extensively, for instance with the base stock model (Hopp and

Spearman, 2011). Other aspects as for instance, the proper release point location, or the optimum release amount are discussed in literature too (Lödding, 2016). In particular, this requires decisions with respect to logistical positioning for optimum performance, where “optimum” depends on the target variable selected.

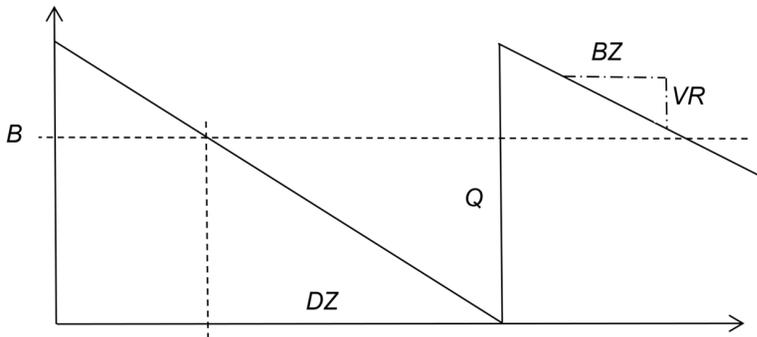


Figure 2: Reorder Point System

2.3 Consumption Variation: VR or BZ

A major difference between the reorder point discussion in literature (Hopp and Spearman, 2011; Chopra and Meindl, 2016) and production reality is that literature suggests use of a consumption rate VR at a fixed cycle time BZ as described in Figure 2. All practical applications on the shop floor will assume a processing time BZ per volume produced (which here is considered one piece at a time only) though. The interpretations for both are quite different:

- In the first case, DZ is an amount of intervals, and in each interval, a batch of parts is consumed.
- In the latter case, DZ is a frequency, at every interval, one part is consumed.

In static condition, the two can easily be derived from one another as reciprocal values. This proves not to be correct under uncertainty.

2.4 Performance Parameter: FR or CSL

The problem is further complicated by different performance measures: The major output parameters often discussed in the literature are CSL and FR (Chopra and Meindl, 2016). CSL measures the replenishment yield whereas FR measures demand yield. From a practical perspective, the customer related FR is much more important than the supplier related CSL . Four different cases can be distinguished now: Observation of CSL or FR and variation of VR or BZ .

First, the typical solution by way of the base stock model for $CSL(B)$ at variation of VR is analyzed: VR is defined by a random variable with expected value μ and standard deviation σ for a given distribution $F(x)$. The parameters describe a pattern for volume; every replenishment cycle is divided into DZ fixed intervals of such demand pattern, all of which considered independent from each other. The optimum reorder point B is then defined as follows:

$$B = F^{-1}(CSL, DZ \mu, \sqrt{DZ} \sigma). \quad (1)$$

Second, the analytical relationship $FR(B)$ at variation of VR is analyzed. Unfortunately, this is more complicated than the above. In fact, a closed analytical solution for the FR under uncertainty is most often limited to special cases like a simple normal distribution of the variation of consumption VR (Hopp and Spearman, 2011). Generally, the following approach is being suggested (Hopp and Spearman, 2011; Zipkin, 2000):

$$FR = 1 - \frac{ESC}{B}, \text{ with } ESC = \int_B^{\infty} (x - B)f(x)dx, \quad (2)$$

utilizing the Expected Shortage Count (ESC), based on the density function $f(x)$ for VR . This can only be resolved for B in special cases.

Though Equation (2) cannot always explicitly be solved, the special case solutions that are available do allow for an interesting observation (Hopp and Spearman, 2011):

- CSL is dependent on replenishment time DZ but independent from replenishment volume Q , for as long as $Q \geq B$ is met.
- FR is dependent on Q , but independent from replenishment time DZ . Therefore, any parameter study with respect to FR will require variation of Q additionally.

Third, $CSL(B)$ at variation of BZ is analyzed: For production, variation in BZ is more common than variation in VR . In order to derive a proper relationship between B and CSL at variation of BZ as opposed to VR above, the statistical experiment must be redefined. Instead of a volume experiment as with the base stock model, now a time experiment must be conducted: CSL is defined by the probability of an uncertain total processing time for a fixed number of parts B not to exceed a given boundary time DZ :

$$CSL = 1 - F(DZ, B, \mu, \sqrt{B} \sigma) \quad (3)$$

with μ and σ now defining the stochastic distribution F of BZ in time. This is unfortunately not quite as elegant if an optimal value of B is sought as no explicit solution is easily available. With some numerical effort, it however does the trick.

The fourth and most interesting combination would describe a relationship between the FR and B at any given variation in BZ . Impact on FR is much more important as a performance criterion than CSL as it reflects customer orientation as opposed to supplier focus. This is however at this time not available and cries for more research.

2.5 Variation in Supply Lead Time DZ and Demand Rate VR

In reality, both supply lead-time DZ and consumption will vary, regardless of whether defined by VR or BZ . For the remainder of the article, neither the BZ case nor the FR analysis can be investigated further as no solution could be offered even for the more simple case of variation in consumption only.

Otherwise literature offers an easy solution for the combined case (Chopra and Meindl, 2016; Hopp and Spearman, 2011): For a consumption process VR described by the expected value μ , and standard deviation σ as above, and a supply process DZ described by the expected value ϱ , and standard deviation η , the overall system behavior is described by

$$\mu_{DZ+VR} = \mu * \rho, \text{ and } \sigma_{DZ+VR} = \sqrt{\rho \sigma^2 + \mu^2 \eta^2}. \quad (4)$$

Here, it is assumed that all consumption cycles are independent and stationary. Then, means are multiplicative and variations are additive and the above equation contains two types of variation: $\sqrt{\rho \sigma^2}$ describes variation within cycles, whereas $\sqrt{\mu^2 \eta^2}$ describes variation between cycles.

3 Simulation Approach

The simulation platform Plant Simulation (PS) (Bangsow, 2011) provides a very comprehensive system for event based queueing systems of large magnitude. Here, the order of magnitude is very small. Nonetheless, it proves a very powerful system for stochastic analysis, and complexity arises from the scope, not from the scale.

The general setup is chosen such that a source supplies pallets full of parts, which are unloaded and stored in the DBR buffer piece-wise, and processed subsequently. Now, all necessary combinations of fixed or variable supply amounts, fixed or variable processing times and fixed or variable processing volumes can be studied. The DBR buffer is under constant observation, and new material is released as soon as buffer level falls below the reorder point level. Throughout the paper it is assumed that replenishment is an on/off condition that cannot be activated several times at once, it is finished with the last part from the pending replenishment cycle being fed into inventory. The drum is implemented by way of a high resolution time triggered function which controls the inventory level and issues new pallets from the source as needed.

3.1 Variation of Demand Processing Time BZ

The most natural approach to modelling a push-oriented supply chain in PS leads to the case of variable processing time BZ as represented by Equation (3), the implementation is shown in Figure 3. Left from buffer B , the process is lot-based; on the right hand side from B , it is executed piece-wise with step size BZ .

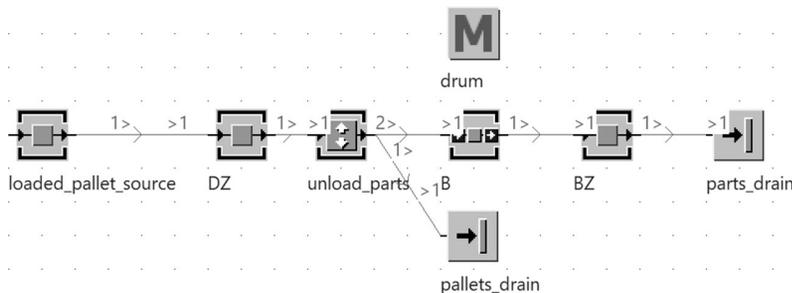


Figure 3: Implementation with variable processing time BZ

3.2 Variation of Demand Rate VR

In case of a demand rate VR variation, the modeling approach is a bit counter-intuitive: The supply side left from B is unchanged. The consumption side is very

different though. At a fixed time rate, the consumption is batch-oriented with a variable number of parts VR . Not always all parts will be available from B , then it is a design choice whether or not to work with a backlog list. Such model corresponds with Equations (1) and (2).

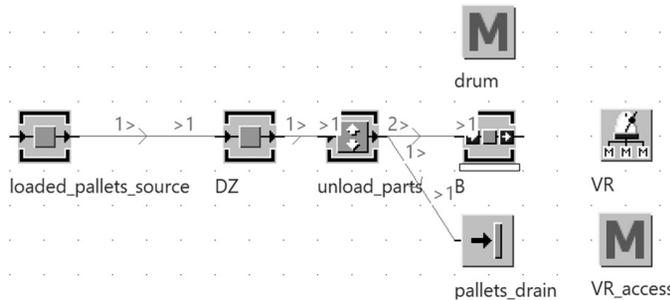


Figure 4: Implementation with variable consumption rate VR

3.3 Means to Measuring FR and CSL

In the corresponding PS model, extra provisions must be taken to be able to collect and analyze relevant statistics:

- For every replenishment cycle the simulation must record whether demand was met during this related cycle, thus impacting CSL.
- For any demand requested from B it must record if the request could be met sufficiently, thus impacting FR.

Either variable is calculated as the sum of successful attempts over the sum of all attempts.

3.4 Stochastic Variability Modeling

In order to introduce variability into the simulation, a proper distribution form has to be chosen. Logistic problems often show one-dimensional exponential behavior (Hopp and Spearman, 2011), in time. To be able to adjust average and standard deviation separately and still be rather true to reality, the logarithmic normal (lognorm) distribution is employed throughout herein:

$$y = e^{\mu_0 + \sigma_0 x} \text{ with } \mu_0 = \ln \frac{\mu}{\sqrt{\left(\frac{\sigma}{\mu}\right)^2 + 1}} \text{ and } \sigma_0 = \sqrt{\ln \left[\left(\frac{\sigma}{\mu}\right)^2 + 1 \right]}. \quad (5)$$

The random variable x is normally distributed with expected value μ_0 and standard deviation σ_0 . The lognorm distribution is limited to $y > 0$, which is a nice property for simulation. Rather unusual, PS uses a definition of the lognorm with expected value μ and related standard deviation σ as opposed to μ_0 and σ_0 . These representations can easily be transformed into either space as shown above.

3.5 Parameter Studies

Throughout the analysis, the impact from different reorder levels B on CSL and FR is analyzed. For this reason, very many different simulations have to be conducted, sweeping through a whole range of values for B , and given the stochastic nature of

the model, every simulation must be carried out a few times. For this, PS's ability to perform simulation study experiments has been used extensively for the below parameter studies.

4 Empirical Reorder Point Analysis

Throughout the studies, the reorder point B has been analysed in single steps from 1 to 100 pieces. All simulations have been carried out with 50.000 cycles.

4.1 Variation of Demand Processing Time BZ

The simulation parameters are chosen as follows: Processing time BZ is lognorm distributed with $\mu = \sigma = 1$ cycles per part (cpp), replenishment time is set to $DZ = 30$ cycles, and demand $VR = 1$ piece per cycle (ppc). The results are shown in Figure 5, comparing the simulation results with the analytical solution from Equation (3). The match is not perfect but acceptable given the large number of numerical operations. Thus, the simulation confirms the above model represented by Equation (3).

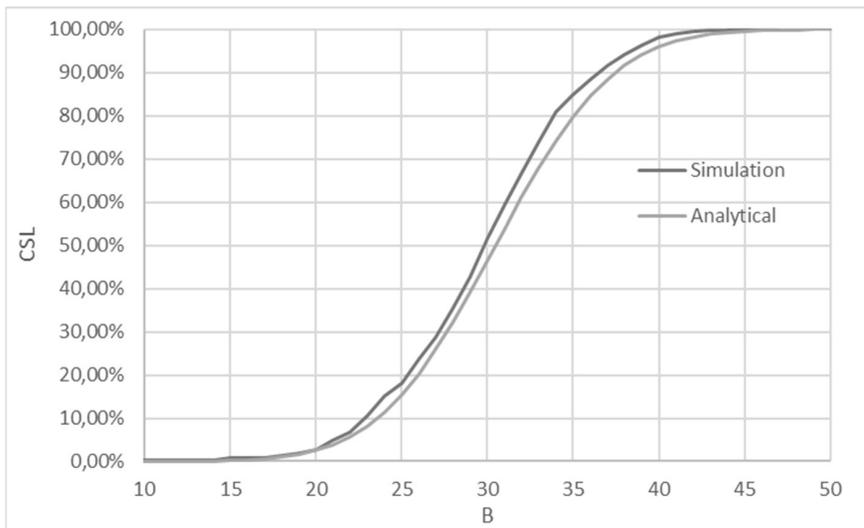


Figure 5: $CSL(B)$ at Variation of Demand Processing Time BZ

4.2 Variation of Demand Rate VR

First, the impact of B on CSL is analyzed, and the following conditions are chosen: $DZ = 8$ cycles, now VR is lognorm distributed with $\mu = \sigma = 5$ ppc. The results are shown in Figure 6, comparing simulation results with the results from Equation (1). Here, the difference between simulation and analytical solution is more pronounced. For further validation, an additional analytical solution is included based on the actual random numbers generated throughout the simulation, which turn out to be $\mu = 5,24$ ppc and $\sigma = 5,13$ ppc instead of 5ppc each, which must be attributed to rounding errors or some implementation error within the simulation tool.

Again the results don't show a perfect match but are considered close enough to validate the model from Equation (1), which is not surprising, as it is a well-

documented textbook case anyways. The parameters of the resulting function $CSL(B)$ as indicated by Equation (1) reverse-engineered from the simulation results show lognorm behavior with $\mu = 5,54\text{ppc}$ and $\sigma = 5,32\text{ppc}$ at a $RMS = 5 * 10^{-4}$ over the 100 data points.

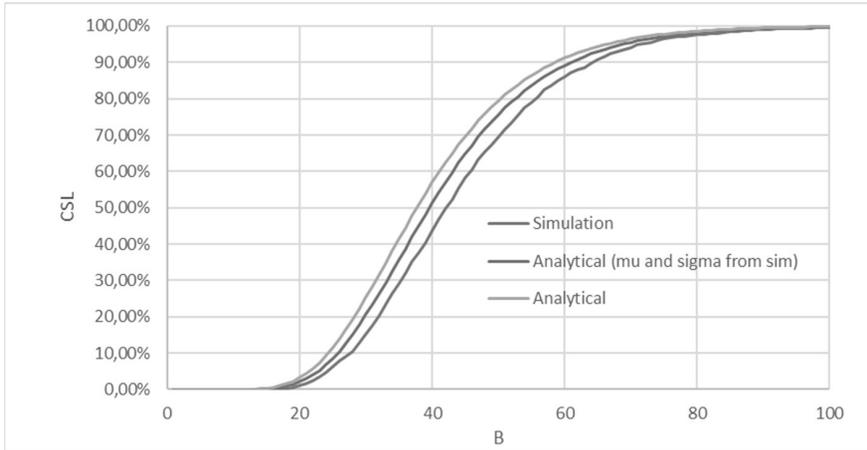


Figure 6: $CSL(B)$ at Variation of Demand Rate VR

Next, the more complex $FR(B)$ case is investigated for which there was no closed analytical solution of Equation (2). It was resolved numerically using a scientific calculation tool for the given parameters based on a parameter sweep similar to the simulation. The comparison of these semi-analytical results with simulation results is shown in Figure 7.

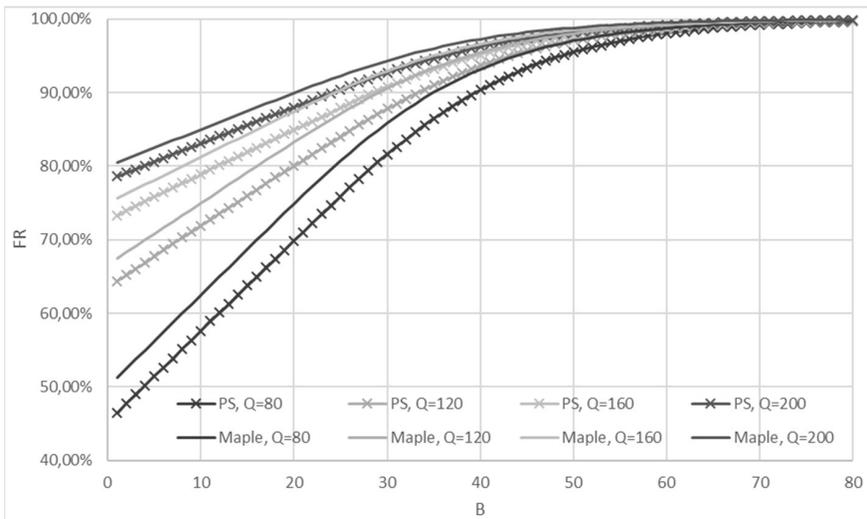


Figure 7: $FR(B)$ at Variation of Demand Rate VR

As stated above, now the replenishment volume Q must also be taken into account, here it is displayed for a range from 80ppc to 200ppc. Again, the match is not perfect but the trend is confirmed and Equation (2) is validated by the experiment.

4.3 Variation in Supply Lead Time DZ and Demand Rate VZ

Finally, the case of variation in both replenishment time DZ and consumption rate VR is analyzed to validate Equation (4). The simulation example from above is being reused: The consumption process VR is defined by a lognorm distribution with $\mu = \sigma = 5\text{ppc}$. The supply process is equally described by a lognorm distribution for DZ with $\varrho = \eta = 8$ cycles. With equation (4) this results in a total system expected value of $\mu_{ges} = 40\text{ppc}$ and $\sigma_{ges} = 42,43\text{ppc}$. A comparison of analytical and simulation results is displayed in Figure 8: An approximation of the simulation result $CSL(B)$ by a lognorm distribution function results in $\mu_{ges} = 50,49\text{ppc}$ and $\sigma_{ges} = 61,7\text{ppc}$ at a RMS of $4,9 \cdot 10^{-3}$ over the 100 data points. Of course, the results are far from identical, but the trend is confirmed. Further analysis for the source of error revealed that again even the random number generation for the distribution functions wasn't accurate, the rest of the deviation to be attributed to error accumulation. These results do however confirm Equation (4).

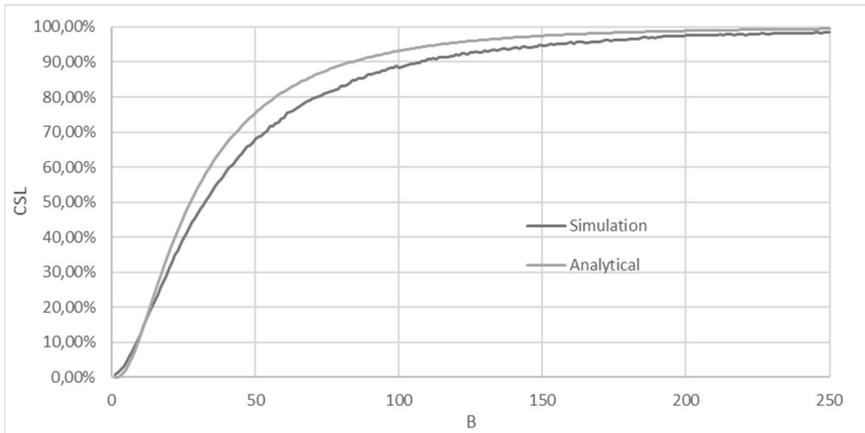


Figure 8: $CSL(B)$ at Variation of Replenishment time DZ and Demand Rate VR

All of the corresponding simulation results have shown substantial deviation from the expected. Overall, the trends could be confirmed, but always with certain unease. This is a continuation of observations already made in another study (Rükgauer; Fabeck, 2020). The random number generation of PS is far from precise and it continues to appear that some mysterious rounding error accumulations would always pollute the outcomes.

5 Conclusion

Main focus of this paper was on the relationship between reorder point B and the resulting logistical performance parameters when analyzing DBR as reorder point system. Among the many interesting questions arising from such approach, here the relationship between reorder point B and the logistical target variables CSL and FR

have been investigated analytically and experimentally by way of stochastic simulation.

Literature predominantly focusses on variation in volume VR and the related impact on CSL and FR. For the latter, no closed explicit solution is available, but this can be resolved using mathematical solvers. Variation in time BZ requires more analytical analysis, It is shown that one is not the reciprocal of the other but requires a completely different stochastic experiment setup. At this point, only a solution with respect to CSL is available. In reality, not only consumption will vary, but also supply, in particular, replenishment time DZ . Literature offers a nice approach to combine VR and DZ , for the combination of BZ and DZ , or even BZ , VR and BZ , so far no solution is available.

Two simulation models have been developed to validate the relationships identified. Simulation of the VR case requires an approach that seems almost counter-intuitive. With some numerical error margin, all formulae have been verified using the simulation models.

Based on this work, some solutions are available to the practitioner trying to accomplish desired logistical performance by selecting a proper reorder point B . A lot of questions remain unanswered though and call for more research, in particular an extension of the models for the most practically relevant case of $FR(B)$ at variation of BZ , and the all-encompassing “super-models” describing $CSL(B)$ and even more importantly, $FR(B)$ at full variation of all of BZ , VR , and DZ at once.

References

- Bangsow, S.: Praxishandbuch Plant Simulation und SimTalk: Anwendung und Programmierung in über 150 Beispiel-Modellen. München: Hanser Verlag 2011.
- Chopra, S.; Meindl, P.: Supply chain management: Strategy, planning, and operation. Boston: Pearson 2016.
- Cox, J.; Schleier, J. (Hrsg.): Theory of constraints handbook. New York: McGraw-Hill Professional 2010.
- Eley, M.: Simulation in der Logistik: Einführung in die Erstellung ereignisdiskreter Modelle unter Verwendung des Werkzeuges "Plant Simulation". Berlin: Springer 2012.
- Goldratt, E.M.; Cox, J.: The goal: A process of ongoing improvement. Great Barrington MA: North River Press 2004.
- Hopp, W.J.; Spearman, M.L.: Factory physics. Long Grove (Illinois): Waveland 2011.
- Lödding, H.: Verfahren der Fertigungssteuerung: Grundlagen, Beschreibung, Konfiguration. Berlin, Heidelberg: Springer Vieweg 2016.
- Rückgauer, A.; Fabeck, G.: Challenges in using the Kingman Equation for Prediction and Validation, <https://opus4.kobv.de/opus4-fhws/frontdoor/index/index/docId/1906>, FHWS Würzburg, 2020, zuletzt geprüft am 28.04.2021.
- Zipkin, P.H.: Foundations of inventory management. Boston: McGraw-Hill 2000.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Die Materialflusssimulation im Kontext eines Produktlebenszyklusmanagements

Material Flow Simulation in the Context of Product Lifecycle Management

Torben Meyer, Carsten Pöge, Volkswagen AG, Wolfsburg (Germany),
torben.meyer@volkswagen.de, carsten.poege@volkswagen.de

Abstract: Product Data Management (PDM) and the associated Product Lifecycle Management (PLM) form the data and process basis for the product development process. Thus, PDM also takes place in production planning, especially in the area of the digital factory, where material flow simulation also has its focus. This paper addresses the problem of how to design links between material flow simulation on the one hand and PDM/PLM on the other. The benefit is a better integration of the material flow simulation into the further operational value creation processes, especially with regard to input data availability, change processes and further use of the simulation results.

1 Einführung

Die Rahmenbedingungen für die Materialflusssimulation in der Automobilindustrie haben sich in den letzten Jahren stark verändert. Eine steigende Anzahl an neuen Fahrzeugmodellen, kürzere Modellzyklen und die Einführung neuer Technologien, für die in der Planung noch wenig Erfahrungswerte vorliegen (frei navigierende FTF, Batteriezellfertigung, etc.), führen zu einer steigenden Anzahl von Simulationsprojekten und der in den einzelnen Projekten zu untersuchenden Planungsvarianten. Gleichzeitig steigt mit der Komplexität der Planungsgegenstände auch die Komplexität der Simulationsmodelle. Um die Arbeitslast für die Simulationsexperten weiterhin beherrschbar zu halten, muss die Projektdurchführung effizienter werden. Eine Möglichkeit, dies zu erreichen, ist der konsequente Einsatz eines Produktdatenmanagements auch im Bereich der Materialflusssimulation, was neben dem Aspekt einer effizienteren Projektdurchführung auch Potentiale im Bereich der Informationssicherheit und bei der Zusammenarbeit in größeren Projektteams bietet.

In diesem Abschnitt erfolgt eine Einführung in das Thema und es wird eine Einordnung in den aktuellen Forschungsstand vorgenommen. Danach werden aus der VDI-Richtlinie 3633 (1993) die Simulationsdaten dem PDM gegenübergestellt. Im

Anschluss werden die zwölf Kernfunktionen von PDM-Systemen nach Eigner und Stelzer (2009) auf die Materialflusssimulation angewendet. Schließlich erfolgt eine Diskussion von industriellen Anwendungsfällen der Materialflusssimulation vor dem Hintergrund des PDM.

Von Fischer et al. (2017, S. 136) wurden Prozess- und Ressourcendaten aus dem PDM-System Teamcenter zum Simulator Plant Simulation zum Zweck der automatischen Modellgenerierung übertragen. Bei Morshedzadeh et al. (2018) wurden im PDM-System Teamcenter Produkt-, Prozess- als auch Ressourcenstrukturen modelliert, die bspw. 3D-Geometrien mit Verfügbarkeitsdaten und Arbeitszeiten beinhalten. Diese Daten wurden ebenfalls zum Simulator Plant Simulation übertragen und importiert. Nach Abschluss der Simulation wurde ein „HTML Report“ als Ergebnisdokument dem PDM-System zur Verfügung gestellt. Von Kühn (2006, S. 1900) wird qualitativ der strategische Nutzen von PLM im Rahmen der Digitalen Fabrik untersucht. Die technische Realisierung einer Schnittstelle zwischen einem PDM-System und einem Simulator beschreibt Bauknecht (2019).

2 Einordnung der Simulationsdaten in das Produktdatenmanagement

Unumstritten sind die Eingangsdaten ein wichtiger Erfolgsfaktor für die Materialflusssimulation. Die „Daten für die Simulation“ werden in der VDI-Richtlinie 3633 (1993, S. 9) strukturiert aufgelistet und sind in Tabelle 1 dargestellt. Gegenübergestellt sind die relevanten Strukturen aus dem PDM. Dabei handelt es sich zuallererst um die Produktstückliste, genauer gesagt um die Produktionssicht auf die Produktstückliste. Die prozessunabhängige Infrastruktur der Fabrik, also die Fabrikstrukturdaten wie z. B. das Layout, werden im PLM-Kontext als PlantBOR organisiert. Alle weiteren Informationen, die insbesondere die zeitliche Dynamik des Simulationsmodells tangieren (z. B. Ablauforganisation), werden als Prozessdaten abgelegt. Eine Stärke von PDM-Systemen ist ausdrücklich die Verwaltung der verschiedenen Strukturen bzw. Stücklisten mit Dokumenten, die im Rahmen der Erstellung eines Unternehmensmodells definiert werden (Arnold et al. 2011, S. 167).

In die Strukturen der Tabelle 1 kann das ganze Simulationsmodell als Dokument abgelegt werden. Unter Dokument verstehen wir dabei datentechnisch „eine Informationseinheit, bestehend aus Datei und Metadaten“ (Schöttner 2017, S. 103). Die Metadaten des Simulationsmodells (Name, Änderungsdatum, Version, etc.) sind dabei im PDM-System indexiert und (entsprechende Berechtigungen vorausgesetzt) lesbar. Für das eigentliche Öffnen der Datei des Simulationsmodells ist jedoch der Simulator notwendig. Dieses Vorgehen ist einzuordnen in das dokumentenbasierte Arbeiten.

In Hinblick auf das modellbasierte Arbeiten kann bei einer tieferen Integrationsstufe das Simulationsmodell auch als Struktur verstanden werden, das heißt, es wird nicht ausschließlich als Ganzes verwaltet (Versionierung, Zugangsberechtigung, etc.), sondern das Simulationsmodell bildet ebenfalls eine Struktur über ableitbare bzw. aussteuerbare Dokumente. Bei bausteinorientierten Simulatoren könnten bspw. Bausteine als einzelne Dokumente modelliert werden. Eine derartige Umsetzung ist die Voraussetzung für die Verbindung einzelner Bestandteile des Simulationsmodells

mit anderen Dokumenten aus dem PDM-System (z. B. Layoutelementen). Die Anwendungsmöglichkeiten werden im folgenden Abschnitt geschildert.

Table 1: Zusammenhang zwischen Simulationsdaten und PDM-Strukturen

Daten für die Simulation (VDI-Richtlinie 3633, 1993)	Relevante Strukturen im PDM
Produktdaten	Manufacturing Bill of Material (MBOM) Produktionssicht auf das Produkt
Fabrikstrukturdaten	Plant Bill of Resource (PlantBOR) Prozessunabhängige Infrastruktur
Auftragseinlastung	
Arbeitszeitorganisation	
Ressourcenzuordnung	
Ablauforganisation	Plant Bill of Process (PlantBOP)
Fertigungsdaten	Fabriksicht der Fertigungsprozesse
Materialflussdaten	
Stördaten	

3 Anwendung der Kernfunktionen eines PDM-Systems auf die Materialflusssimulation

Von Eigner und Stelzer (2009) wurden zwölf Kernfunktionen beschrieben, die in der Mehrzahl der PDM-Systeme implementiert sind. Zu den Kernfunktionen werden ebenda gezählt: Stamm- und Strukturverwaltung, Dokumentenmanagement, Gruppentechnik, Projektmanagement, Workflow-Management, Freigabe- und Änderungsmanagement, Viewing/Redlining/DMU, I/O-Management, Publisher, Archiv und Backup, Datenreplikation sowie Integrationen.

In Mayer et al. (2020) wurden viele Anwendungsgebiete der Materialflusssimulation am Beispiel der Automobilindustrie vorgestellt. In diesem Abschnitt sollen die zwölf Kernfunktionen von PDM-Systemen auf das Anwendungsgebiet der Planung von Karosseriebauanlagen angewendet und der entsprechende Mehrwert dargestellt werden.

Die Funktion **Stamm und Struktur** der Karosseriebauanlagen wird im Rahmen der Produktionsplanung dem PDM-System zur Verfügung gestellt. Unter Struktur versteht man in diesem Zusammenhang die hierarchische Anordnung (PlantBOR) von dem Werk über das Gebäude, die Linie, die Station usw. bis hin zu einzelnen Komponenten, deren CAD-Konstruktion ebenfalls im PDM-System abgelegt wird. Der Stamm (respektive die Stückliste) bildet dabei die Liste über einen definierten (Fabrik-)Bereich. Die Materialflusssimulation wählt entsprechend der simulativen Systemgrenzen ihren Bereich aus. Zu beachten ist dabei, dass die Verknüpfung zwischen Informationsobjekt im PDM-System und Informationsobjekt in der Materialflusssimulation (z. B. einem Simulationsbaustein) modelliert wird und erhalten bleibt (Datendurchgängigkeit), um einen durchgängigen

Informationsaustausch zwischen der Materialflusssimulation als Expertensystem auf der einen Seite und dem PDM-System als zentraler Datendrehscheibe auf der anderen Seite konsistent zu ermöglichen. Auszutauschende Informationen können dabei z. B. die Verfügbarkeiten von einzelnen Anlagen sein, die der Anlagenplaner ermittelt und neben dem Layout im PDM-System abgelegt hat und die vom Simulationsexperten als Eingangsinformationen genutzt werden. Darüber hinaus sollen nach Abschluss der Simulationsstudie relevante Informationen (z. B. die tatsächliche Auslastung einer Anlage) an das entsprechende Informationsobjekt im PDM-System zurückgespielt werden.

Mit der Kernfunktion **Dokumentenmanagement** können für die Prozesskette relevante Dokumente im PDM-System an einem zentralen Ort und im Kontext der Fabrikstruktur (PlantBOR) abgelegt werden. Nach Abschluss der Materialflusssimulation bei Karosseriebauanlagen wird ein Abschlussbericht bzw. eine -präsentation erstellt, der bzw. die im PDM-System verfügbar gemacht werden muss und so späteren Prozessschritten (sowie Umbauten und Brown Field Planung) die Ergebnisse der Simulation zugänglich und die Suche nach Ansprechpartnern einfacher macht.

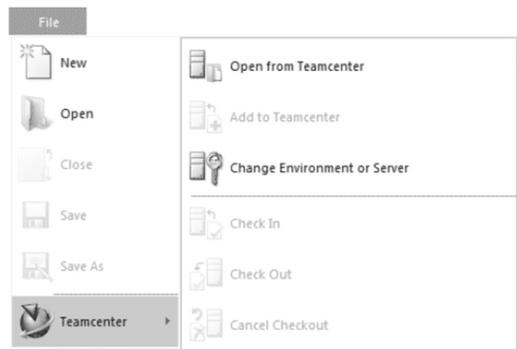


Abbildung 1: PDM-Integration auf Dokumentenebene in Plant Simulation zu Teamcenter

Unter **Integration** versteht man im PLM-Kontext das Zusammenwirken und den Datenaustausch verschiedener Expertensysteme (z. B. Materialflusssimulator) und der betrieblichen Datenbackbones (PLM und ERP). Zuallererst geht es hierbei um die Bereitstellung von Dateioperationen (Öffnen, Speichern, etc.) in der Materialflusssimulation (Abbildung 1), damit die Anwender des Expertensystems sich auf ihre Applikationen konzentrieren können und die Komplexität des PDM-Systems im Tagesgeschäft des Simulationsexperten ausgeblendet werden kann. Auf Basis einer vorhandenen Integration können Abhängigkeiten zu Eingangsdaten (Tabellen mit Verfügbarkeitsdaten, Layouts mit Materialflüssen, etc.) sowie genutzte Simulationsbibliotheken, die im Rahmen eines PLM-Managementansatzes im PDM-System verwaltet werden, modelliert und dem Simulator zugänglich gemacht werden. Im Sinne eines Lebenszyklus sollen dann die Ergebnisdaten (ggf. mit Unterstützung der Kernfunktion Publisher) zugänglich gemacht werden.

Bei Wick et al. (2020, S.30) wird erwähnt, dass sowohl „die Eingangsdaten, als auch die Ergebnisdaten (...) im Simulationsmodell gespeichert werden (müssen). Dafür bietet sich eine zentrale Struktur an, um die Übersichtlichkeit zu erhöhen“. Aus Sicht des einzelnen Prozessschritts „Materialflusssimulation“ ist diese Aussage mit Sicherheit richtig. Aus der Sichtweise eines durchgängigen Prozesslebenszyklusmanagements, insbesondere der Kernfunktion **Freigabe- und Änderungsmanagement**, stellt diese lokale Datenhaltung jedoch eine Gefahr da. Im Falle einer Änderung der Karosseriebauanlage ist bei einer getrennten, lokalen Datenhaltung nicht sichergestellt, dass diese Änderung auch Eingang in die Simulation findet. Daher sollte die geforderte zentrale Datenhaltung im PDM-System stattfinden, auf das alle Beteiligten Zugriff haben. Von der Materialflusssimulation sollte nur eine Referenzierung erfolgen. Ein weiterer Vorteil in diesem Zusammenhang ist, dass Daten in PDM-Systemen einem kontrollierten Änderungsprozess (engl. Engineering Change Management, ECM) unterliegen, in dem alle Beteiligten abgefragt und über Änderungen informiert werden. Dies erleichtert auch die Pflege und Aktualisierung der Daten durch Kollegen ohne Simulations-Know-How und ermöglicht eine Versionierung der Daten, respektive Modelle, sowie die regelkonforme Archivierung. Grundlage für einen automatisierten ECM-Prozess ist natürlich die informationstechnische Vernetzung der beteiligten Parteien.

Im Bereich des Wissensmanagements mit PDM kann die **Gruppentechnik** angewendet werden. Hierbei werden abgelegte Simulationsmodelle oder -ergebnisse mit Schlagworten versehen oder klassifiziert, damit sie z. B. durch eine gezielte Suche schneller gefunden werden können. Das PDM-System dient dabei als Wissensdatenbank. Typische Schlagworte oder Klassifizierungen in der Automobilbranche sind das simulierte Gewerk (z. B. Schlagwort „Inbound Logistik“), das produzierende Werk (Klassifizierungsmerkmal Werk ist gleich „Wolfsburg“), etc. Für alle abgelegten Simulationsmodelle bzw. -ergebnisdokumente kann die Klassifizierung systemseitig als verbindlich konfiguriert werden, so dass ein fester Satz von Klassifizierungsmerkmalen (Werk, Gewerk, Stückzahl pro Stunde, Methodik, usw.) definiert werden muss, um eine spätere Suche zu erleichtern.

Die Materialflusssimulation ist im fachlichen Ablauf vieler Produktionsplanungen fest verankert und findet eine entsprechende Repräsentation im betrieblichen **Projektmanagement** als ein oder mehrere Projektvorgänge, die Vorgänger- und Nachfolgerbeziehungen zu anderen Vorgängen aufweisen. Durch die Implementierung von Projektmanagementmodulen in vielen PDM-Systemen kann eine direkte Verknüpfung zwischen Projektvorgang und den dazugehörigen Dokumenten bzw. fachspezifischen Workflows erfolgen. Dies ist ein klarer Vorteil ggü. dedizierter Projektmanagementsoftware, bei der zumeist eine Trennung zwischen den Inhalten des Projektmanagements (im Sinne von Zeit, Qualität und Kosten) sowie den fachlichen Inhalten der Projektvorgänge (hier das Simulationsmodell bzw. -ergebnisdokument), die häufig auf Netzwerklaufwerken, Dokumentenmanagementsystemen oder lokal gespeichert werden, erfolgt. Durch die Integrationsplattform des PDM-Systems sind das Projektmanagement und die fachlichen Inhalte direkt verbunden und werden nicht getrennt verwaltet.

Ein **Workflow** ist ein verbindlicher und systemgestützter Arbeitsablauf von meist mehreren Prozessbeteiligten auf operativer Ebene. Häufig wird ein Workflow einem Projektvorgang unisono zugeordnet. Mit Beginn des Projektvorgangs

„Materialflusssimulation (Anlage) erstellen“ wird zum Beispiel der gleichnamige Workflow (Abbildung 2) ausgeführt und die zugehörigen Aufgaben werden den Bearbeitern zugewiesen. Im Rahmen von Workflows werden häufig dokumentierte Reviews (z. B. Abstimmung der Eingangsdaten) oder Reifegrade mit Status (Simulation abgeschlossen) durchgeführt.

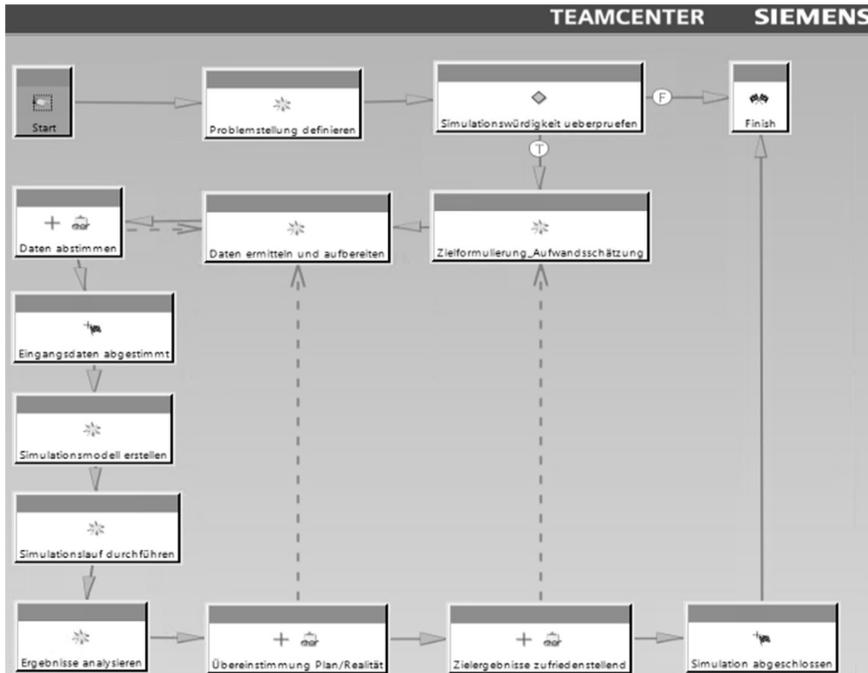


Abbildung 2: Workflow für Materialflusssimulation in Anlehnung an VDI-Richtlinie 3633 (1993) im PDM-System

Eine wichtige Eigenschaft des PLM-Managementansatzes ist das Bereitstellen von Informationen für die gesamte Prozesskette, wovon die Materialflusssimulation nur einen Prozessschritt bildet. Die eigentliche Simulationsanwendung nutzt dabei zumeist ein proprietäres Dateiformat zur Speicherung des Simulationsmodells und der -ergebnisse, so dass weite Teile der Prozesskette diese Dateien nicht öffnen respektive verwenden können (**Viewing**). Die Lösung dafür bilden Autorensystem-unabhängige Neutralformate, zum Beispiel Simulation Data Exchange (SDX, Sly 2001) oder Core Manufacturing Simulation Data (CMSD) mit allen Vor- und Nachteilen, die diese Neutralformate bieten (Bergmann et. al 2011). Dazu zählt auch, dass die Übertragung der simulator-spezifischen Programmiersprachen in Neutralformat in der praktischen Anwendung nicht gelöst ist.

Neben einer Übertragung des Simulationsmodells in Neutralformate ist auch eine Übertragung der 2D- oder 3D-Repräsentation des Simulationsmodells in Neutralformate (JT, VRML, etc.) zum Zweck des Viewings möglich. Um die Simulationsergebnisse für die Prozesskette zu visualisieren, könnten grafische Diagramme oder Kennzahlen direkt im 3D-Raum neben den Anlagen dargestellt werden. Unter Einsatz der PDM-Kernfunktion **Redlining**, womit ursprünglich auf

technischen Zeichnungen mit einem roten Stift Kommentierungen vorgenommen wurden, werden Kommentare in den 3D-Repräsentationen im Neutralformat durch Prozesskettenteilnehmer hinterlassen. Dadurch können Anmerkungen und Kommentare an der Simulation vorgenommen werden. Viewing und Redlining sowie Digital Mock-Up (DMU) sind einige Beispiele für das **Input/Output-Management**. I/O-Methoden, wie das Drucken, sind – insbesondere im Kontext der Materialflusssimulation – zunehmend seltener anzutreffen.

Unter der Kernfunktion **Publishing** versteht man im engeren Sinne das Ableiten von Dokumenten aus Produktmodellen. Beispielweise kann es sich hierbei um die (ggf. automatische) Erstellung einer Bedienungsanleitung für Fahrzeuge handeln. Angesichts von Dokumentationspflichten sowie varianten- und revisionsgenauer Ausleitung et cetera ist diese Aufgabe keinesfalls als trivial anzusehen.

Die Vorgehensweise des Publishings lässt sich auch auf die Materialflusssimulation im Kontext des PLM übertragen. Den Abschluss einer Simulationsstudie bildet klassischerweise ein Abschlussdokument bzw. eine Abschlusspräsentation. Im Vordergrund stehen dabei sicherlich managementgerechte, kompakte Aussagen und Handlungsempfehlungen (so genanntes Executive Summary). Darüber hinaus bestehen die Abschlussdokumente im Beispiel der Karosseriebauten zum großen Teil aber aus Diagrammen und Tabellen, deren manuelle Aktualisierung und Anpassung fehleranfällig wäre. An dieser Stelle können mit Hilfe des Publishings Dokumente aus Simulationsdaten generiert und damit die Nachverfolgbarkeit zwischen Dokument und Modell sichergestellt werden. Eine Umsetzung mit der Bezeichnung „TeamcenterReport (HTML)“ wurde von Bauknecht (2019) für Plant Simulation vorgestellt.

Die Kernfunktionen **Archiv und Backup** sowie **Datenreplikation** sind in zeitgemäßen PDM-Installationen als gegeben anzusehen. Die Archivierung und das Backup beugen Datenverlust vor und sorgen für eine Langzeitverfügbarkeit der Daten, um ggf. gesetzliche Anforderungen zu erfüllen. Die automatische Datenreplikation sorgt für die standortübergreifende und latenzarme Verfügbarkeit von Daten. Besondere Anforderungen der Simulationsdaten, die über die Anforderungen der klassischen Produktdaten hinausgehen, sind nicht zu erkennen. Beide Kernfunktionen sind im Kontext eines professionellen Produktdatenmanagements (auch bei Simulationsdaten) als selbstverständlich anzusehen.

4 Integrationstiefe

Die Erfahrungen aus vielen IT-Projekten haben gezeigt, dass das Wasserfallmodell mit der detaillierten Definition von Anforderungen bis zur vollständigen Realisierung eines komplexen Vorhabens in der Regel nicht vollständig umgesetzt wird, unter anderem weil sich Anforderungen im Projektverlauf ändern. Daher ergibt die Definition von kleineren, in sich abgeschlossenen Umfängen Sinn. In diesem Sinne soll die Einbindung der Materialflusssimulation in ein PDM in diesem Abschnitt in fünf Teilumfänge gegliedert werden (Abbildung 3), deren Integrationstiefe ansteigt. Die Realisierung der einzelnen, aufeinander aufbauenden, Schritte ergibt jeweils einen Mehrwert. Die passende Integrationstiefe muss firmen- und projektabhängig definiert werden.

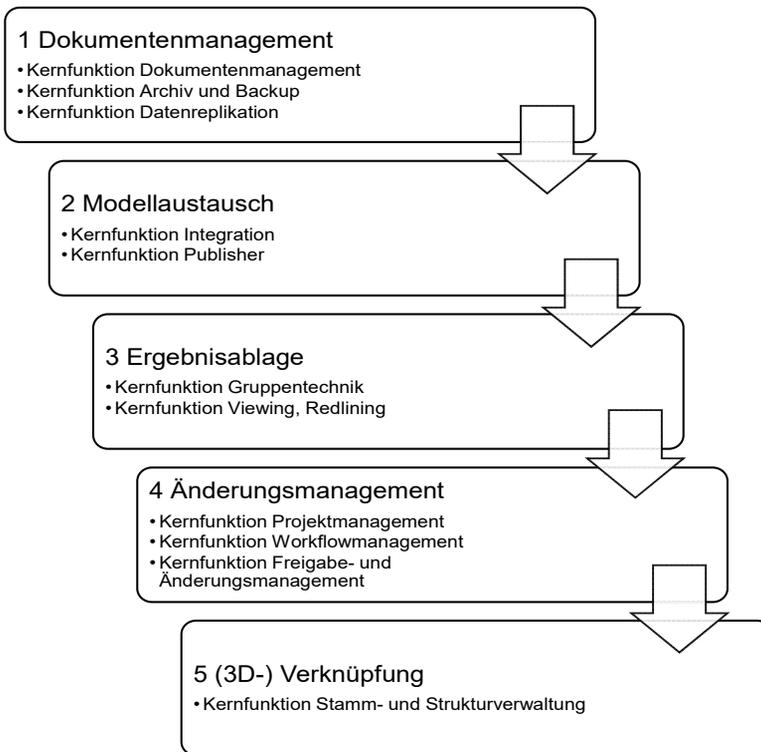


Abbildung 3: Integrationsschritte

Der erste und einfachste Schritt, um die Simulationsmodelle und Ergebnisdokumente im PDM-System verfügbar zu machen, ist das manuelle Datenmanagement. Nachdem das Datenmodell des PDM-Systems um die spezifischen Dateitypen des Materialflusssimulators und die ggf. notwendige Erweiterung um einen dedizierten Objekttyp für diese Simulationsausprägung ergänzt wurde, können die Daten manuell hochgeladen werden. Der Vorteil dieses ersten Schrittes ist, dass die Simulationsdaten zentral verfügbar und revisionierbar sind. Standardfunktionen wie Archivierung, Backup und Datenreplikationen finden automatische Anwendung.

Als Kritik am ersten Schritt könnte vorgebracht werden, dass sich die Simulationsexperten nun mit einem weiteren System zur reinen Datenablage beschäftigen müssen. Als Konsequenz bietet es sich in einem zweiten Schritt an, aus dem Materialflusssimulator unmittelbar ins PDM-System zu speichern (Kernfunktion Integration). Darüber hinaus ist in diesem Zuge auch die automatische Ablage der erzeugten Ergebnisdokumente (Publisher) sinnvoll. Im Sinne des PLM-Gedankens treffen diese Ergebnisdokumente auf größeres Verständnis und Interesse in der Prozesskette als das Simulationsmodell.

Nachdem die Simulationsergebnisse erfasst wurden, kann im dritten Schritt die Verwendung im PDM erfolgen. Auf der einen Seite können für das Wissensmanagement die abgeschlossenen Simulationsprojekte klassifiziert werden (Gruppentechnik), so dass die Suche nach vergleichbaren Simulationsprojekten

schnell und zielgerichtet erfolgen kann. Auf der anderen Seite kann während der Durchführung des Simulationsprojekts aus der Prozesskette heraus mittels Kommentierung des Simulationsergebnisses (Viewing und Redlining) Feedback erfasst und direkt an den Simulationsdaten gespeichert werden.

Die verstärkte Einbindung in die Prozesse ist ein weiterer Integrationsschritt um die betriebliche Wirklichkeit auch in den IT-Systemen abzubilden. Änderungsprozesse, die z. B. Layoutanpassungen, Veränderungen der Verfügbarkeiten, etc. betreffen, werden durch Vorgaben- und Ergebnisrelationen mit der Materialflusssimulation auf Revisionsebene im PDM verknüpft. Auf diesem Wege ist eindeutig nachvollziehbar, welcher Datenstand im Simulationsmodell berücksichtigt wurde. Für die standardisierte Durchführung von Simulationsstudien unter Anwendung von Änderungsprozessen kann auch die Workflow-Funktion des PDM-Systems eingesetzt werden. Auch die Einbettung der Materialflusssimulation in das zentrale Projektmanagement, sofern es die Gesamtorganisation bereits im PDM-Kontext nutzt, ergibt in diesem Schritt Sinn, da Workflow- und Projektmanagementfunktionen eng vernetzt sind.

Der fünfte Schritt beschreibt eine intensive Verknüpfung der 3D-Projektdateien (Layout) mit den Simulationsdaten. Wenn diese Verknüpfung nicht automatisiert im Sinne einer automatischen Modellgenerierung erzeugt wird, ist sie manuell aufwendig und fehleranfällig. Die Vorteile einer durchgängigen Verknüpfung der Stamm- und Strukturdaten zwischen PDM und Simulationsmodell ist, dass z. B. Layoutänderungen übernommen werden können. Darüber hinaus kann die Ableitung der im Projekt verwendeten Betriebsmittel zurückverfolgt werden, so dass z. B. die Verfügbarkeiten der angezogenen Standardbetriebsmittel (Bibliothekselemente) in der Simulation automatisch verwendet werden. Simulationsergebnisse können nicht nur als Ergebnisdokument, sondern mittels Redlining direkt am 3D-Layout annotiert werden. Die Zusammenarbeit zwischen Produktionsplaner und Simulationsexperte erfolgt so ohne die Verwendung von temporären Arbeitsdokumenten und -modellen und steigert so die Effizienz.

5 Zusammenfassung

In diesem Beitrag wurde die Zusammengehörigkeit der Materialflusssimulation und des Produktdatenmanagements auf der Datenebene (Tabelle 1) und der Funktionsebene gezeigt. Darüber hinaus wurde eine schrittweise Einführung (Abbildung 3) mit steigender Integrationstiefe vorgestellt.

Schon durch das aufwandsarme Dokumentenmanagement der Simulationsmodelle und -ergebnisse kann für die Prozesskette ein Mehrwert im Sinne einer besseren Informationsverfügbarkeit geschaffen werden. Die Einbindung in die Änderungsprozesse (vergleichbar zum Produkt) birgt darüber hinaus ebenfalls Potentiale insbesondere für die Darstellung der Abhängigkeiten zu den Simulationseingangsinformationen. Die detaillierte, revisionsgenaue Verknüpfung einzelner Objekte zwischen Produktionsplanung und Materialflusssimulation (fünfte Integrationsstufe) sollte jedoch in Abhängigkeit von den betriebsindividuellen Aufwänden kritisch dem erwarteten Nutzen gegenübergestellt werden.

Insgesamt lässt sich festhalten, dass die Einbindung der Materialflusssimulation in das durchgängige Produktlebenszyklusmanagement und die zentrale Datenhaltung in

PDM-Systemen Vorteile für die gesamte Prozesskette im Hinblick auf Informationsverfügbarkeit bietet. Die Simulations- und Projektergebnisse können besser an die Prozesspartner kommuniziert werden, den Simulationsexperten erwarten akzeptable Aufwände bei der Bereitstellung der Simulationsmodelle sowie Vereinfachungen bei der Beschaffung der Simulationseingangsdaten. Eine Einbettung in das Unternehmensdatenmodell ist erstrebenswert.

Ergebnisse, Meinungen und Schlüsse dieser Publikation sind nicht notwendigerweise die der Volkswagen Aktiengesellschaft.

Literaturverzeichnis

- Arnold, V.; Dettmering, H.; Engel, T.; Karcher, A.: Product Lifecycle Management beherrschen. Berlin, Heidelberg: Springer Berlin Heidelberg 2011.
- Bauknecht, R., 2019: Teamcenter Wizard. Tecnomatix Plant Simulation Worldwide User Conference, 2019.
- Bergmann, S.; Stelzer, S.; Strassburger, S.: Initialization of simulation models using CMSD. In: Jain, S.; Creasey, R.R.; Himmelspach, J.; White, K.P.; Fu, M. (Hrsg.): 2011 Winter Simulation Conference, 2011, S. 2223–2234.
- Eigner, M.; Stelzer, R.: Product Lifecycle Management: Ein Leitfaden für Product Development und Life Cycle Management. Berlin, Heidelberg: Springer Berlin Heidelberg 2009.
- Fischer, J.; Willers, R.; Sinnwell, C.: Materialflusssimulation für die CTPS Entwicklung. In: Eigner, M.; Koch, W.; Muggeo, C. (Hrsg.): Modellbasierter Entwicklungsprozess cybertronischer Systeme. Berlin, Heidelberg: Springer Berlin Heidelberg 2017, S. 131–138.
- Kühn, W.: Digital Factory - Simulation enhancing the product and production engineering process. In: Perrone, L.F.; Wieland, F.P.; Liu, J.; Lawson, B.G.; Nicol, D.M.; Fujimoto, R.M. (Hrsg.): Proceedings of the 2006 Winter Simulation Conference, 2006, S. 1899–1906.
- Mayer, G.; Pöge, C.; Spieckermann, S.; Wenzel, S. (Hrsg.): Ablaufsimulation in der Automobilindustrie. Berlin, Heidelberg: Springer Berlin Heidelberg 2020.
- Morshedzadeh, I.; Oscarsson, J.; Ng, A.; Aslam, T.; Frantzen, M.: Multi-level management of discrete event simulation models in a product lifecycle management framework. *Procedia Manufacturing* 25 (2018) 7–8, S. 74–81.
- Schöttner, J.: Umsatz gut, Rendite mangelhaft - das Kostenproblem der Fertigungsindustrie: Warum IT, Digitalisierung, PLM & Co. allein nichts ändern - Ursachen und Lösungen. München: Hanser 2017.
- Sly, D.; Moorthy, S.: Simulation data exchange (SDX) implementation and use. In: Peters, B.A.; Smith, J.S.; Medeiros, D.J.; Rohrer, M.W. (Hrsg.): 2001 Winter Simulation Conference, 2001, S. 1473–1477.
- VDI: Richtlinie 3633 Blatt 1 - Simulation von Logistik-, Materialfluß- und Produktionssystemen.
- Wick, C.; Lüdemann, M.; Mayer, G.: Ablaufsimulation von Karosseriebauten. In: Mayer, G.; Pöge, C.; Spieckermann, S.; Wenzel, S. (Hrsg.): Ablaufsimulation in der Automobilindustrie. Berlin, Heidelberg: Springer Berlin Heidelberg 2020, S. 21–33.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Optimierung der Auslastung eines Matrix-Montage-Systems durch Konzeptionierung und Implementierung eines Genetischen Algorithmus für Maschinenbelegungsplanung

Optimization of the utilization of a matrix assembly system by conceptual design and implementation of a Genetic Algorithm for Machine Scheduling

Jana Purgander, Julius Lohse, Johannes Hinckeldeyn, Technische Universität Hamburg, Institut für Technische Logistik, jana.purgander@tuhh.de, julius.lohse@tuhh.de, johannes.hinckeldeyn@tuhh.de

Abstract: Individualization and an increasing number of variants confront manufacturing systems with new challenges. Due to their flexibility, matrix assembly systems have shown great potential to solve many resulting difficulties. However, the control of such production systems is very complex due to a large number of possible scheduling plans for the products to be manufactured. Decentralized agent controllers can route orders through the production system while also observing target values such as utilization maximization. However, a full utilization is unlikely, because there is no consideration of the whole system. In this article, a Genetic Algorithm is used to create a machine scheduling plan for a matrix assembly system. This can be used to implement a centralized control system that optimizes the production system as a whole in terms of utilization. Relationships between parameters of the algorithm and achievable workloads are to be established. The results for a practical use case are evaluated and recommendations for further experiments are derived.

1 Einleitung

Das Phänomen der Mass Customization führte zu einer Nachfrage nach neuen Fertigungssystemen, die eine Vielfalt an unterschiedlichen Produkten effizient herstellen können (Koren 2010). Dafür eignet sich die Auftragsfertigung, bei der modulare Komponenten entsprechend den Kundenanforderungen verbaut werden und so eine hohe Bandbreite an Kundenwünschen abgedeckt werden. Die dafür notwendigen Montagelinien mit Taktzeiten bewähren sich, wenn Arbeit in einzelne Arbeitsgänge mit gleichen Durchlaufzeiten aufgeteilt werden. Jedoch erhöht sich durch individuelle Kundenbestellungen die Variantenvielfalt drastisch. Dies führt auch zu einer zunehmenden Variation des Arbeitsinhalts für jeden Auftrag. In der

starrten Struktur eines klassischen Fließbandes wird mit Taktzeiten gearbeitet und die Aufträge warten so lange, bis die Bearbeitung an allen anderen Stationen abgeschlossen ist, bevor sie die nächste Station ansteuern können. Bei uneinheitlichen Bearbeitungszeiten führt dies zu Leerlaufzeiten und damit zu einer geringeren Auslastung und Effizienz der einzelnen Stationen und folglich der gesamten Montagelinie. Die daraus resultierende geringere Auslastung führt zu unnötigen Kosten und Effizienzverlusten.

Die veränderten Marktbedingungen motivieren die Untersuchung des Konzepts der Matrix-Montagesysteme (MMS). Dieses Konzept verspricht, die Produktivitätsstärke klassischer Montagelinien mit der Flexibilität der Werkstattfertigung zu kombinieren (Schönemann et al. 2015). Dies wird durch einen flexiblen Materialfluss zwischen den Stationen erreicht. Die Fähigkeiten der Stationen werden durch Redundanz und Mehrfachbelegung so konfiguriert, dass Arbeitsinhalt und Kapazität für eine Reihe von Produkten ausgeglichen werden können (Greschke 2015). So sind MMS eine konzeptionelle Antwort auf die tatsächlichen Herausforderungen der Mass Customization. Gleichzeitig steigt die Komplexität der Steuerung im Vergleich zu anderen Fertigungstypen (Greschke und Herrmann 2014).

Für die Planung der Maschinenbelegung von MMS kommen heute vorwiegend simple Heuristiken zum Einsatz (siehe Stand der Technik). In der Produktionsplanung und -steuerung finden fortgeschrittene Algorithmen Anwendung, wie beispielsweise Genetische Algorithmen (GA). Diese konnten bei anderen Produktionssystemen Vorteile gegenüber Heuristiken vorweisen (Aytug et al, 2003). Es stellt sich die Frage, ob Metaheuristiken wie GA auch für MMS Vorteile mit sich bringen verglichen mit simplen Heuristiken. Dies zu untersuchen ist Ziel dieser Arbeit. Um diese Frage zu untersuchen, wird in diesem Artikel zunächst der Stand der Technik vorgestellt. Anschließend werden die Vorgehensweise und das Konzept des GA erläutert, die Ergebnisse der Implementierung beschrieben und ausgewertet und letztlich diskutiert.

2 Stand der Technik

Das Konzept der Matrix-Montage-Systeme zeichnet sich durch ein schachbrettähnliches Layout aus, bei welchem die einzelnen Arbeitszellen (AZ) untereinander flexibel durch Fahrwege vernetzt sind. Häufig kommen Fahrerlose Transportfahrzeuge (FTF) zum Einsatz, sodass für jeden Auftrag ein individueller Produktionspfad durch das System entsteht (Popp, 2018, S.72). So kann jeder Auftrag entsprechend seiner Variante und der benötigten Arbeitspakete (AP) eine flexible Sequenz von Arbeitszellen anfahren. Diese Vielzahl möglicher Produktionspfade bedeutet einen erhöhten Steuerungsaufwand für die Auftrags- und Belegungsplanung.

Bisher kamen dafür hauptsächlich Heuristiken zum Einsatz (Greschke 2015, Popp 2018). Diese können zwar keine optimalen Lösungen garantieren, erreichen jedoch in kurzer Zeit bereits zulässige Lösungen. Optima werden dabei nicht gezielt angestrebt. Es bleibt also fraglich ob diese Heuristiken Optimierungspotenziale bei MMS ungenutzt lassen. Für die Maschinenbelegungsplanung in komplexen Produktionssystemen, wie z.B. Werkstattfertigungen, kommen auch Metaheuristiken wie Ameisenalgorithmen und Schwarmalgorithmen zum Einsatz (Echler Minguillon, 2020). Auch bei der Maschinenbelegungsplanung eines MMS handelt es sich um ein komplexes kombinatorisches Optimierungsproblem. Bei vergleichbaren Problemstellungen konnten Genetische Algorithmen (GA) bereits

erfolgsversprechende Erfolge vorweisen (Aytug et al, 2003). Es stellt sich daher die Frage, ob durch den Einsatz eines GA eine Verbesserung der erreichbaren Auslastung eines MMS im Vergleich zu den bekannten Heuristiken erzielt werden kann, indem eine effektivere Belegungsplanung realisiert wird.

3 Vorgehensweise und Konzept des Genetischen Algorithmus

In der durchgeführten Simulationsfallstudie wurde der Umbau eines existierenden Produktionssystems von einer taktgesteuerten Linienfertigung hin zu einem MMS betrachtet. Für den geplanten Umbau wurde ein Konzept eines Matrix-Montage-Systems erstellt, anhand dessen der Vergleich zwischen zwei Heuristiken (Push/Pull) und eines GA stattfand. Anschließend wurden Simulationen sowohl für die Heuristiken als auch für den GA durchgeführt und diese miteinander verglichen. Der nachfolgende Abschnitt stellt die Fallstudie und die verwendeten Algorithmen und Heuristiken vor.

Die bisherige Produktion beruht auf dem Konzept der getakteten Linienfertigung (siehe Abbildung 1) mit 25 Stationen und einer Auslastung von 83 Prozent. Es handelt sich hier um ein aus einer realen Fertigung abgeleitetes, idealisiertes Modell eines taktgesteuerten Linien-Montage-Systems. Dieses wurde als Vergleichsmaßstab für die Bewertung des konzeptionellen Matrix-Montage-Systems verwendet.

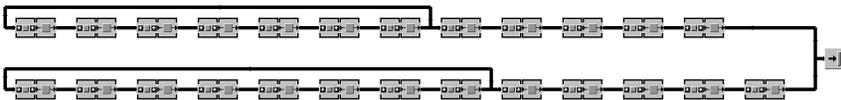


Abbildung 1: Layout der Fließfertigung

Das betrachtete Layout des MMS enthält aus Platzgründen nur 22 Arbeitszellen, an welchen jeweils ein Werker arbeitet (siehe Abbildung 2). Links und rechts vom Arbeitsplatz eines Werkers befinden sich jeweils ein Puffer, auf welchem ankommende Aufträge bearbeitet werden können oder auf ihre Bearbeitung warten. In einem Arbeitsschritt an einer AZ wird genau ein Arbeitspaket am Auftrag durchgeführt. Unterschiedliche AZ können unterschiedliche AP durchführen. Dabei kann jede AZ stets mehrere Arbeitspakete ausführen. Dies hat den Hintergrund, dass Redundanzen und Mehrfachbelegungen im System für eine größere Robustheit gegenüber Störungen sorgen.

Die im Layout realisierten Wege sind zu keinem Zeitpunkt gegenläufig, sondern stets nur in eine Richtung befahrbar. Die zu bearbeitenden Fahrzeuge, welche die Aufträge im Anwendungsfall des Fahrzeugherstellers repräsentieren, werden dabei auf Fahrerlosen Transportsystemen bewegt. Des Weiteren ist es den FTF nicht möglich, über freie Puffer zu fahren, selbst wenn dadurch die Wegstrecke zwischen dem Start- und Zielpunkt des Auftrags minimiert werden würde. Daraus lässt sich schließen, dass Aufträge auf dem Weg von einer Arbeitszelle zur nächsten stets denselben kürzesten Weg nutzen können. Die Zeit, die für diesen kürzesten Weg benötigt wird, wird als Wegzeit zwischen den Arbeitszellen bezeichnet.

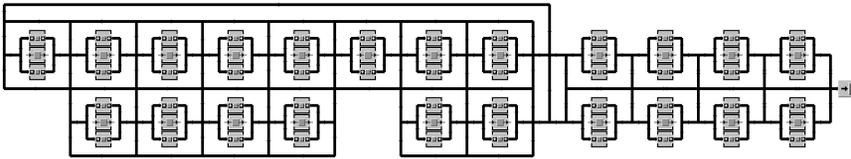


Abbildung 2: Layout der MMS

Für diese Simulationsfallstudie zur Untersuchung des MMS-Layouts (siehe Abbildung 2) wurden zwei unterschiedliche Heuristiken implementiert. Bei der **Push**-Steuerung wird für jeden Auftrag nach Beendigung des aktuellen Bearbeitungsschritts für alle möglichen nachgelagerten Stationen nun der frühestmögliche Bearbeitungszeitpunkt ermittelt und verglichen. Die Station, welche das Fahrzeug am frühesten bearbeiten kann, erhält den Zuschlag. Bei der **Pull**-Steuerung bestellt eine Arbeitsstation, die in absehbarer Zeit kein Fahrzeug mehr zur Bearbeitung hat, ein neues Fahrzeug bei ihren vorgelagerten Zellen. Diejenige vorgelagerte Station, welche am frühesten der Anfrage nachkommen kann, leitet ihren Auftrag entsprechend weiter. Die restlichen Zellen werden über die Hinfälligkeit der Anfrage informiert. Sollte keine vorgelagerte Zelle der Anfrage nachkommen können, so wartet die betrachtete Arbeitsstation so lange, bis sich dieser Zustand ändert.

Bei der Verwendung eines Genetischen Algorithmus zur zentralen Steuerung erhalten die FTF von einer zentralen Recheneinheit die Information, welche Arbeitszelle sie als nächstes anfahren, damit ihr Auftrag dort bearbeitet werden kann. Es wird ein Maschinenbelegungsplan erstellt, in welchem festgehalten wird, welchen Produktionspfad die Aufträge durch das System nutzen sollen. Es kann somit von einem Leitstand gesprochen werden. Der GA stellt eine zentrale Schnittstelle zwischen allen FTF und Arbeitsstationen dar. Im Leitstand werden alle verfügbaren Informationen gebündelt und ausgewertet, sodass eine Betrachtung des Gesamtsystems realisiert werden kann. Anders als bei den Heuristiken kann eine Vielzahl an Parametern einbezogen werden, wie etwa die Auslastung der Produktionsstationen. Der GA erstellt zunächst eine Population von 800, 1600 oder 3200 gültigen Maschinenbelegungsplänen. Für jeden Plan (Individuum) wird dann die erreichbare Auslastung ermittelt und als Fitnesswert hinterlegt. Entsprechend des Fitnesswertes werden die Pläne für den Übergang in die nächste Generation selektiert. Dabei wurden drei unterschiedliche Selektionsmechanismen betrachtet:

- **RouletteSelektion:** Selektionswahrscheinlichkeit ist proportional zum Anteil der Fitness des Individuums
- **RangSelektion:** alle Individuen werden ihrer Fitness nach sortiert; Selektionswahrscheinlichkeit ist proportional zum Rang
- **WettkampfSelektion:** es werden zufällig zwei Individuen für einen Wettkampf gegeneinander ausgewählt; das fittere Individuum gewinnt mit einer Wahrscheinlichkeit von 75% und wird selektiert

Die selektierten Individuen werden dann miteinander rekombiniert, indem jeder Nachkomme von zwei der selektierten Individuen bis zu einem zufälligen Punkt die Erbinformationen des einen Elternteils und danach die Informationen des anderen

Elternteils übernimmt (1-Punkt-Crossover) (Buttelmann und Lohmann, 2004). Des Weiteren werden die fittesten 50 oder 100 Individuen einer Generation direkt in die Folgegeneration übernommen. Man spricht hierbei vom Elitismus. Es entsteht eine neue Generation mit der gleichen Individuenanzahl wie die Ausgangspopulation. Es folgt die Mutation der rekombinierten Individuen der neuen Generation, bei welcher jedes Individuum mit einer Wahrscheinlichkeit von 15 Prozent, 30 Prozent oder 60 Prozent mutiert. Dabei wird zufällig eine Maschinenzuweisung im Plan geändert. Jeder Simulationsdurchlauf bestand aus 750 Iterationen, sodass bei jedem Durchlauf 750 Generationen entstanden. Bei jeder Kombinationsmöglichkeit wurden fünf Durchläufe durchgeführt. Die Ergebnisse des GA wurden anschließend mit den Ergebnissen der Heuristiken verglichen.

4 Ergebnisse und Auswertung

Der Vergleich von GA und Heuristiken wurde anhand einer Simulation durchgeführt. Beim GA werden 50 Aufträgen betrachtet, welche in einer festen Auftragsfreigabereihenfolge aus den vorgelagerten Produktionsprozessen in das MMS eingeschleust werden. Jeder Auftrag hat eine Menge an zwölf Arbeitspaketen (AP), die bis zu seiner Fertigstellung durchgeführt werden müssen. Die Abfolge der durchnummerierten AP ist dabei fest und kann nur in aufsteigender Reihenfolge durchgeführt werden (AP1 – AP2 – AP3 – ... – AP12). Die Bearbeitung der AP findet auf 22 Arbeitszellen (AZ) mit einem Werker und zwei Puffern je AZ statt. Entsprechend wurden Redundanzen und Mehrfachbelegungen geschaffen (siehe Abbildung 3). Störungen wurden dabei außer Acht gelassen.

Für den beschriebenen Anwendungsfall wurden diese Heuristiken in Form einer Simulation erstellt, welche in Tecnomatix PlantSimulation 13.1 implementiert wurde. Dabei erreichte die **Push**-Steuerung eine Maschinenauslastung von 94,09% und die **Pull**-Steuerung 93,08%. Das GA Konzept wurde aus Gründen der Laufzeit in Java implementiert, erhält seine Inputparameter jedoch weiterhin aus dem existierenden PlantSimulation-Modell. Zu diesen Inputparametern gehören:

- Wegzeiten für die Strecken zwischen den Maschinen
- Prozesszeiten für die Bearbeitung der jeweiligen Arbeitspakete
- Auftragsinformationen (individueller Code, Materialtyp)
- Maschineninformationen (AP-Fähigkeiten)

	AP_0	AP_1	AP_2	AP_3	AP_4	AP_5	AP_6	AP_7	AP_8	AP_9	AP_10	AP_11	AP_12
AZ_1	1.00	1.00											
AZ_2		1.00	1.00										
AZ_3		1.00	1.00										
AZ_4			1.00	1.00									
AZ_5			1.00	1.00									
AZ_6				1.00	1.00								
AZ_7				1.00	1.00								
AZ_8					1.00	1.00							
AZ_9					1.00	1.00	1.00						
AZ_10						1.00	1.00						
AZ_11							1.00	1.00					
AZ_12							1.00	1.00					
AZ_13								1.00	1.00				
AZ_14								1.00	1.00				
AZ_15									1.00	1.00			
AZ_16									1.00	1.00			
AZ_17										1.00	1.00		
AZ_18										1.00	1.00		
AZ_19											1.00	1.00	1.00
AZ_20											1.00	1.00	1.00
AZ_21												1.00	1.00
AZ_22												1.00	1.00

Abbildung 3: Redundanzen und Mehrfachbelegungen im System

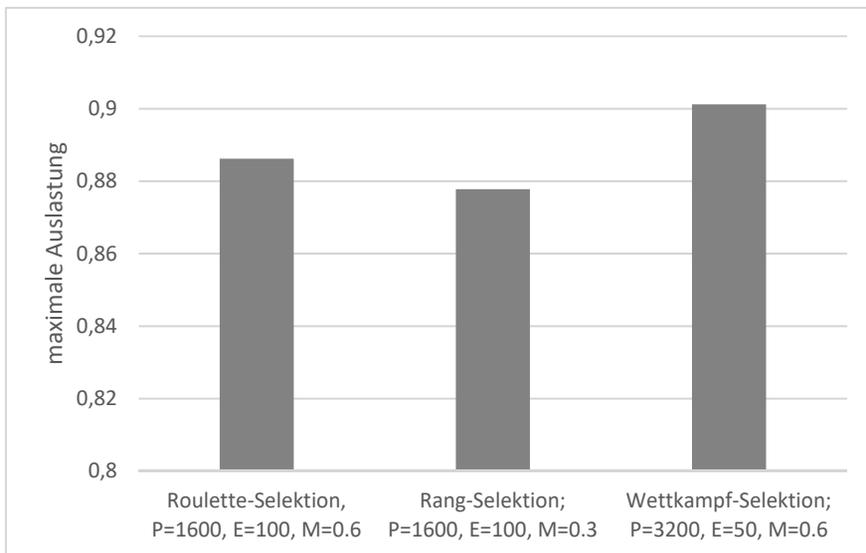
Zur Übermittlung dieser Informationen wird ein Web-Socket verwendet, bei welchem das Modell als Client und der in Java implementierte GA als Server agiert. Im PlantSimulation-Modell werden die Informationen von SimTalk in JSON geparkt, über das Socket übermittelt und anschließend von JSON wieder in Java übertragen. Die Informationen stehen dann dem GA für seinen weiteren Verlauf zur Verfügung.

Die Bearbeitung der Inputinformationen durch den Genetischen Algorithmus variiert durch die Nutzung unterschiedlicher Parameterkombinationen, welche verschiedene Ergebnisse liefern. Dabei zeigten sich für die verwendeten Selektionsmechanismen jeweils vorteilhafte Parameterzusammenstellungen. Tabelle 1 zeigt die besten Ergebnisse der jeweiligen Selektionsmechanismen und die entsprechenden Parameterkombination aus Populationsgröße (P), Elitenanzahl (E) und Mutationsrate (M). Die besten Ergebnisse pro Parameterkombination sind jeweils in fett gekennzeichnet.

Table 1: Maximale Auslastung je Selektionsansatz

Parameterkombination	Roulette-Selektion	Rang-Selektion	Wettkampf-Selektion
P = 800, E = 100, M = 0,3	0,873	0,862	0,874
P = 800, E = 100, M = 0,6	0,886	0,876	0,885
P = 1600, E = 100, M = 0,3	0,874	0,878	0,893
P = 1600, E = 100, M = 0,6	0,886	0,858	0,899
P = 3200, E = 50, M = 0,3	0,851	0,813	0,900
P = 3200, E = 100, M = 0,15	0,880	0,833	0,887
P = 3200, E = 100, M = 0,3	0,866	0,853	0,900
P = 3200, E = 100, M = 0,6	0,874	0,838	0,901

Die Wettkampf-Selektion weist bei vorteilhaften Parameterkombinationen die erfolgreichsten Ergebnisse auf und soll deshalb als Vergleichswert mit den Heuristiken dienen (siehe Abbildung 4).


Abbildung 4: Ergebnisse der Selektionsmechanismen

Nach Betrachtung der Ergebnisse des GA lässt sich festhalten, dass auch die besten erreichten Ergebnisse bei der Wettkampfselktion mit manchen Durchläufen über 90,1 Prozent Auslastung dennoch schlechter sind als die realisierten Werte durch Push- und Pull-Steuerung, siehe Abbildung 5. Dort konnten Maschinenauslastungen von 93,08 Prozent und 94,09 Prozent erreicht werden. Wird jedoch in Betracht gezogen, dass der GA lediglich mit einer begrenzten Anzahl an Parameterwerten und

-kombinationen lief, liegt die Vermutung nahe, dass die erreichbare Auslastung durch andere Parameterwerte noch weiter erhöht werden kann. Es sind weitere Experimente notwendig, um dahingehend eine genauere Aussage treffen zu können.

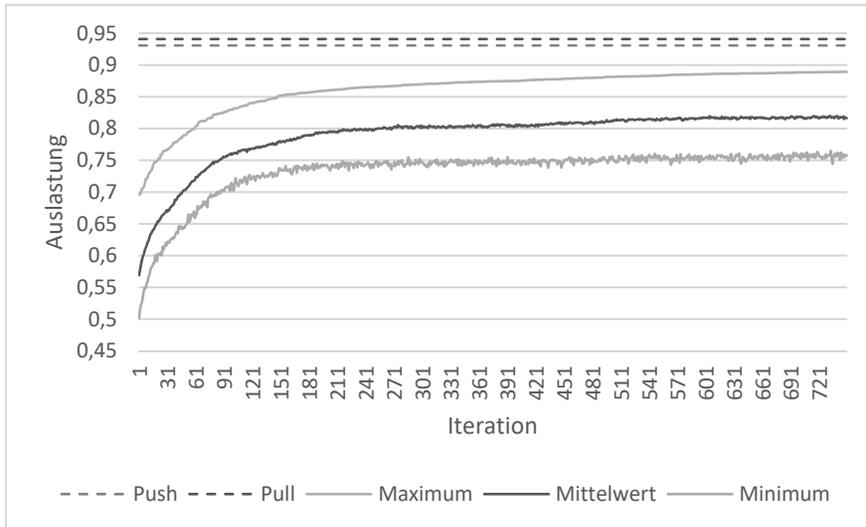


Abbildung 5: Ergebnisse der Parameterkombination Populationsgröße $P = 3200$, Elitenanzahl $E = 50$, Mutationsrate $M = 0.6$ bei der Wettkampf-Selektion

Es können jedoch Empfehlungen getätigt werden, welche Werte für die Parameter abhängig von der Selektionsmethode noch weiter betrachtet werden sollten. Bei Verwendung der Roulette-Selektion als Selektionsmethode erwies sich der höhere Wert für die Elitenanzahl als vorteilhaft. In weiteren Experimenten könnte untersucht werden, wie stark die Elitenanzahl ansteigen kann, bis der positive Effekt aufgrund zu geringer Diversität wieder schwindet. Vermutlich besteht eine Abhängigkeit dieses Maximums zur gewählten Populationsgröße. Da sich im Zusammenspiel mit vielen Eliten auch eine hohe Mutationsrate als vorteilhaft hinsichtlich der erreichbaren Fitnesswerte erwiesen, sind auch in diesem Zusammenhang weitere Experimente notwendig. Möglicherweise führt mit steigender Elitenanzahl auch eine weiter steigende Mutationsrate zu besseren Ergebnissen.

Wurde die Rang-Selektion gewählt, so erwies sich eine geringe Populationsgröße als vorteilhaft. Inwieweit sich dies mit anderen maximalen und minimalen Erwartungswerten verändert, muss untersucht werden. Durch einen anderen maximalen und minimalen Erwartungswert lässt sich regulieren, wie viel größer die Selektionswahrscheinlichkeit des fittesten Individuums im Vergleich zum am wenigsten fitten Individuum ist. Durch einen höheren maximalen Erwartungswert lässt sich der Selektionsdruck signifikant erhöhen. Vor allem im Zusammenhang mit höheren Elitenanzahlen können hier vermutlich noch bessere Auslastungen realisiert werden.

Die Wettkampf-Selektion führte bei den getesteten Werten für die unterschiedlichen Parameter zu den besten Ergebnissen. Vereinzelt wurden Maschinenauslastungen

über 90% erreicht, was den Werten der Push- und Pull-Steuerung von allen Ergebnissen am nächsten ist. Die Auswirkung anderer Siegwahrscheinlichkeiten neben 75% muss in weiteren Durchläufen erfasst werden und Populationen mit mehr als 3200 Individuen könnten zu besseren Ergebnissen führen.

5 Diskussion des Vergleichs

Die vorgestellten Ergebnisse zeigen nicht, dass der Genetische Algorithmus den Heuristiken bei der Maschinenbelegungsplanung von Matrix-Montage-Systemen überlegen ist. Letztere erreichen Maschinenauslastungen von 93,08 Prozent und 94,09 Prozent. Der GA hingegen konnte bei der beschriebenen Umsetzung Auslastungen von 90,01 Prozent realisieren. Der Grund für das schlechtere Abschneiden des GA deutet darauf hin, dass nur eine kleine Anzahl an Lösungen betrachtet wird. Hinzu kommt die Tatsache, dass der GA deutlich rechenaufwändiger ist als die Heuristiken. In Anbetracht der Rechenzeit liefern Heuristiken anscheinend schon gute Ergebnisse und sind deshalb praktikabler. Dennoch ist durch weitere Optimierung und Parameterisierung nicht ausgeschlossen, dass noch eine bessere Lösung mit dem GA gefunden werden kann. Dahingehend besteht ein Bedarf weiterer Verbesserung auch hinsichtlich der Rechenzeit.

6 Fazit und Ausblick

Es konnten Zusammenhänge zwischen den Selektionsmechanismen und den Kombinationen der Modellparameter hergestellt werden. So unterscheidet sich die vorteilhaften Populationsgrößen je nach Selektionsmechanismus, während eine höhere Elitenanzahlen und Mutationsraten stets positive Auswirkungen zur Folge haben. Insgesamt wurden die besten Ergebnisse mit der Wettkampfselektion erreicht, welche bei einigen Parameterkombinationen Auslastungen von über 90% in Aussicht stellt. Die vergleichenden Heuristiken erreichten Maschinenauslastungen von 93,08 Prozent und 94,09 Prozent. Der GA schnitt folglich schlechter ab, birgt aber aufgrund seiner Vielzahl an Parameterkombinationen weiteren Forschungsbedarf. Aufgrund der erkannten Zusammenhänge zwischen Parameterwerten und den erreichbaren Auslastungen lässt sich vermuten, dass der GA weiteres Verbesserungspotential hat und auch bessere Ergebnisse bringen könnte. Neben den reinen Modellparametern sollte auch hinsichtlich der Rekombinationsmechanismen weiter experimentiert und geforscht werden. Auch die Rechenzeit ist ein weiterhin zu berücksichtigender Faktor. In der durchgeführten Simulation wurde bei einer Berechnung mit einem Kern eine Rechenzeit von circa 55 Minuten für einen Durchlauf der vorteilhaftesten Parameterkombination benötigt.

Die zentrale Steuerung von Matrix-Montage-Systemen hat das Potential, höhere Auslastungen als dezentrale Steuerungsmethoden zu erreichen, da sie in der Lage ist, das Gesamtsystem zu betrachten. Ein GA ist geeignet, um gute Maschinenbelegungspläne zu erstellen. Der in dieser Arbeit vorgestellte Entwurf konnte zwar die Ergebnisse der dezentralen Steuerungsmechanismen nicht übertreffen, doch es konnten Richtungen für weitere Experimente aufgezeigt werden. Durch weitere Forschung kann es möglich sein, dass die Auslastungen des GA die der simpleren Steuerungsalgorithmen übertreffen und darüber hinaus von positiven Eigenschaften zentraler Steuerungen profitiert werden kann.

7 Literatur

- Aytug, H.; Khouja, M.; Vergata, F. E.: Use of genetic algorithms to solve production and operations management problems: a review, in: *International Journal of Production Research*, 41:17 (2003) 3955–4009.
- Buttelmann, Maik; Lohmann, Boris: Optimierung mit Genetischen Algorithmen und eine Anwendung zur Modellreduktion, in: *at – Automatisierungstechnik*, Ausgabe 52 (2004), S. 151–163.
- Echsler Minguillon, Fabio: *Prädiktiv-reaktives Scheduling zur Steigerung der Robustheit in der Matrix-Produktion*, Stuttgart 2020.
- Greschke, Peter: *Matrix-Produktion als Konzept einer taktunabhängigen Fließfertigung*, Vulkan Verlag, Essen 2016.
- Greschke, Peter; Herrmann, Christoph: Das Humanpotenzial einer taktunabhängigen Montage, in: *ZWF*, Ausgabe 109 (2014), S. 687–690.
- Koren, Yoram: *The Global Manufacturing Revolution*. Hoboken, NJ, USA: John Wiley & Sons, Inc 2010.
- Popp, Julian: *Neuartige Logistikkonzepte für eine flexible Automobilproduktion ohne Band*, Dissertation Universität Stuttgart, Stuttgart 2018.
- Schönemann, Malte; Herrmann, Christoph & Greschke, Peter & Thiede, Sebastian: Simulation of matrix-structured manufacturing systems, in: *Journal of Manufacturing Systems*, Ausgabe 37, (2015) S. 104–112.
- VDI 6224 Blatt 1: *Bionische Optimierung – Evolutionäre Algorithmen in der Anwendung*, in: *VDI-Fachbuch Bionik* 2012.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Modularer Simulationsbaukasten zur Modellierung und Optimierung der Instandhaltungsabläufe in Zugdepots

Modular simulation toolkit for maintenance processes in rail vehicle depots

Toni Donhauser, Siemens Mobility GmbH, Erlangen (Germany),
toni.donhauser@siemens.com

Niklas Natterer, Siemens Advanta Consulting, München (Germany),
niklas.natterer@siemens.com

Emin Genc, Siemens Mobility GmbH, München (Germany),
emin.genc@siemens.com

Abstract: The planning of plants for the maintenance of rail vehicles is subject to a wide range of influences with a high degree of mutual dependencies. If these remain unconsidered during the planning phase, the calculation is affected on the one hand by unnoticed erroneous assumptions that lead either to an overpriced offer or to unplanned costs when in operation. On the other hand, these planning errors cannot be corrected after the construction of the depot and result in inefficiencies or additional costs. For the validation of bid calculations, Siemens Mobility reproduces the planned layout and the intended workflows in the maintenance depot in a simulation model. By means of such a dynamic approach, bottlenecks and system deficiencies can be identified and avoided at the very beginning of the planning phase in order to submit a reliable offer with regard to viability and cost security. Due to the developed modular simulation toolkit, any depot can be modeled with very little effort to verify its suitability for the maintenance contract or to optimize existing operations.

1 Anforderungen an die Planung von Instandhaltungsabläufen für Schienenfahrzeuge

Trotz ihres Charakters eines Werkstattbetriebes weist die Instandhaltung von Schienenfahrzeugen aufgrund der wiederkehrenden Arbeitspläne Parallelen zu einer Serienproduktion auf. Aus diesem Grund bietet die Anwendung etablierter Ansätze zur Material- und Informationsflussoptimierung aus dem Produktionsumfeld (z.B. Lean, PPS) erhebliches Potenzial zur Steigerung der Effizienz. Um derartige Einsparungseffekte zu erzielen, hat die Siemens Mobility GmbH das „Maintenance Operational Excellence“-Programm aufgesetzt. In diesem Rahmen werden in

verschiedenen Workstreams die Erkenntnisse aus dem Shopfloor mit den verantwortlichen Bereichen für die Planung von Instandhaltungsdepots zusammengeführt. Hieraus entstehen Best Practices zur durchgängigen Ausübung einer effizienten Instandhaltung über alle Projekte hinweg, die entsprechend als Standards beschrieben werden.

Im Bereich des Workstreams „Produktionssystem“, der sich unter Einbezug aller Depotfunktionen für eine effiziente Aufbau- und Ablauforganisation und den zielgerichteten Einsatz digitaler Unterstützungstools verantwortlich zeigt, sind viele Fragestellungen hinsichtlich des Depotbetriebs isoliert nicht vollumfänglich bewertbar. Vielmehr ist es hierbei erforderlich, Zusammenhänge aus verschiedenen Bereichen (Disposition, Personal, Logistik u.a.) zu berücksichtigen und deren Wechselwirkungen zu quantifizieren.

Einen bewährten Ansatz zur Untersuchung derartiger Fragestellungen stellt die Ablaufsimulation dar, da diese im Zeitverlauf auftretende dynamische Wechselwirkungen zwischen einzelnen Systembestandteilen berücksichtigt. In einem Instandhaltungsdepot für Schienenfahrzeuge treten diese insbesondere durch das Rangieren der Fahrzeuge, durch unvorhersehbare Ereignisse, wie Infrastrukturdefekte oder Ablaufverzögerungen, oder durch Verspätungen bei der Bereitstellung von Fahrzeugen auf. Die dadurch bedingten Einflüsse auf die Arbeitsabfolge sowie Wartezeiten an den einzelnen Bearbeitungsstationen lassen sich in statischen Planungsverfahren nicht abbilden und führen in der Folge zu Ungenauigkeiten oder überflüssigen Puffern als Kapazitätsreserve.

2 Systembeschreibung und Problemstellung

Von der Planungsphase bis hin zum operativen Betrieb treten die geschilderten Herausforderungen in jedem Depot in unterschiedlichen Ausprägungen auf. Verteilt über das Gelände eines Instandhaltungsdepots erfolgen sowohl die klassischen Instandhaltungstätigkeiten wie planmäßige, korrektive oder werterhaltende Maßnahmen einschließlich Abdrehen der Radsätze, als auch weitere Services wie Außenreinigung, Innenreinigung, Enteisung oder Vorbereitungsdienste. Eine wesentliche Anforderung bei der Abarbeitung der entsprechenden fahrzeugindividuellen Arbeitspläne besteht darin, die Verfügbarkeit der Fahrzeuge für die täglichen Betriebsumläufe sicherzustellen. Dies kann beispielsweise in Nachtzeiten zu einem dichten Verkehrsaufkommen führen, das einen sehr hohen Dispositionsaufwand nach sich zieht. Dieser Umstand wird dadurch verstärkt, dass aufgrund des Gleislayouts ein zeitgleiches Rangieren mehrerer Fahrzeuge in bestimmten Bereichen oftmals nur eingeschränkt möglich ist. Aufgrund der Vielfalt an Einflussgrößen, wie Ankunftszeiten, Fahrzeuganzahl, Arbeitspläne oder Personalverfügbarkeit, ist ohne Simulation nur schwer zu ermitteln, ob die geplante Arbeitslast innerhalb des verfügbaren Zeitrahmens umzusetzen ist.

Gleichmaßen lässt sich ohne Berücksichtigung der Zeitdynamik der maßgebliche Engpass im Depotlayout oftmals nicht eindeutig bestimmen, genauso wie die erforderliche Personalkapazität an den einzelnen Arbeitsstationen. Dies erschwert beispielsweise die Bewertung, ob eine Änderung im Schichtmodell eine Entzerrung der Auslastung an einem Engpass herbeiführen kann.

3 Lösungsansatz: Simulationsbaukasten

Um die erforderliche Transparenz für die genannten Zielstellungen zu erreichen, ist für ein Instandhaltungsdepot ein Simulationsmodell zu erstellen und entsprechend der genannten Freiheitsgrade anzupassen, um in beliebigen Szenarien die Realität möglichst exakt nachzubilden. Anylogic bietet für die Simulation des Schienenverkehrs eine zugeschnittene Bausteinbibliothek (Eisenbahnbibliothek - Anylogic Simulationssoftware, 2021) an, die jedoch nicht auf die spezifischen Instandhaltungsprozesse ausgelegt ist, welche beispielsweise im Bereich der Komponentenüberholung einer klassischen Werkstattfertigung ähneln. Gleichmaßen verhält es sich bei den für den Eisenbahnbetrieb entwickelten Simulationsprogrammen OpenTrack (Nash & Huerlimann, 2004) und RailSys (Bendfeldt, Mohr, & Müller, 2000), die ihre Stärken in der Bewertung von Fahrplänen für ein vorgegebenes Netz haben, jedoch nicht für die Simulation von Instandhaltungsprozessen bzw. der Abläufe in einem Depot vorgesehen sind. Erste Ansätze zur Abbildung der Instandhaltungsabläufe in einer hierfür geeigneten ereignisdiskreten Simulation wurden von (Greasley, 2000; Hani, Amodeo, Yalaoui, & Chen, 2008) veröffentlicht, welche jedoch fallspezifische Individualmodelle beschreiben oder im Fall von (Bury, Spieckermann, Wortmann, & Hübler, 2018) durch die Betrachtung der Interaktion von Fahrzeugumläufen mit Instandhaltungsslots eine andere Abstraktionsebene ansprechen.

Daher wurde zur Minimierung des Erstellungs- und Anpassungsaufwands über diverse Projekte hinweg eine Bausteinbibliothek entwickelt, die einen Instandhaltungsbetrieb vom Geländelayout bis zur Abarbeitung von Arbeitsplänen integriert abbilden lässt. Diese enthält spezifische Standardbausteine für die Instandhaltung von Schienenfahrzeugen, die die folgenden Funktionen beinhalten:

- Ausführung spezifischer Arbeitspläne: planmäßige Inspektionen, Instandsetzungsmaßnahmen, Revisionen am Fahrzeug und in der Komponentenwerkstatt
- Instandhaltungs- und Serviceeinrichtungen/-anlagen: Außenreinigungsanlage, Werkstattgleise, Unterflurdrehbank, automatisierte Fahrzeuginspektionsanlage, Enteisungsanlage, Servicegleise, Abstellgleise
- Parameterabhängige Rangierlogik
- Generator für Instandhaltungspläne auf Basis von wiederkehrenden Zugaufenthalten: Ankunfts-/Bereitstellungszeiten, Arbeitsaufträge, Zugtyp, Betreiber u.a.
- Personelle Ressourcen: Schichtmodelle, Einsatzplanung, Kompetenzen
- Automatisierter Standard-KPI-Report zur Auswertung der Simulationsläufe

Durch den generischen Aufbau der Bausteine sowie eine zentrale Steuerungslogik lassen sich die einzelnen Module beliebig zusammensetzen und ohne weitere Modellierung in den Simulationsablauf integrieren. Insbesondere ermöglicht dies eine aufwandsarme Generierung neuer Szenarien, was vor allem bei der Bewertung von möglichen Anpassungen des Depotlayouts von Vorteil ist.

4 Modellierung und Implementierung

Der beschriebene Modulbaukasten wurde auf Basis eines Referenzmodells für die Abläufe in einem Zugdepot in Tecnomatix Plant Simulation implementiert. Grundsätzlich lassen sich Instandhaltungsdepots hiermit nach zwei unterschiedlichen Modellierungsansätzen abbilden. Liegt der Fokus weniger auf dem exakten Ablauf

der Zugsbewegungen hinsichtlich Kollisionsvermeidung und Wartepunkten, eignet sich die in Abbildung 1 dargestellte Modellierungsform, bei der anstelle exakter Distanzen und Geschwindigkeiten reine Zeitannahmen für Rangierfahrten angesetzt werden. Dies ermöglicht kürzere Laufzeiten und somit umfangreichere Untersuchungen. Der in Abbildung 2 dargestellte Ansatz basiert auf exakteren Dimensionen und fahrphysikalischen Eigenschaften und eignet sich somit für detailliertere Untersuchungen in Bezug auf die Fahrwege der Züge.

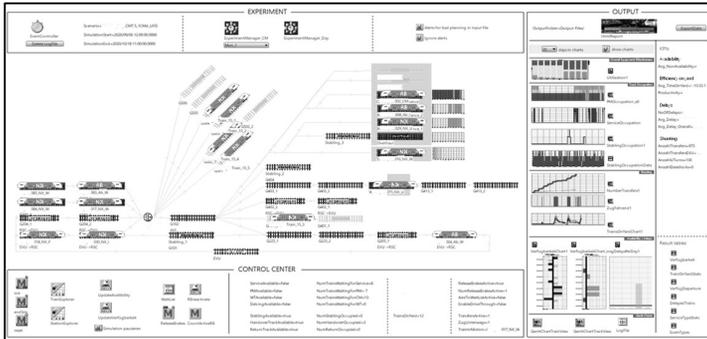


Abbildung 1: Ereignisdiskretes Simulationsmodell der Instandhaltungsabläufe in einem Zugdepot modelliert in Tecnomatix Plant Simulation

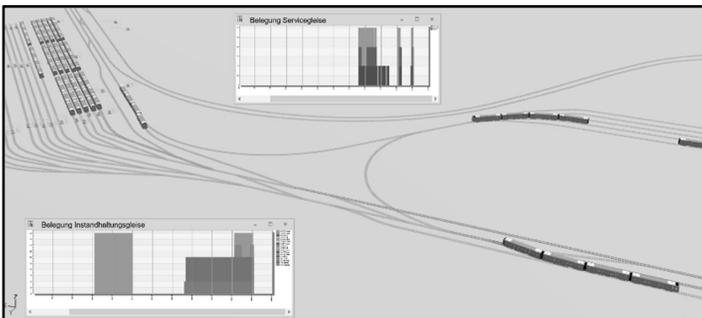


Abbildung 2: Ereignisdiskretes Simulationsmodell eines Zugdepots in 3D-Visualisierung (Tecnomatix Plant Simulation)

Modellbausteine

Für den Grundaufbau des Modells werden Standardbausteine aus Plant Simulation herangezogen. Sie werden in generischer Form auf den jeweiligen Einsatzzweck in der Zuginstandhaltung hin zugeschnitten, sodass diese Objekte für die entsprechende Verwendung in einem spezifischen Depot instanziiert werden können. So basiert eine Zugeinheit auf dem Fahrzeug-Baustein und ein Gleiselement auf einem Weg-Objekt. Dies hat zum Vorteil, dass die implementierte automatisierte Routenfindung direkt weiterverwendet werden kann. Die besagten Objekte können je nach gewählter Ausprägungsform (z. B. Abstell-, Übergabe-, Instandhaltungsgleis u. a.) bestimmte Eigenschaften annehmen und Methoden aufweisen.

An alle Gleiselemente, auf denen Arbeitsaufträge abgearbeitet werden, sind über Eingangsmethoden oder Sensoren Arbeitsstationen angebunden, die die Ausführung der dort vorzunehmenden Tätigkeiten steuern. Diese führen schließlich alle Daten und Informationen mit, die zur Bewertung der Effizienz des Gesamtsystems oder einzelner Bereiche heranzuziehen sind. Für diese Auswertungen sind für einen Instandhaltungsbetrieb spezifizierte Standarddiagramme angelegt, die je nach Typ der Arbeitsstation deren Einsatzgrad visualisieren.

Ablauflogik

Um die Fahrzeuge innerhalb des Depots an die erforderlichen Arbeitsstationen zur Erfüllung der zugeordneten Arbeitsaufträge zu disponieren, sind Termindaten (Ankunftszeit, späteste Ausfahrzeit u. a.), Arbeitspläne, Fahrzeugeigenschaften sowie Steuerungsregeln erforderlich. Abbildung 3 stellt die im Baukasten implementierte Steuerungslogik in einem Sequenzdiagramm dar.

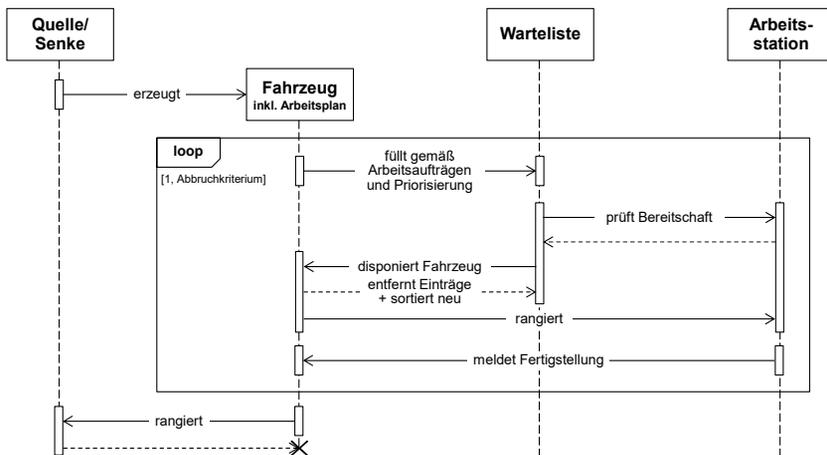


Abbildung 3: Steuerungslogik für Fahrzeuge zur Erfüllung der Arbeitsaufträge

Ausgehend von einem Terminplan für den zu simulierenden Zeitraum werden zu den festgelegten Zeitpunkten in einer Quelle Fahrzeuge generiert und ihnen der hinterlegte Arbeitsplan sowie spezifische Fahrzeugeigenschaften angehängt. Daraufhin wird das Fahrzeug zu einem der dedizierten Übergabegleise navigiert. Anhand der dortigen Eingangssteuerung wird der hinterlegte Arbeitsplan an die Warteliste kommuniziert und anhand von depotspezifisch festgelegten Priorisierungsregeln in die bereits bestehenden Fahraufträge eingeordnet. Mittels verschiedener Trigger wird die Warteliste kontinuierlich durchlaufen und hierbei geprüft, ob ein Fahrauftrag ausgeführt werden kann.

Ist für ein Fahrzeug nach verschiedenen Fahraufträgen und Aufenthalten an Arbeitsstationen ein Abbruchkriterium erfüllt (z. B. vorgegebene Abfahrtszeit erreicht, alle Arbeitsaufträge abgeschlossen), so wird es, bei Bedarf über Übergabegleise, einer Senke (z. B. Depotausfahrt) zugeführt. Während des gesamten Depotaufenthalts werden sämtliche Bewegungen und Aufenthalte aufgezeichnet und stehen im Nachgang zur Auswertung des jeweiligen Szenarios zur Verfügung.

5 Exemplarische Fallstudien

In der gewählten Form eignet sich der Baukasten für eine Vielzahl verschiedener Fragestellungen, die ohne Simulationseinsatz nur unzureichend bewertet werden können. Dies kann in verschiedenartigen Fallstudien nachgewiesen werden.

5.1 Bottleneck-Analyse

Bei der Planung neuer Depots sind Instandhaltungseinrichtungen und Gleisanlagen in optimaler Menge und Lage auf begrenzter Grundstücksfläche zu verorten, um effiziente Abläufe bei geringen Investitions- und Betriebskosten zu gewährleisten. Da eine nachträgliche Anpassung erheblichen Aufwand erfordert, wird bei unzureichender Transparenz bezüglich der zu erwartenden Abläufe mit großen Kapazitätspuffern gearbeitet. Um diese zu vermeiden, wird in einem neuen Instandhaltungsprojekt das geplante Depotlayout anhand der künftigen Umlaufpläne und Einbestellungen simulativ im Voraus bewertet und an den identifizierten Engstellen angepasst.

Dank der Simulation ist ersichtlich, aufgrund welcher Arbeitsstation Züge wie lange auf eine Weiterfahrt warten müssen bzw. in wie vielen Fällen diese auf ein Abstellgleis ausweichen müssen. Hierzu werden die Aufenthaltsorte der Züge sowie die entsprechenden Entscheidungen in der Steuerungslogik in Bezug auf die geplante Weiterfahrt geloggt und ausgewertet. Ferner können belastbare Aussagen in Bezug auf die erforderliche Anzahl an Abstellgleisen, an Triebfahrzeugführern sowie die erzielbare Außenreinigungsfrequenz für die Flotte getroffen werden.

5.2 Einfluss der Rangierstrategie auf die Fahrzeugverfügbarkeit

In einem bereits bestehenden Depot sind viele Freiheitsgrade aufgrund des starren Depotlayouts frühzeitig festgelegt. Hier sind vorrangig die operativen Entscheidungsvariablen zu optimieren, um die Effizienz im Depotbetrieb zu erhöhen. Hierzu kann der Einfluss der Disposition der Züge zu den einzelnen Arbeitsstationen innerhalb des Depots untersucht werden. Als Variablen für die Generierung auszuwertender Szenarien sind beispielsweise die Priorisierung einzelner Arbeitsstationen oder einzelner Zugtypen, die Nutzung von Abstellgleisen als Ausweichrouten oder unterschiedliche Schichtmodelle anzusetzen. Ziel solcher Simulationsstudien ist die Ermittlung einer robusten Strategie zur Erreichung einer maximalen Verfügbarkeit der Fahrzeugflotte für den Betrieb.

6 Ergebnisse der Fallstudien

Als Nachweis für die Anwendbarkeit und den Nutzen des Simulationsbaukastens für komplexe Fragestellungen bei der Planung der Aufbau- und Ablauforganisation sowie des Layouts eines Instandhaltungsdepots, werden nachfolgend umgesetzte Fallstudien vorgestellt. Als Vergleichsreferenz dienen klassische Planungsansätze ohne Einsatz von Simulationsverfahren.

Kostenminimale Erweiterung zur Erreichung des Zieldurchsatzes

Bei neuen Instandhaltungsprojekten gelingt basierend auf Erfahrungswerten eine gute Annäherung an erforderliche Ressourcen zur vollständigen Erledigung des

Arbeitspensums. Sind hierbei jedoch wechselseitig einschränkende Einflussfaktoren zu berücksichtigen (z. B. Depotlayout, Schicht- und Umlaufpläne, variierende Prozesszeiten), so sind für die endgültige Festlegung Sicherheiten vorzusehen. In einer Fallstudie waren in einem zweiwöchigen Referenzzeitraum in der Nacht 182 Fahrzeuge in einem Depot zu behandeln. Der Untersuchungsfokus lag auf einer rechtzeitigen Betriebsfreigabe aller Fahrzeuge bei möglichst geringen Wartezeiten durch kostenminimale Erweiterung der Arbeitsstände innerhalb der Geländegrenzen.

Ausgehend von einem Referenzszenario wurden zusätzliche Abstell- und Innenreinigungsgleise als günstige Erweiterungen auf deren Einfluss auf den Durchsatz untersucht. Diese wurden anhand von Engpassanalysen als maßgebliche Verursacher von Verzögerungen identifiziert. Tabelle 1 zeigt, dass durch ein zusätzliches Abstellgleis weiterhin viele Fahrzeuge die vorgesehene Abfahrtszeit nicht einhalten können. Ein zusätzliches Gleis für die Innenreinigung hingegen führt zu einem frühen Zeitpunkt zu einer Reduzierung der Aufstauung und so zu einer zügigen Weiterbearbeitung. Dies lässt in der Folge lediglich sechs der 182 Fahrzeuge übrig, deren pünktliche Ausfahrt nicht ohne Weiteres sichergestellt werden kann. In einer Einzelfallbetrachtung konnten jedoch organisatorische Maßnahmen identifiziert werden, um auch diese Fahrzeuge rechtzeitig freizugeben.

Tabelle 1: Szenarienvergleich zur Identifizierung erforderlicher Erweiterungen

	Referenzszenario	Ref.-sz. + Abstellgleis	Referenzszenario + Innenreinigungsgleis
Freigabe kritisch [# Fzg.]	35	23	6
Ø Wartezeit im Depot	86 min	71 min	37 min

Nachweis der zu erreichenden Übergabetaktung in den Umlauf

In vielen Lightrail-Projekten erfolgen sowohl die Abstellung als auch die Reinigung und Instandhaltung auf demselben Depotgelände. Dies bringt die Anforderung mit sich, einen Großteil der Fahrzeuge zu Beginn der Hauptverkehrszeiten sehr kurz getaktet an den Betrieb zu übergeben. Aufgrund großer Distanzen sowie Restriktionen wie Maximalgeschwindigkeiten oder abschnittswise Einzelzugbewegungen kann dies eine große Herausforderung darstellen. In einem entsprechenden Projekt sind zur Bedienung des morgendlichen Passagieraufkommens 18 Fahrzeuge im 90-Sekunden-Takt in den laufenden Betrieb einzugliedern. Aufgrund der dort auftretenden Interaktion des Depotverkehrs mit dem Fahrzeugumlauf ist die Eignung der geplanten Streckenführung und Signalgebung im Depot simulativ zu bewerten.

Ausgehend von einem Referenzszenario mit zehn verspäteten Übergaben wurden verschiedene Strategien entwickelt, um die Übergabeanforderungen vollständig zu erfüllen. Durch eine veränderte Routenführung, ein verlängertes Ausziehgleis sowie die vorzeitige Bereitstellung einzelner Fahrzeuge an den Endhaltestellen ist dies abschließend gelungen.

Analyse des Einflusses von Schichtmodellen

Der erhebliche organisatorische Aufwand zur Änderung eines Schichtmodells sollte vorab hinsichtlich der erwarteten Effizienzerhöhung abgesichert werden. In der

vorliegenden Fallstudie war zu analysieren, welche Auswirkungen die Vermeidung von Nacharbeiten in der Instandhaltung auf die Verfügbarkeit der Fahrzeuge für den Betrieb hat. Durch die Entzerrung des nächtlichen Rangierbetriebs sind geringere Wartezeiten und somit eine rechtzeitige Betriebsfreigabe für die zur Außenreinigung einbestellten Fahrzeuge zu erwarten.

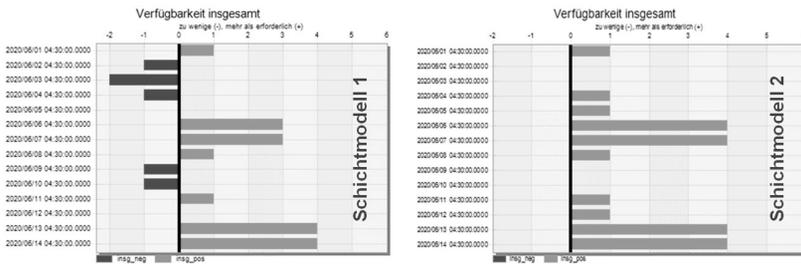


Abbildung 4: Fahrzeugverfügbarkeit in Abhängigkeit des Schichtmodells

Abbildung 4 zeigt die Einhaltung der geforderten täglichen Fahrzeugverfügbarkeit über zwei Musterwochen in Abhängigkeit des gewählten Schichtmodells. Ein nächtlicher Instandhaltungsbetrieb (Schichtmodell 1) führt zu erhöhtem Verkehrsaufkommen und somit zu Verzögerungen in der Abwicklung der kurz getakteten Rangierfahrten zwischen den Reinigungsanlagen. Ändert man lediglich das Schichtmodell ohne Erhöhung des Personaleinsatzes, so treten keine Nicht-Verfügbarkeiten mehr auf und es verbleiben überdies Puffer für die Kompensation von unerwarteten Ereignissen oder zur Optimierung des Ressourceneinsatzes.

Literaturverzeichnis

- Bendfeldt, J.-P., Mohr, U., & Müller, L. (2000). RailSys, a system to plan future railway needs. *WIT Transactions on The Built Environment*(50).
- Bury, H., Spieckermann, S., Wortmann, D., & Hübler, F. (2018). A case study on simulation of railway fleet maintenance. *Winter Simulation Conference (WSC)* (S. 2851-2860). IEEE.
- Eisenbahnbibliothek - Anylogic Simulationssoftware. (19. Juli 2021). (The AnyLogic Company) Von <https://www.anylogic.de/resources/libraries/rail-library/> abgerufen
- Greasley, A. (2000). Using simulation to assess the service reliability of a train maintenance depot. *Quality and Reliability Engineering International*(3), S. 221-228.
- Hani, Y., Amodeo, L., Yalaoui, F., & Chen, H. (2008). Simulation based optimization of a train maintenance facility. *Journal of intelligent manufacturing*(3), S. 293-300.
- Nash, A., & Huerlimann, D. (2004). Railroad simulation using OpenTrack. *WIT Transactions on The Built Environment*(74).



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Collaborative Virtual Reality System for Industrial based Assembly Training

Kollaboratives Virtual-Reality-System für industrielles Montagetraining

Rainer Müller, Leenhard Hörauf, Max Eichenwald, ZeMA gGmbH, Saarbrücken (Germany), rainer.mueller@zema.de, leenhard.hoerauf@zema.de, m.eichenwald@zema.de

Abstract: The handling of increasing product complexities and a raising number of variants demands a high degree of expertise and adaptability from employees in the assembly area. Influences such as the advancing demographic change or employee absences due to illness require the short-term assumption of assembly tasks by other employees. In order to avoid production downtimes, effective and quickly adaptable training systems must be available. Virtual Reality (VR) based simulations allow the training on a digital twin in a virtual environment. This paper describes a developed system that allows an interactive, manual assembly process. The focus is on the simulation of interdepartmental and cross-location interaction and collaboration. Additionally, a user study (N=12) was conducted to identify potential benefits of VR-based collaborative training simulations over a comparable training scenario via an online communication platform, showing that the use of the developed system results in lower assembly time and number of errors.

1 Introduction

Factors such as increased product complexity, shorter time-to-market, older workforces due to advancing demographic change and the short-term assumption of tasks by specialised employees in the event of illness or retirement from the company require the efficient training of employees to meet these challenges. The further development of modern tools for mapping training simulations is therefore becoming increasingly important (Thomas et al. 2018). One possibility that is favoured by strong technological developments is the mapping of training scenarios using Virtual Reality (VR). Simulating real environments in a virtual simulation environment, based on the three-dimensional representation of relevant products can be a useful tool for employee training. Ultimately, the goal of such simulation environments is always to establish the transfer between virtual and real environments. Workflows and previously simulated scenarios should be recognised and intuitively adopted by the user in the real environment (Learning by doing (Klemmer et al. 2006)).

In the following, the state of the art of current developments in VR-based training will be discussed. Section 3 presents a developed multi-user VR system. In addition, a user study was conducted, which aims at comparing collaborative VR training to "classical" online training. Section 4 gives an overview of the goal, procedure and results of this study. Discussion and conclusion is highlighted in Section 5.

2 State of the Art

The first VR-based training system was introduced in 1999 by Jayasekera et al. (Jayaram et al., 1999) under the name Virtual Assembly Design Environment (VADE). Interaction possibilities between user and objects were given, as well as the possibility to assemble or disassemble a product. Based on this system a large number of VR based training systems for validation, visualisation and simulation of assembly processes can be identified today (Abidi et al. 2019; Gomes de Sá and Zachmann 1999; Jayasekera and Xu 2019; Im et al. 2017). Furthermore, Al-Ahmari et. al (Al-Ahmari et al. 2016) have shown that the implementation of tactile and auditory feedback in addition to the purely visual form of presentation can increase the performance and learning rate of the user.

An increasing trend is currently seen in network-based collaboration within virtual reality environments. Bringing together multiple users within a VR environment and the collaborative handling of objects enables new possibilities in a variety of different industries. Developments and applications of collaborative VR systems can be clustered in: medical, civil engineering and industry.

Medicine: Collaborative VR simulators are already used in medical fields. Collaborative and immersive environments provide reality based scenarios for education and assessment of surgical teams. With the goal to improve medical skills, especially in rare or complicated situations multiple users can be guided by a trainer within the scenario (Chheang et al. 2019; Paiva et al. 2015; Schild et al. 2018).

Civil Engineering: With the goal of efficient and simplified collaboration between several stakeholders of a joint project, VR based environments of buildings are used. The combination with BIM (Building Information Modelling) can provide additional information regarding necessary specifications and implementations. Du et al. (Du et al. 2016) describe a system in which a common walkthrough of all stakeholders is possible.

Industry: The development of commercial simulation tools in the industrial environment is also progressing steadily. As an example, the company *Raumtänzer* (Raumtänzer GmbH 2020) offers training systems that enable a gathering of several employees in a virtual environment.

In addition, especially for industrial applications (i.e. assembly and disassembly processes), systems have to be created that enable realistic interaction between users and objects. Automated assistance, but also location-independent communication and collaboration options between several users are required. The aim of the system presented below is to simplify direct cooperation and collaboration between employees. On the one hand, a visual interaction between employees should be possible, independent of the available end devices of the users. On the other hand,

the point of direct collaboration is examined in more detail on the basis of a shared digital twin. The implementation of feedback in the form of visualisation and speech should facilitate work within the virtual environment.

3 System Development

This chapter describes the developments of a VR-based system for the training of employees. Therefore, the setup and basic functionalities of the implemented user-communication and -collaboration will be described.

3.1 Setup

The setup for the virtual environment includes Head-Mounted Displays (HMD) (here: Oculus Quest). The integrated hand tracking of the HMD is used for the single-user application described in Section 3.2. On one hand, this enables intuitive handling of the application, and on the other hand, it allows to assign specific functionalities in different gestures. The multi-user application described in Section 3.3 requires at least two HMD's and the additional use of the controllers, which facilitate user interaction and movement. The system is implemented in the Unity development environment. The developments listed below build on previous implementations (Müller et al. 2020). The training environment consists of an assembly system and the product to be assembled. To give the user a realistic impression of the working environment, the workplace was digitized with the help of a 3D laser scanner. The generated point cloud was converted into an *.obj File (including necessary colour information) and embedded in Unity. For the product to be assembled (here: transmission), an existing 3D-CAD model could be used.

3.2 Network-based communication Tool for single-user Training

The functionality is divided into two modes:

- **Guided assembly:** The part/subassembly to be assembled in the next step is highlighted using a shader. In addition, a panel points to the object and provides further information regarding object name, process- and resource information.
- **Unguided assembly:** The employee does not receive any assistance and has to implement the previously learned. For subsequent validation, the throughput time and the assembly sequence is recorded.

In order to give the employee additional support, the possibility to start a video chat is given. The implementation of the chat function is based on the Agora SDK for Unity. When the video chat is started, users of the VR as well as of a computer or smartphone can join the chat room. The videos of the camera/webcam of the computer and smartphone are streamed as well as a live recording from the virtual environment. A visual and auditive communication is thus given. This functionality enables the behaviour and movements of an employee within the VR to be displayed by other devices (computer/smartphone) independent of location. Supervisors can follow the training process of a new/inexperienced employee and give him/her assistance and instructions. Figure 1 gives an overview of the communication and

visualisations in the individual systems. In Figure 1 (a) the view of the user of the virtual environment becomes clear. By selecting the *Videochat*-function the user can enter the chat room. Other users are displayed next to the main panel. A similar application is provided for desktop/smartphone users. Figure 1 (b) shows the application from the PC user's point of view. The own camera image (bottom left) as well as a smartphone user (top left) and VR user (top right) can be displayed.

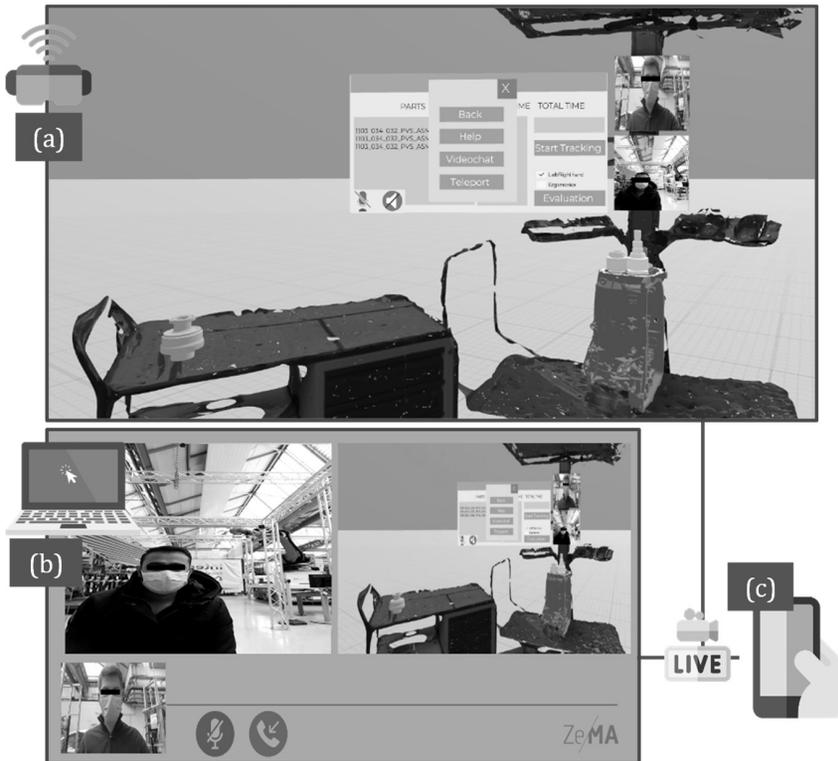


Figure 1: Live video chat functionality. Connection and visualization of multiple participants in (a) VR system, (b) desktop application, (c) smartphone

3.3 Multi-User Collaboration Platform

In addition to the single-user training, a training for multi-user applications is implemented to create a virtual, collaborative environment. The goal is to enable location-independent collaboration of several employees to create an effective and efficient training character. Different functional scopes and visualisations are offered depending on a previously selected role (employee on shop floor level (worker), engineer, assembly planner, etc.). For a realistic visualisation, the focus is put on the visualisation and manipulation of objects in real time, and on the communication in the form of a voice chat. In a first step a suitable Unity extension was selected, which enables network-based collaboration. The following systems are examined in detail: Unet, Bolt, Photon Unity Network (PUN), Photon Realtime and TriCat. With

regard to necessary specifications, the target system should i.e. be cloudbased, have a high number of possible users and Unity support. By performing a pairwise comparison and a utility analysis (VDI-Richtlinie Blatt 3), PUN was identified as a suitable package.

Interaction

The user/trainee should be able to interact with other users/teacher. Two types of interaction were implemented: **Voice communication** through which users can talk to each other inside the virtual environment. Provided is an open, distance and position sensitive communication. A realistic communication in which the direction of the sound depends on the position of the users to each other. Conversations at long distances appear quieter than those at short distances from each other. Furthermore, the direction from which the sound is heard depends on the orientation of the users to each other. **Tactile communication** through which users can highlight (here: via colored shader effect) grabbed or targeted parts. Instructions and assembly sequences are thus easier to see. Figure 2 illustrates the implementation of the collaborative development environment. It is shown how the interaction between a teacher (red) and a trainee (white) can take place. In point (d) of the Figure the previously presented *tactile communication* is illustrated.

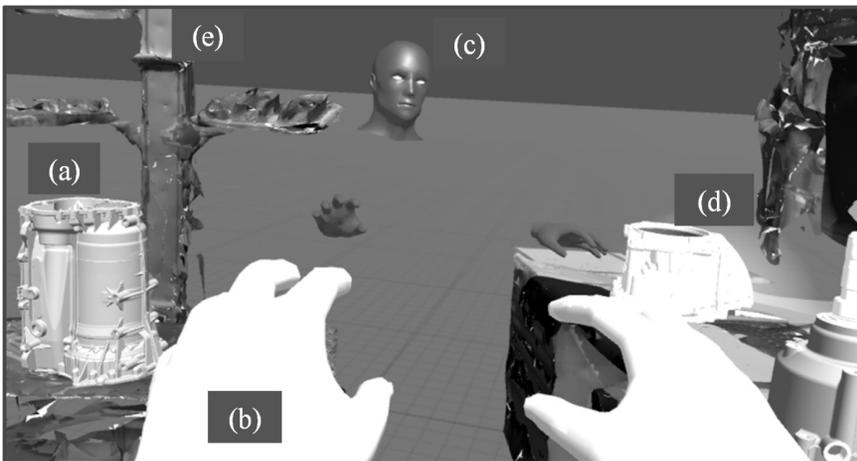


Figure 2: Multi-user training scenario, (a) product, (b) learning employee, (c) teacher, (d) visualizing next part by teacher (shader effect), (e) scanned real environment

4 User Study and Assessment

In the second part of the work the developed multi-user system (see Sec. 3.3) is evaluated within a user study. The focus is on the collaboration between employees within a common environment. The goal is to make a statement regarding the effectiveness of such virtual collaboration in an industrial environment.

4.1 Experiment Task

An eight-speed automatic transmission is used to perform the user study (see Fig. 3 (c), (d)). A total of six parts/subassemblies had to be mounted by the participants. The assembly is carried out by joining. Operating resources such as screwdrivers are not necessary. In order to simulate the scenario for the participants within the VR simulation in the best possible way, the transmission and workstation were imported into the simulation close to reality. The parts/subassemblies to be mounted are placed at the same position in the simulation as well as in the real environment (see Fig. 3 (a), (b)).

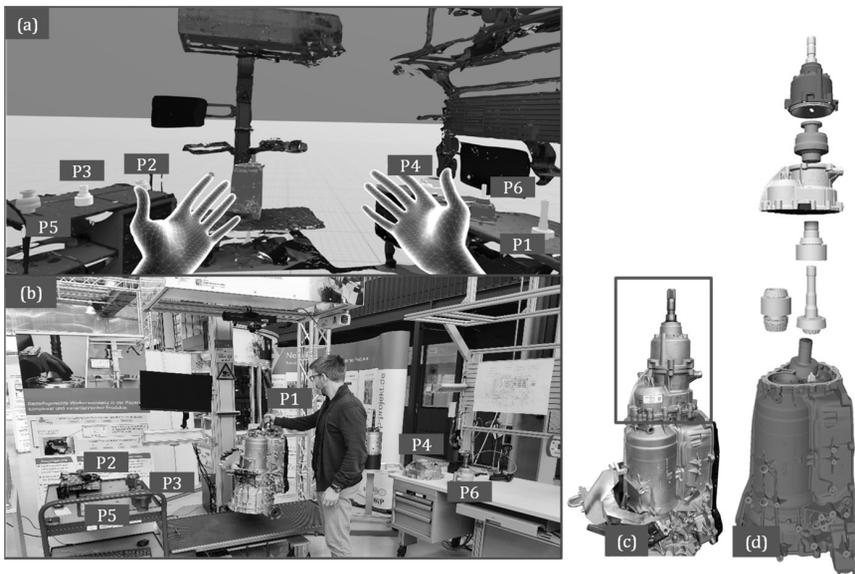


Figure 3: Representation of a (a) virtual workstation modelled on the (b) real environment; P1-P6 illustrate the components to be assembled / Visualisation of (c) real and (d) virtual (CAD) transmission

4.2 Procedure

To conduct the study, the participants are randomly divided into two groups. Six people per group, consisting of a mixture of men and women with technical experienced and non-technical experienced. The procedure is as follows (see Fig. 4): Both groups are provided with a declaration of consent before the start of the user study. Afterwards the participants are guided through a training session. Participants of group 1 (VR-System) receive training with the developed VR system under the guidance of a trainer. Participants in group 2 (Baseline) also received digital training. This took place using the communication tool Microsoft Teams. On the one hand, they are shown a presentation that included pictorial instructions for assembly, and on the other hand, the expert is also available for questions and explanations. After the training for both groups is finished, the real assembly took place. The time required and the number of errors made by the participants is measured.

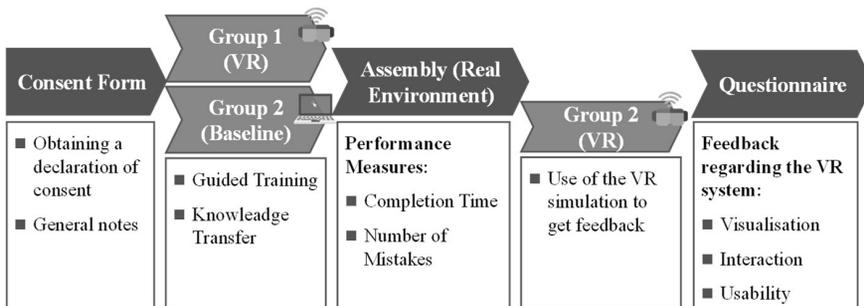


Figure 4: Schematic process of user study

4.3 Analysis and Results

In order to make a statement regarding the advantage of collaborative VR simulations over training processes with the help of a communication platform, two hypothesis are defined regarding assembly time and number of errors:

1. Time:

Null Hypothesis H_0 : The average assembly time achieved by both groups is equal

Alternative Hypothesis H_1 : The average assembly time achieved by participants from group 1 is smaller than the average assembly time achieved by participants from group 2

2. Errors:

Null Hypothesis H_0 : The average number of errors (two types of errors (EC) were considered here) made during the assembly process are equal for both groups

Alternative Hypothesis H_1 : The average number of errors made during the assembly process by participants from group 1 is smaller than the average number of errors by participants from group 2

Figure 5 gives an overview of the data collected during the conducted user study. Participants 1, 2, 3, 8, 10 and 12 are assigned to the group VR. Participants 4, 5, 6, 7, 9, 11 are assigned to the group Baseline. The upper part of the Figure shows a total overview of the time required, broken down into individual process times, per participant for the assembly of the transmission. The average value for the VR group is 87,17 s, the average value for the Baseline group is 121,83 s. The lower part of the Figure describes the errors made during assembly. A distinction was made between error EC1 (Wrong object gripped) and error EC2 (Mounted in wrong pose). The average values in relation to the participants of the individual groups are also given.

The calculation of the p-value (ANOVA analysis) was used to check the hypotheses previously established. The p-value describes whether the difference between the calculated mean values of several samples (VR group compared to baseline group) is statistically significant. The comparison of the p-value with a previously defined significance level (here a standard value of $\alpha = 0,05$ (Fisher 1992) is assumed, thus the calculated p-value needs to be smaller than α to proof statistical significance) provides an indication of whether the null hypothesis is true or should be rejected.

Table 1 gives an overview of the p-values determined and thus of the statistical significance.

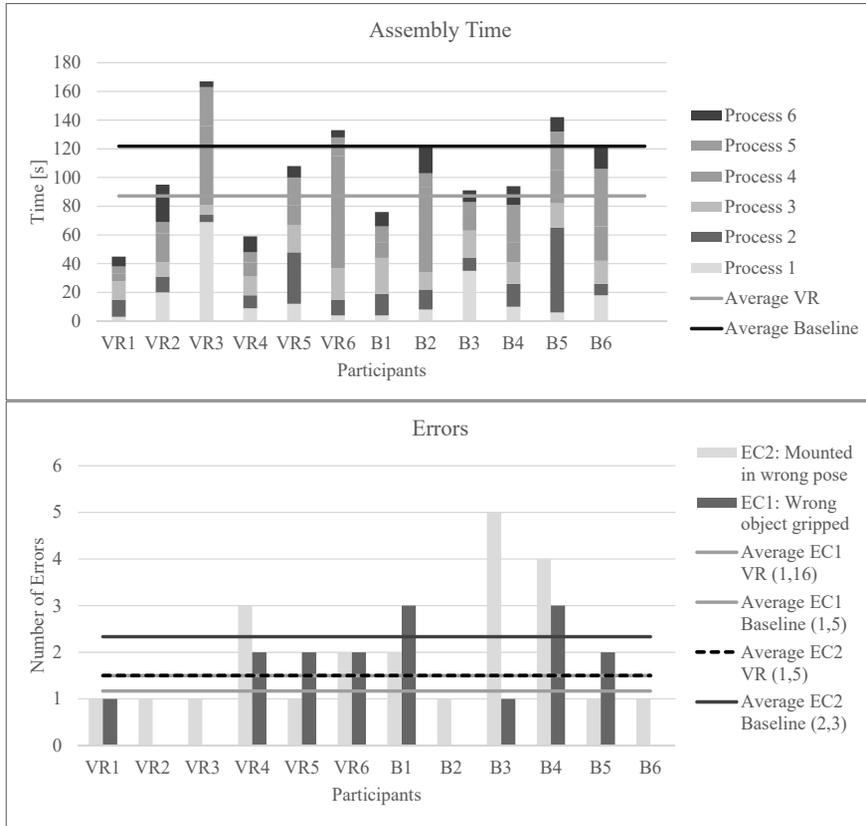


Figure 5: Measured time per process per participant and number of errors for VR Training group and Baseline Training group (VR group participants: VR1-VR6/ Baseline group participants: B1-B6)

Table 1: Statistical significance

Null Hypothesis	p-value	Significance	Null Hypothesis
Time	0,001<p>0,005	High Significance	Reject
Error – EC1	0,5<p<0,75	Low Significance	Not Reject
Error – EC2	0,5<p<0,75	Low Significance	Not Reject

5 Discussion and Conclusion

This paper presents a virtual reality based system for the simulation of manual assembly activities in an industrial environment, enabling a potential user to assemble a defined product (here a transmission). The focus was on cooperation and collaboration between employees. Furthermore, a user study was carried out to find out to what extent the collaborative VR-based system has advantages compared to a remote training via Microsoft Teams. Thereby 12 participants were randomly assigned to two groups. Group 1 received instructions within the virtual environment. Group 2 was instructed via the communication platform Microsoft Teams. With the help of an ANOVA analysis (calculation of the p value) the following conclusion could be drawn:

Group 1 participants who received virtual reality based training were able to complete the assembly of the real product in a shorter time than group 2. The average time to complete assembly operations is 87,17 s for group 1 and 121,83 s for group 2. Group 1 participant who received virtual reality based training, were able to complete the assembly of the real product with fewer errors than group 2. The wrong part/subassembly was gripped less seldom and an attempt to assemble an actually correct object in the wrong pose was made less often. Furthermore it could be shown, that VR-based training has a statistical significance in terms of assembly time compared to the baseline group, meaning that VR based training leads to a lower assembly time of the real product. Statistical significance in terms of the number of errors made could not be demonstrated. However, in order to obtain more accurate results, the number of participants should be expanded for further investigation. It could be shown that the implementation of a collaborative training form using VR technologies can bring advantages for the training or further education of employees regarding assembly processes in an industrial environment.

Acknowledgement

This paper was written in the framework of the research project KomZetSaar (Funding code 01MF117003A), which is funded by the Federal Ministry of Economic Affairs and Energy (BMWi) and supervised by the lead partner German Aerospace Center (DLR).

References

- Abidi, M.H.; Al-Ahmari, A.; Ahmad, A.; Ameen, W.; Alkhalefah, H.: Assessment of virtual reality-based manufacturing assembly training system. *The International Journal of Advanced Manufacturing Technology* 105 (2019) 9, pp. 3743–3759.
- Al-Ahmari, A.M.; Abidi, M.H.; Ahmad, A.; Darmoul, S.: Development of a virtual manufacturing assembly simulation system. *Advances in Mechanical Engineering* 8 (2016) 3, 168781401663982.
- Chheang, V.; Saalfeld, P.; Huber, T.; Huettl, F.; Kneist, W.; Preim, B.; Hansen, C.: *Collaborative Virtual Reality for Laparoscopic Liver Surgery Training* (2019).
- Du, J.; Shi, Y.; Mei, C.; Quarles, J.; Yan, W.: *Communication by Interaction: A Multiplayer VR Environment for Building Walkthroughs* (2016), pp. 2281–2290.
- Fisher, R.A.: *Breakthroughs in statistics* (1992), pp. 66–70.

- Gomes de Sá, A.; Zachmann, G.: Virtual reality as a tool for verification of assembly and maintenance processes. *Computers & Graphics* 23 (1999) 3, pp. 389–403.
- Im, T.; An, D.; Kwon, O.-Y.; Kim, S.-Y.: A Virtual Reality based Engine Training System - A Prototype Development & Evaluation (2017), pp. 262–267.
- Jayaram, S.; Jayaram, U.; Wang, Y.; Tirumali, H.; Lyons, K.; Hart, P.: VADE: a Virtual Assembly Design Environment. *IEEE Computer Graphics and Applications* 19 (1999) 6, pp. 44–50.
- Jayasekera, R.D.; Xu, X.: Assembly validation in virtual reality—a demonstrative case. *The International Journal of Advanced Manufacturing Technology* 105 (2019) 9, pp. 3579–3592.
- Klemmer, S.R.; Hartmann, B.; Takayama, L.: How bodies matter. In: Carroll, J.M.; Bødker, S.; Coughlin, J. (Hrsg.): *Proceedings of the 6th ACM conference on Designing Interactive systems - DIS '06*, University Park, PA, USA, 26.06.2006 - 28.06.2006, 2006, pp. 140.
- Müller, R.; Hörauf, L.; Bashir, A.; Karkowski, M.; Eichenwald, M.: Virtual Reality based Assembly Process Validation and Rework Assistance with consistent Data Exchange. In: Schüppstuhl, T.; Tracht, K.; Henrich, D. (Hrsg.): *Annals of Scientific Society for Assembly, Handling and Industrial Robotics*. Berlin, Heidelberg: Springer Berlin Heidelberg 2020, pp. 1–10.
- Paiva, P.V.; Machado, L.d.; Batista, T.V.: A Collaborative and Immersive VR Simulator for Education and Assessment of Surgical Teams (2015), pp. 176–185.
- Raumtänzer GmbH, 2020: Flux VR: Virtuelle Mitarbeiterschulungen. Online verfügbar unter <https://www.raumtaenzer.com/flux-vr-virtual-reality-software-mitarbeiterschulung/>, zuletzt geprüft am 23.10.2020.
- Schild, J.; Lerner, D.; Misztal, S.; Luiz, T.: EPICSAVE — Enhancing vocational training for paramedics with multi-user virtual reality (2018), pp. 1–8.
- Thomas, O.; Metzger, D.; Niegemann, H. (Hrsg.): *Digitalisierung in der Aus- und Weiterbildung: Virtual und Augmented Reality für Industrie 4.0*. Berlin: Springer Gabler 2018.
- Verein Deutscher Ingenieure VDI-Richtlinie Blatt 3: VDI 2225 - Design engineering methodics: VDI Verlag Düsseldorf, 1998.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Simulationsbasierte Konfiguration der Stationen von Matrix-Produktionssystemen

Simulation-based Configuration of Working Stations in Matrix-Structured Manufacturing Systems

Sven Völker, Richard Verbeet, Technische Hochschule Ulm, Ulm (Germany),
sven.voelker@thu.de, richardverbeet@yahoo.de

Abstract: When designing matrix-structured manufacturing systems, it must be determined which station should be equipped with which capabilities. This paper shows how this planning task can be solved by means of simulation-based optimization. Simulation experiments show the effectiveness of the method and the influence of the assignment of capabilities to stations on the performance of these systems. An analytical model for configuring matrix-structured manufacturing systems is proposed and used for comparison.

1 Einleitung

Für die Montage von variantenreichen Produkten in hohen Stückzahlen sind getaktete Variantenfließlinien üblich. Dieses Konzept stößt bei wachsender Produktvarianz und sich im laufenden Betrieb veränderndem Produktmix an seine Grenzen. Seit einigen Jahren wird das Konzept der Matrix-Produktion (Matrix-Structured Manufacturing Systems, MSMS) als Alternative diskutiert (Greschke 2016). MSMS sind flexible Fertigungssysteme (ElMaraghy 2005), bei denen die Arbeitsstationen in Form einer Matrix angeordnet sind und die Werkstücke mittels fahrerloser Transportfahrzeuge zwischen den Arbeitsstationen transportiert werden (Abb. 1). Die Arbeitsstationen in MSMS sind nicht auf die Ausführung bestimmter Arbeitsgänge festgelegt, sondern können ohne Rüstaufwand flexibel eingesetzt werden (ElMaraghy 2005).

Die Gestaltung von MSMS erfordert neue Entwurfsmethoden (Filz et al. 2019, Kern et al. 2016). Von besonderem Interesse ist dabei die Frage, welche Stationen zur Ausführung welcher Arbeitsgänge ausgerüstet werden sollen: Es ist weder erforderlich noch praktikabel, sämtliche Stationen zur Ausführung aller Arbeitsgänge zu befähigen, da die Ausstattung Flächen verbraucht und Kosten verursacht.

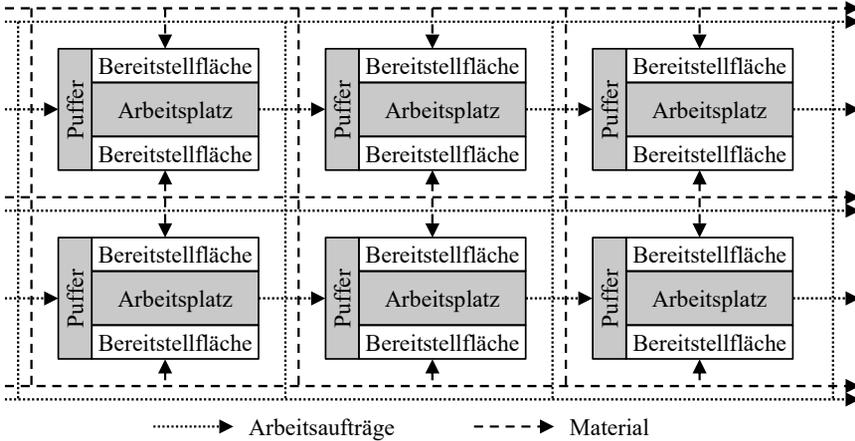


Abbildung 1: Struktur eines MSMS (vgl. Filz et al. 2019)

Bisherige Arbeiten zu MSMS konzentrieren sich häufig auf die operative Steuerung: Müller und Schmitt (2020) betrachten den Einfluss der Auftragsfreigabe, Ranke und Bauernhansl (2021) untersuchen verschiedene Bestellverfahren und Bányai et al. (2019) schlagen eine Heuristik für die Logistikplanung vor. Mehrere Arbeiten behandeln maschinelles Lernen im Kontext der Steuerung von MSMS (Göppert et al. 2020, Hofmann et al. 2020, May et al. 2021). Regelmäßig wird die Methode der Materialflusssimulation angewandt, um die Dynamik von MSMS zu analysieren (Schönemann et al. 2015, Filz et al. 2020).

Im Gegensatz zu Fragen der operativen Steuerung wurde das Problem der Verteilung von Fähigkeiten auf Arbeitsstationen bisher kaum untersucht. Burggräf et al. (2020) schlagen dafür ein kapazitätsbasiertes Verfahren vor, das sich durch eine detaillierte Modellierung von Fähigkeiten und ihre Zusammenfassung zu Fähigkeitsprofilen auszeichnet. Die Fähigkeitsprofile werden den Stationen manuell zugeordnet.

Der vorliegende Beitrag untersucht die Auswirkung der statischen Zuordnung von Fähigkeiten zu Arbeitsstationen auf die Leistungsfähigkeit eines MSMS und stellt ein Zuordnungsverfahren auf der Basis simulationsbasierter Optimierung vor.

2 Problembeschreibung

Grundlage der Auslegung eines MSMS sind Annahmen über die in einer Periode zu fertigenden Mengen von Produkten und Produktvarianten sowie über die zugehörigen Arbeitspläne. Die Arbeitspläne legen die Dauer jedes Arbeitsgangs und die zu seiner Ausführung benötigte Fähigkeit fest. Unter „Fähigkeit“ wird das Vorhandensein aller für das Ausführen eines Arbeitsgangs erforderlichen Werkzeuge, Betriebsmittel und Mitarbeiterqualifikationen verstanden. Es wird angenommen, dass jeder Arbeitsgang genau eine Fähigkeit erfordert. Mehrere Arbeitsgänge können dieselbe Fähigkeit erfordern, aber unterschiedlich lange dauern.

Als Vorbereitung zur Lösung des Planungsproblems werden in einer statischen Analyse der gegebenen Daten eine Unterschranke für die Gesamtzahl benötigter Arbeitsstationen berechnet und für jede Fähigkeit ermittelt, wie viele Stationen

mindestens mit dieser Fähigkeit ausgestattet werden müssen. Auf der Grundlage dieser unteren Schranken wird festgelegt, aus wie vielen Stationen das MSMS bestehen soll und wie viele Stationen jeweils mit den einzelnen Fähigkeiten ausgestattet werden sollen. In diese Festlegung fließen Erwägungen zur angestrebten Auslastung und zur funktionalen Redundanz der Stationen ein.

Nun ist ein ganzzahliges Optimierungsproblem zu lösen: Gegeben sind eine Menge von Arbeitsstationen M sowie eine Menge von Fähigkeiten S . Die binären Entscheidungsvariablen $x_{m,s}$ geben an, ob eine Station $m \in M$ über die Fähigkeit $s \in S$ verfügt ($x_{m,s} = 1$) oder nicht ($x_{m,s} = 0$). Dabei gelten folgende Nebenbedingungen:

$$\sum_{s \in S} x_{m,s} \leq s_{\max} \quad \forall m \in M \quad (1)$$

$$\sum_{m \in M} x_{m,s} = n_s \quad \forall s \in S \quad (2)$$

Die erste Nebenbedingung (Gl. 1) begrenzt die Anzahl der Fähigkeiten, mit denen eine Station ausgestattet werden kann, auf einen Maximalwert s_{\max} . Die zweite Nebenbedingung (Gl. 2) stellt sicher, dass die Anzahl der mit der Fähigkeit s ausgestatteten Stationen der Vorgabe n_s entspricht. Der Wert von n_s ist dabei proportional zur Nachfrage nach dieser Fähigkeit, die aus dem Produktionsprogramm resultiert. Dadurch soll erreicht werden, dass keine Fähigkeit zum Engpass für das gesamte System wird.

Eine optimale Auslegung des MSMS ermöglicht einen maximalen Durchsatz bzw. eine maximale Kapazitätsauslastung. Wird ein konkreter Auftragsbestand betrachtet, entspricht das einer Minimierung des spätesten Fertigstellungszeitpunkts aller Aufträge (Gesamtdurchlaufzeit). Alle drei Größen hängen nicht nur von der Auslegung des MSMS, sondern auch von der operativen Steuerung ab. Zudem müssen in vielen Anwendungsfällen stochastische Einflüsse berücksichtigt werden, insbesondere schwankende Bearbeitungszeiten, aber auch das Ausfallverhalten der Stationen. Daher wird die Zielfunktion durch ein Simulationsmodell des MSMS repräsentiert. Dieses Modell bildet die Grundlage für eine simulationsbasierte Optimierung.

Um die Güte der simulationsbasierten Optimierung einschätzen zu können, wird zum Vergleich ein analytisches Modell mit folgender Zielfunktion gelöst:

$$\min \left| \sum_{m \in M} x_{m,s} \cdot n_s - \frac{1}{|S|} \sum_{s \in S} n_s \right| \quad (3)$$

Die Zielfunktion beruht auf folgender Überlegung: Für jede Station m kann ein Redundanzgrad

$$R_m = \sum_{s \in S} x_{m,s} \cdot n_s \quad (4)$$

angegeben werden, der die Häufigkeit beschreibt, mit der die Fähigkeiten der Station m im MSMS insgesamt vorkommen. Selten vorhandene Fähigkeiten sollen nicht durch Aufträge blockiert werden, die diese Fähigkeiten nicht benötigen. Deshalb sollen alle Stationen einen ähnlichen Redundanzgrad aufweisen. Daher minimiert die Zielfunktion (Gl. 3) die mittlere absolute Abweichung des Redundanzgrads R_m vom mittleren Redundanzgrad

$$\bar{R}_m = \frac{1}{|S|} \sum_{s \in S} n_s. \quad (5)$$

Allerdings fließen wichtige Aspekte bei der Auslegung von MSMS nicht in die Zielfunktion ein: Insbesondere wird nicht angestrebt, Fähigkeiten, die häufig in aufeinanderfolgenden Arbeitsgängen benötigt werden, auf denselben Stationen zu bündeln oder zumindest räumlich benachbarten Stationen zuzuordnen. Es ist daher damit zu rechnen, dass eine simulationsbasierte Optimierung zu besseren Ergebnissen führt als das analytische Modell.

3 Simulationsmodell und Optimierungsverfahren

Die folgenden Abschnitte beschreiben das Simulationsmodell (Abb. 2) sowie das Optimierungsverfahren zur Konfiguration der Stationen von MSMS.

3.1 Simulation von MSMS

Ein MSMS dient der Herstellung einer definierten Menge von Produkten. Jedes Produkt besitzt einen produktspezifischen Arbeitsplan, der eine Folge von Arbeitsgängen beschreibt. Für die Ausführung eines Arbeitsgangs wird jeweils eine bestimmte Fähigkeit benötigt. Die Arbeitsgänge müssen in der Reihenfolge ausgeführt werden, in der sie im Arbeitsplan definiert sind.

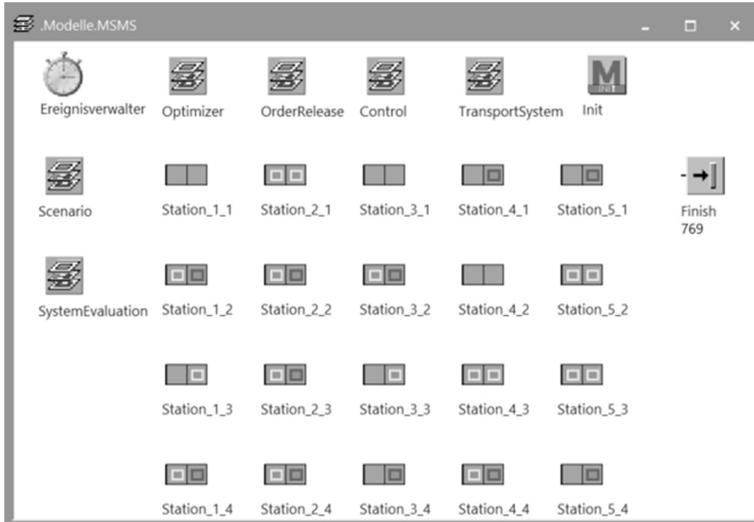


Abbildung 2: Simulationsmodell eines MSMS in Plant Simulation

Aufträge werden durch bewegliche Elemente dargestellt. Jeder Auftrag verlangt die Herstellung einer Einheit eines bestimmten Produkts. Die Aufträge werden vom Baustein *OrderRelease* freigegeben, der eine CONWIP-Steuerung realisiert.

Die *Station* repräsentiert eine Arbeitsstation innerhalb des MSMS. Jede Station besteht aus einem Puffer, der maximal einen Produktionsauftrag aufnimmt, sowie

einem Arbeitsplatz, an dem jeweils ein Auftrag bearbeitet werden kann. Der Arbeitsplatz verfügt über mehrere Fähigkeiten. Rüstzeiten beim Wechsel der eingesetzten Fähigkeit fallen nicht an, da MSMS als Spezialfall flexibler Fertigungssysteme betrachtet werden. Trifft ein Auftrag an einem Arbeitsplatz ein, wird geprüft, ob mit den vorhandenen Fähigkeiten mehrere aufeinanderfolgende Arbeitsgänge ausgeführt werden können. Ist dies der Fall, wird diese Möglichkeit genutzt, um die Anzahl der Transporte zu minimieren.

Nach Abschluss der Bearbeitung eines Auftrags an einem Arbeitsplatz wird die zentrale Steuerung *Control* aufgerufen. Sie identifiziert alle aktuell transportfähigen Aufträge, ermittelt jeweils die Stationen, die den nächsten Arbeitsgang ausführen könnten, wählt eine Station aus und erteilt dem *TransportSystem* die entsprechenden Transportaufträge. Transportfahrzeuge werden im Modell nicht explizit abgebildet, da angenommen wird, dass jedes Werkstück dauerhaft von einem Transportfahrzeug begleitet wird oder sich selbstständig bewegen kann (Burggräf et al. 2020). Die Transportzeit ergibt sich aus der Manhattan-Distanz zwischen den Arbeitsstationen und der Transportgeschwindigkeit. Ein Transport beginnt erst, wenn der Pufferplatz der Zielstation frei ist. Falls nach Beendigung eines Arbeitsgangs keine geeignete Nachfolgestation aufnahmefähig ist, verbleibt der Auftrag auf dem Arbeitsplatz und blockiert diesen, bis ein Transport möglich ist.

Sind für den nächsten Arbeitsgang mehrere Stationen verfügbar, muss eine Auswahl getroffen werden. Im Modell sind dafür folgende Prioritätsregeln implementiert:

- *Höchster Redundanzgrad*: Es wird die Station mit dem höchsten Redundanzgrad R_m (Gl. 4) gewählt. Bei Gleichwertigkeit wird als zweites Kriterium die minimale Transportentfernung herangezogen. Diese Regel zielt darauf ab, seltene Fähigkeiten nicht zu blockieren.
- *Niedrigste Auslastung*: Es wird die Station gewählt, die im bisherigen Simulationsverlauf am niedrigsten ausgelastet wurde. Diese Regel strebt an, die Stationen durch die zugewiesenen Aufträge gleichmäßig zu belasten.

Die Architektur von MSMS schließt das Auftreten von Verklemmungen nicht aus. Es gibt Situationen, in denen ein Auftrag nach der Bearbeitung an einer Station nicht an eine andere Station abgegeben werden kann, weil die Puffer aller geeigneten Stationen belegt sind. Sind diese Stationen ebenfalls blockiert und resultiert daraus eine zyklische Blockade, entsteht eine Verklemmung. Die Steuerung *Control* erkennt Verklemmungen und löst sie auf, indem alle beteiligten Aufträge gleichzeitig zur jeweiligen Nachfolgestation transportiert werden. Dabei wird unterstellt, dass das Transportsystem verklemmungsfrei arbeitet.

Der Simulationsanwender definiert das zu untersuchende *Scenario* durch die Angabe der folgenden Parameter:

- Layout und Transportsystem (Anzahl und Abstände der Stationen, Anzahl der Fähigkeiten pro Station, Transportgeschwindigkeit)
- Produktionsprogramm (Anzahl der Fähigkeiten im System, Anzahl und Nachfragewahrscheinlichkeiten der Produkte, Eigenschaften der produktspezifischen Arbeitspläne)
- Anzahl der Produktionsaufträge
- Operative Steuerung (Höhe des Umlaufbestands für die CONWIP-Regelung, Prioritätsregel zur Auswahl der Folgestation)

Zu einem Layout können verschiedene Produktionsprogramme und zu einem Produktionsprogramm mehrere Auftragsbestände existieren. Der Baustein *Scenario* generiert das Layout und die Auftragsdaten auf Basis der gegebenen Parameter.

3.2 Simulationsbasierte Optimierung

Ziel der simulationsbasierten Optimierung ist es, die Fähigkeiten so auf die Arbeitsstationen zu verteilen, dass die Gesamtdurchlaufzeit zur Bearbeitung eines gegebenen Auftragsbestands minimiert wird. Der Baustein *Optimizer* realisiert dafür eine Kombination aus Tabu Search und Simulated Annealing (Chopard und Tomassini 2018): Zu Beginn eines Optimierungslaufs wird zunächst eine zulässige Verteilung der Fähigkeiten auf die Stationen als Startlösung generiert. Ausgehend von der Startlösung werden iterativ neue Lösungen durch lokale Veränderungen erzeugt. Bei jeder lokalen Veränderung werden zufällig zwei Stationen gewählt und zwei ebenfalls zufällig gewählte Fähigkeiten zwischen den Stationen ausgetauscht. Falls die so entstandene Lösung in einer mitgeführten Tabu-Liste der zuletzt untersuchten Lösungen enthalten ist, wird sie verworfen. Andernfalls wird in einem Simulationslauf die Güte der Lösung – also die Gesamtdurchlaufzeit des Auftragsbestands – ermittelt und die stochastische Akzeptanzregel von Simulated Annealing angewandt. Die Optimierung endet, wenn eine gegebene Anzahl aufeinanderfolgender Iterationen keine Verbesserung erreicht.

4 Simulationsstudie zur Auslegung von MSMS

Das analytische Modell zur Auslegung von MSMS wurde in CPLEX Optimization Studio 12.8.0 realisiert und das Verfahren zur simulationsbasierten Optimierung in Tecnomatix Plant Simulation 16.0.3 umgesetzt.

4.1 Untersuchte Szenarien

Gegenstand der Untersuchung ist ein MSMS mit 20 Stationen, die in einem 4×5-Raster mit einem Rasterabstand von 5 m angeordnet sind. Die Transporte erfolgen mit einer Geschwindigkeit von 1,5 m/s. Das MSMS stellt 10 verschiedene Produkte her, die mit gleicher Wahrscheinlichkeit nachgefragt werden. Insgesamt verfügt das MSMS über 20 verschiedene Fähigkeiten, wobei jede Station 4 Fähigkeiten besitzt.

Die Arbeitspläne der Produkte umfassen jeweils 10 Arbeitsgänge. Da sich die in einem Produktionssystem hergestellten Produkte technologisch meist ähneln, werden die produktspezifischen Arbeitspläne folgendermaßen generiert: Als Vorlage dient eine zufällige Sequenz von 10 Fähigkeiten. Ein produktspezifischer Arbeitsplan entsteht, indem die Vorlage an zwei zufällig gewählten Stellen modifiziert wird, d. h. zwei der benötigten Fähigkeiten werden durch andere ersetzt. Dabei wird sichergestellt, dass jede Fähigkeit des MSMS mindestens von einem Produkt in Anspruch genommen wird. Die Bearbeitungszeiten der Arbeitsgänge sind Erlang-verteilt mit einem Mittelwert von 60 s und einer Standardabweichung von 20 s. In jedem Simulationslauf wird die Bearbeitung von 1.000 Produktionsaufträgen simuliert. Die Zufallszahlen für die Bearbeitungszeiten sind in den auftragsspezifischen Arbeitsplänen hinterlegt, so dass sich verschiedene Simulationsläufe mit unterschiedlichen Auslegungen des MSMS in dieser Hinsicht nicht unterscheiden.

Insgesamt wurden 10 Szenarien erzeugt. Für jedes Szenario wurden 500 zufällige Auslegungen generiert, eine Auslegung mit dem analytischen Modell bestimmt und 10 Auslegungen durch Simulated Annealing ermittelt.

Einige Parameter wurden in Voruntersuchungen festgelegt und in den Simulationsexperimenten nicht variiert. Für die CONWIP-Steuerung hat sich ein Umlaufbestand von 30 Aufträgen als geeignet erwiesen. Der Temperatur-Parameter von Simulated Annealing wird mit dem Zielfunktionswert der Startlösung initialisiert und in jeder Iteration um 0,5 % reduziert. Die Tabu-Liste umfasst die letzten 100 Lösungen. Die Optimierung wird nach 200 Iterationen ohne Verbesserung abgebrochen.

4.2 Simulationsergebnisse

Tabelle 1 zeigt die Gesamtdurchlaufzeiten für die Erfüllung der Produktionsaufträge unter Verwendung der Auslegungen, die mit den verschiedenen Verfahren jeweils bei Einsatz beider Prioritätsregeln bestimmt wurden. Für die zufällige Auslegung und Simulated Annealing wird jeweils der Mittelwert aus 500 bzw. 10 Simulationsläufen angegeben. Die von Simulated Annealing gefundenen Auslegungen sind im Durchschnitt über alle Szenarien und beide Steuerstrategien etwa 10 % besser als die mittlere zufällige Auslegung. Das analytische Modell hat sich dagegen nicht bewährt: Die mit diesem Modell erzeugten Auslegungen sind eher schlechter als zufällig gewählte Auslegungen.

Tabelle 1: Simulationsergebnisse: Gesamtdurchlaufzeit (hh:mm)

Szenario	Höchster Redundanzgrad			Niedrigste Auslastung		
	Zufällige Auslegung	Analytisches Modell	Simulated Annealing	Zufällige Auslegung	Analytisches Modell	Simulated Annealing
1	11:46	12:10	10:40	11:24	11:30	10:45
2	12:47	13:23	11:06	12:46	12:38	11:27
3	11:43	12:08	11:02	11:42	11:49	11:13
4	13:43	17:47	11:25	12:36	12:37	11:25
5	12:35	13:58	10:57	12:17	12:43	10:50
6	11:31	11:55	10:51	11:31	11:38	10:59
7	12:37	12:48	11:20	12:21	12:35	11:19
8	12:56	14:47	11:04	12:30	12:44	11:06
9	13:38	14:22	11:56	14:13	14:58	12:11
10	11:60	13:01	10:54	11:47	12:06	10:58
∅	12:32	13:38	11:07	12:19	12:32	11:13

Wird die mittlere Gesamtdurchlaufzeit über alle Szenarien betrachtet, haben die Prioritätsregeln nur geringen Einfluss auf die Güte der Auslegungen, die zufällig erzeugt oder von Simulated Annealing gefunden wurden. Das könnte ein Indiz dafür sein, dass Auslegungen vergleichsweise unabhängig von den Regeln zur operativen Steuerung von MSMS festgelegt werden können. Allerdings gilt dies nur in der

Gesamtbetrachtung, für einzelne Szenarien liefern die beiden Prioritätsregeln durchaus unterschiedliche Ergebnisse. Auch die Resultate für das analytische Modell deuten in eine andere Richtung: Bei diesen Auslegungen hat ein Wechsel der Prioritätsregel deutliche Auswirkungen auf die Systemleistung.

Abbildung 3 zeigt exemplarisch für ein Szenario die Verteilung der Zielfunktionswerte für die zufällig gewählten Auslegungen des MSMS. Zum Vergleich sind die Zielfunktionswerte t_{AM} und t_{SA} für Auslegungen, die durch das analytische Modell (AM) und von Simulated Annealing (SA) gefunden wurden, eingetragen. Simulated Annealing liefert stabil gute Auslegungen: Der Variationskoeffizient des Zielfunktionswerts aus den 10 Optimierungsläufen liegt im Durchschnitt aller Szenarien bei 1,2 %. Daher ist in Abbildung 3 nur die mittlere Güte dieser 10 Auslegungen eingetragen.

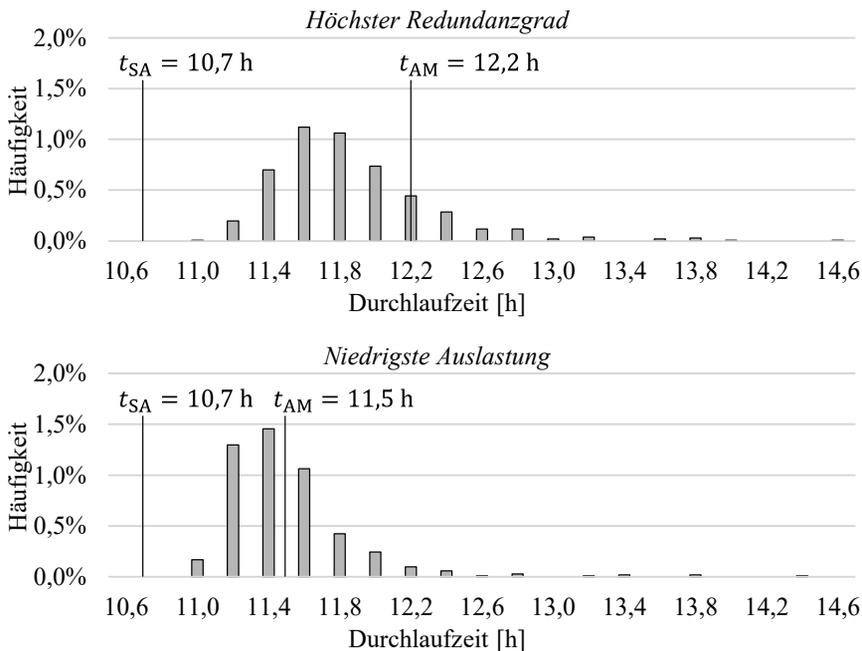


Abbildung 3: Verteilung der Gesamtdurchlaufzeit der Auslegungen für ein Szenario

Die Prioritätsregel *Niedrigste Auslastung* führt tendenziell zu etwas kürzeren Gesamtdurchlaufzeiten als die Regel *Höchster Redundanzgrad*, wobei die von Simulated Annealing gefundenen Auslegungen robuster gegenüber der Steuerstrategie sind. Demgegenüber sind die mit dem analytischen Modell gefundenen Auslegungen sehr empfindlich gegenüber der eingesetzten Prioritätsregel. Daraus ist zu schließen, dass der im analytischen Modell nicht berücksichtigte Aspekt der Häufung bestimmter Abfolgen von Fähigkeiten in den Arbeitsplänen und damit die Möglichkeit der

Bündelung aufeinanderfolgender Arbeitsgänge auf derselben Station für die Auslegung von MSMS eine wesentliche Rolle spielt. Dieses Optimierungsziel wird von Simulated Annealing implizit berücksichtigt.

Abbildung 4 zeigt beispielhaft für ein Szenario die Auslastung der Stationen für eine Auslegung, die mit dem analytischen Modell festgelegt wurde, sowie für die beste mit Simulated Annealing gefundene Konfiguration. Obwohl das analytische Modell die Fähigkeiten so auf die Stationen verteilt, dass eine gleichmäßige Auslastung der Stationen begünstigt wird, verteilt nur die Prioritätsregel *Niedrigste Auslastung* die Arbeit tatsächlich gleichmäßig auf die Stationen, während die Prioritätsregel *Höchster Redundanzgrad* zu einer deutlichen Ungleichverteilung führt. Die von Simulated Annealing gefundene Konfiguration liefert für beide Prioritätsregeln nahezu gleich gute Ergebnisse (mittlere Auslastung von 77,6 % für *Höchster Redundanzgrad* bzw. 77,0 % für *Niedrigste Auslastung*). Demgegenüber ist die analytisch ermittelte Konfiguration empfindlicher gegenüber der Prioritätsregel (mittlere Auslastung 67,0 % für *Höchster Redundanzgrad* bzw. 74,5 % für *Niedrigste Auslastung*).

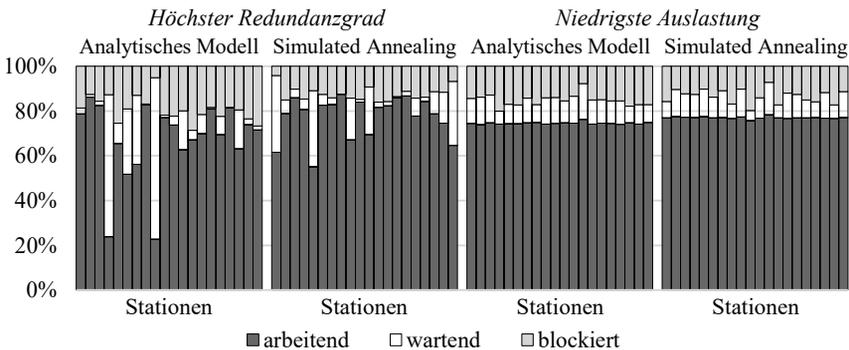


Abbildung 4: Einfluss der Steuerstrategie auf die Auslastung der Stationen

5 Zusammenfassung und Ausblick

Der Beitrag untersucht den Einfluss der Verteilung von Fähigkeiten auf die Stationen von MSMS auf die Leistungsfähigkeit derartiger Systeme und stellt ein Verfahren zur simulationsbasierten Konfiguration der Stationen von MSMS vor. Das Verfahren wurde in Tecnomatix Plant Simulation realisiert und in Simulationsexperimenten getestet. Es konnte gezeigt werden, dass das Verfahren zuverlässig Auslegungen findet, die die zufällige Systemkonfigurationen deutlich übertreffen und robust gegenüber einem Wechsel zwischen den beiden betrachteten Prioritätsregeln zur Auftragsverteilung sind. Als Vergleich wurde ein mathematisches Modell zur Optimierung der Auslegung herangezogen, das sich jedoch nicht bewährt hat.

Der Beitrag bildet die Grundlage für Folgearbeiten, in denen unterschiedliche Steuerstrategien und die Auswirkungen von Materialbereitstellung und Rüstzeiten in MSMS betrachtet werden. Dabei ist auch eine Sensitivitätsanalyse hinsichtlich der charakteristischen Eigenschaften des MSMS, z. B. das Verhältnis von Transport- und Bearbeitungszeit, vorgesehen. Außerdem kann die zugrundeliegende Optimierung durch andere Zielfunktionen und weitere Nebenbedingungen, z. B. die eingeschränkte Kombinierbarkeit von Fähigkeiten, ergänzt werden.

Literatur

- Bányai, Á.; Illés, B.; Glistau, E.; Coello Machado, N. I.; Tamás, P.; Manzoor, F.; Bányai, T.: Smart cyber-physical manufacturing: Extended and real-time optimization of logistics resources in matrix production. In: *Applied Sciences* 9 (2019) 7.
- Burggräf, P.; Dannapfel, M.; Adlon, T.; Kahmann, H.; Schukat, E.; Keens, J.: Capability-based assembly design: An approach for planning an agile assembly system in automotive industry. In: *Procedia CIRP* 93 (2020), S. 1206-1211.
- Chopard, B.; Tomassini, M.: *An Introduction to Metaheuristics for Optimization*. Cham: Springer 2018.
- ElMaraghy, H. A.: Flexible and Reconfigurable Manufacturing Systems Paradigms. *International Journal of Flexible Manufacturing Systems* 17 (2005), S. 261-276.
- Filz, M.-A.; Gerberding, J.; Herrmann, C.; Thiede, S.: Analyzing different material supply strategies in matrix-structured manufacturing systems. In: *Procedia CIRP* 81 (2019), S. 1004–1009.
- Filz, M.-A.; Hermann, C.; Thiede, S.: Simulation-based Data Analysis to Support the Planning of Flexible Manufacturing Systems. In: *Simulation Notes Europe SNE* 30 (2020) 4, S. 131-137.
- Göppert, A.; Rachner, J.; Schmitt, R. H.: Automated scenario analysis of reinforcement learning controlled line-less assembly systems. In: *Procedia CIRP* 93 (2020), S. 1091–1096.
- Greschke, P.: *Matrix-Produktion als Konzept einer taktunabhängigen Fließfertigung*. Dissertation Technische Universität Braunschweig, Institut für Werkzeugmaschinen und Fertigungstechnik, Essen: Vulkan 2016.
- Hofmann, C.; Krahe, C.; Stricker, N.; Lanza, G.: Autonomous production control for matrix production based on deep Q-learning. In: *Procedia CIRP* 88 (2020), S. 25–30.
- Kern, W.; Rusitschka, F.; Bauernhansl, T.: Planning of workstations in a modular automotive assembly system. In: *Procedia CIRP* 57 (2016), S. 327–332.
- May, M. C.; Kiefer, L.; Kuhnle, A.; Stricker, N.; Lanza, G.: Decentralized Multi-Agent Production Control through Economic Model Bidding for Matrix Production Systems. In: *Procedia CIRP* 96 (2021), S. 3–8.
- Müller, D.; Schmitt, T. V.: Production planning in autonomous and matrix-structured assembly systems: Effects of similarity of precedence graphs on order release sequencing. In: *Procedia CIRP* 93 (2020), S. 1358–1363.
- Ranke, D.; Bauernhansl, T.: Evaluation of Material Supply Strategies in Matrix Manufacturing Systems. In: Weißgraeber, P.; Heieck, F.; Ackermann, C. (Hrsg.): *Advances in Automotive Production Technology – Theory and Application*. Berlin, Heidelberg: Springer Vieweg 2021, S. 80-88.
- Schönemann, M.; Herrmann, C.; Greschke, P.; Thiede, S.: Simulation of matrix-structured manufacturing systems. In: *Journal of Manufacturing Systems* 37 (2015) 1, S. 104–112.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Rückwärtssimulation als Instrument zur Produktionsplanung – Erkenntnisse aus einer praxisbezogenen Fallstudie

Backward Simulation for Production Planning – Recent Advances in a Real-World Use-Case

Christoph Laroque, Madlene Leißau, Westsächsische Hochschule Zwickau,
Zwickau (Germany), christoph.laroque@fh-zwickau.de, madlene.leissau.gel@fh-
zwickau.de

Wolfgang Scholl, Germar Schneider, Infineon Technologies Dresden GmbH & Co.
KG, Dresden (Germany), wolfgang.scholl@infineon.com,
germar.schneider@infineon.com

Abstract: The focus on customer orientation and on-time production as well as delivery characterise the competitive environment for manufacturing companies. Customer-specific products must be manufactured and delivered in specified lead times, according to promised delivery dates. In this context, questions as to whether the production programme is feasible and whether all previously promised delivery dates will be met are often answered with backward-oriented planning approaches without taking into consideration any uncertainty or alternatives, that arise during operations. Regarding complex manufacturing systems (here semiconductors with re-entry cycles), these questions can be answered in a more detailed and robust manner by a discrete event-based simulation (DES) approach used in a backward-oriented manner. Research results show that the taken approach can be applied successfully for the scheduling of customer-specific orders in a real-world setting.

1 Motivation

Die Intensivierung globaler Unternehmung und eine stärker kundenorientierte sowie termingerechte Produktion und Logistik sind prägende Wettbewerbsfaktoren für produzierende Unternehmen. Die Entwicklung der Vision von Industrie 4.0 – zukünftig wohl auch Industrie 5.0 (Breque et al. 2021) – und das Konzept einer „smarten Fabrik“ für individualisierbare Produkte in kleinen Losgrößen stellen die Arbeitsvorbereitung wie auch die operative Produktionsplanung fortwährend vor neue Herausforderungen. Moderne, komplexe und hochautomatisierte Produktionssysteme müssen aber in einem möglichst „optimalen Betriebszustand“ betrieben werden, um wirtschaftlich erfolgreich zu sein. Zugesagte Liefertermine und in

Rahmenvereinbarungen festgelegte Durchlaufzeiten sind sicherzustellen und bedingen in der täglichen Durchführung eine permanente (wirksame) Anpassung der Produktionsplanung und -steuerung. Alle weiteren Rahmenbedingungen einer wirtschaftlichen Produktion bleiben unverändert und gelten auch weiterhin.

Im Vergleich zu anderen Branchen besitzen die Produktionssysteme und -prozesse der in diesem Beitrag adressierten Halbleiterfertigung ein überdurchschnittlich hohes Maß an Komplexität. Die eingesetzten Fertigungstechnologien im Mikro- und Nanometerbereich sind im Hinblick auf die Prozessstabilität sehr sensitiv und beinhalten komplexe Steuerungslogiken. In Abhängigkeit verschiedener im Vorfeld festgesetzter Merkmale erfordern einzelne Produktionslose teilweise mehr als tausend Prozessschritte. Stellenweise müssen einzelne Produktionslose innerhalb des bestellten Produktmix mehrfach mit hohem Automationsgrad und unter Reinraumbedingungen über spezielle und teilweise dieselben Maschinen und Transportrouten prozessiert werden (Re-Entry-Cycles). Die Komplexität resultiert nicht selten in Ausschuss angefertigter Produkte in relevanter Größenordnung, der kurzfristig durch zusätzliche Einschleusungen ausgeglichen werden muss. Fragestellungen im Rahmen der Produktionsfeinplanung können mittels existierender Werkzeuge zur Generierung von Ablaufplänen und dem folgend herkömmlicher Planungsverfahren in den meisten Fällen nur unzureichend oder gar nicht beantwortet werden. Nicht kalkulierbare Rückwirkungen einer solchen mangelhaften Ablaufplanung haben erheblichen Einfluss auf die Gesamtleistung und bergen mit Blick auf die zunehmende Konkurrenzsituation und die eigene Stellung am Markt ein immenses Optimierungspotential für produzierende Unternehmen im Bereich der Halbleiterfertigung wie auch allgemein.

Im Rahmen des EU-ECSEL-Forschungsprojektes iDEV40 (Darstellung des Gesamtprojektes unter www.idev40.eu) werden Einsatzszenarien der rückwärtsorientierten Materialflusssimulation in der Halbleiterfertigung erschlossen. In relevanten Vorarbeiten konnte die prinzipielle Machbarkeit dieser Methode auf Fabrikebene zwar nachgewiesen werden (Arakawa et al. 2002; Graupner et al. 2004), musste im Projekt aber auf die Spezifika der Halbleiterbranche adaptiert und neu verprobt werden. Der hier vorliegende Beitrag beschreibt als wesentliche Erweiterung zu eigenen Vorarbeiten der Autoren (vgl. bspw. Scholl et al. 2014) und einer ersten Publikation aus dem Forschungsprojekt mit einem Testmodell (Laroque et al. 2020) neuere Ergebnisse auf Basis eines „real-world use-case“ und zeigt damit spezifischer als bisher Potentiale und Grenzen der Methode auf. Nach einer kurzen Darstellung des wissenschaftlichen Stands der Technik sowie einer Erläuterung des prinzipiellen Lösungsansatzes sollen Projektergebnisse detailliert beschrieben und dargestellt werden. Eine Zusammenfassung beschreibt abschließend die nächsten Schritte im Projekt.

2 Rückwärtssimulation

Entsprechend einer Sicherstellung der Wettbewerbsfähigkeit konzentrieren sich die Produktionsplanung und -steuerung (PPS) heute auf eine Reihe von Schlüsselmaßnahmen. Darunter zählen neben der Verkürzung der Durchlaufzeit, der Erfüllung von Qualitätsanforderungen bei möglichst niedrigen Beständen und der Einhaltung zugesagter Liefertermine gleichermaßen die Erhöhung des Durchsatzes sowie der aktuellen Verfügbarkeit und Wertschöpfung einzelner Produktionsanlagen

(Overall Equipment Effectiveness). Diese Zielgrößen werden entscheidend durch Produktionsplanungs- und Terminierungsmechanismen beeinflusst, wodurch übergreifende Optimierungsansätze notwendig werden, um eine merkliche Verbesserung herbeizuführen (Jain und Chan 1997).

Als Ergänzung zu existierenden Methoden der gemischten-ganzzahligen Optimierung, unterschiedlichen (simulationsbasierten) Heuristiken, einfacher Vorwärts- oder Rückwärtsterminierung (mit oder ohne Kapazitätsbeschränkungen) und Methoden des maschinellen Lernens (bspw. Kuhnle et al., 2019 und 2021) beschäftigt sich dieser Beitrag mit einem Ansatz der diskreten ereignisorientierten Simulation (DES) in Form der Rückwärtssimulation. Modelle zur DES können eine genaue Nachahmung eines realen Systems entsprechend dessen Funktionsweise über die Zeit darstellen, sind gut zu parametrisieren und berücksichtigen die Variabilität der Realität, indem zufällige Einwirkungen über stochastische Bestandteile in die Modelle integriert werden können (Banks 1998; Law und Kelton 2000). Darüber hinaus kann DES gleichauf verschachtelte Ressourcenbeziehungen, Wartungsvorgänge und spezifische Ablauf-, Prioritäts-, Batch- oder Rüstregeln abbilden. Ausgehend davon und einer Eingabe eines konkreten Produktionsprogramms in ein solches Simulationsmodell, kann DES angewendet werden, um Fragestellungen bezüglich einer Realisierbarkeit des Produktionsprogramms und einer Einhaltung im Vorfeld zugesagter Liefertermine (bestmöglich) zu beantworten. Eine derartige Anwendung eignet sich dabei besonders für die Planung im Bereich der Halbleiterfertigung.

Der Ansatz zur Untersuchung zeitlich rückwärtsgerichteter Planungsprobleme (im Folgenden: Rückwärtssimulation) konkretisiert eine Umkehrung der Ablauflogik einer Simulation und die daraus resultierende rückwärtsgerichtete Ausführung derselben. Danach kommen die Vorteile der Simulation auch in der Anwendung der rückwärtsorientierten Planung zum Tragen (Huang und Wang 2009; Schumacher und Wenzel 2000). Die Rückwärtssimulation kann nach Jain und Chan (1997) als ein effizientes Werkzeug für die Implementierung der Rückwärtsterminierung angesehen werden. Dem folgend vereint eine Reihenfolge- und Einschleusplanung nachkommend der Rückwärtssimulation die Lösungsqualität herkömmlicher Terminierungsansätze und die Ausführungsgeschwindigkeit simulationsbasierter Terminierungsansätze. Erste Anwendungsstudien, bei denen Aufträge mit Hilfe einer Rückwärtssimulation zeitlich rückwärtsgerichtet eingeplant werden, sind bereits seit mehr als fünfzehn Jahren verfügbar. Watson et al. (1993 und 1997), Ying und Clark (1994) und Jain und Chan (1997) nutzen solche Verfahren, um die Freigabezeitpunkte von Aufträgen oder Losen auch unter stochastischen Charakteristika der Modelle berechnen zu können.

Eine Modellierung einer rückwärtsgerichteten Ausführung einer Ablaufsimulation verlangt im Vorfeld einige sorgfältige Überlegungen, um eine korrekte Umkehrung im Kontext des abzubildenden Materialflusses vornehmen zu können und sich von der Denkweise einer vorwärtsgerichteten Modellierung zu lösen. Diese Überlegungen beziehen sich insbesondere auf eine Umkehrung einzelner Produktionsprozesse (bspw. einer Montage in eine Demontage) und vorgesehener Steuerungsregeln (Jain und Chan 1997). Letztgenannte Steuerungsregeln lassen sich jedoch nicht immer einzu-eins in das entsprechende Rückwärts-Pendant übertragen. Analog zur Rückwärtsterminierung in PPS-Systemen wird auch die Rückwärtssimulation in Kombination mit vorwärtsgerichteten Simulationsläufen durchgeführt, um die

resultierenden Pläne nochmals abzusichern. Nach Graupner et al. (2004) können anhand einer derartigen kombinierten Ausführung die Vorteile beider Simulationen vereinigt mögliche Modellierungsdiskrepanzen der Rückwärtsimulation ausgeglichen werden.

Eine Modellierung einer rückwärtsorientierten Ausführung einer Ablaufsimulation zieht immer auch eine Umkehrung der Quelle-Senke-Beziehung nach sich. Konkret heißt das, dass Aufträge oder Lose an den Stellen in das System eingeschleust werden, an der sie dieses in der vorwärtsorientierten Ausführung verlassen. Andersherum verlassen sie das Rückwärts-Pendant an den Einschleusstellen der vorwärtsgerichteten Ablaufsimulation („from product to raw material“). Die Rückwärtsimulation ist dennoch nicht als reine „Umkehrfunktion“ der Vorwärtsimulation zu verstehen. So verweisen Ying und Clark (1994) darauf, dass Vorwärts- und Rückwärtsimulation zur selben berechneten Simulationszeit nicht denselben Zustand aufweisen müssen.

In der Domäne der Halbleiterfertigung konnten durch die Autoren in den vergangenen Jahren erste Beispiele erfolgreich realisiert und publiziert werden (Scholl et al. 2014; Laroque et al. 2020). Nachfolgend soll das in diesem Beitrag betrachtete Modell daran anknüpfend nun ebenso spezielle Eigenschaften der Halbleiterfertigung (s.o.) berücksichtigen.

3 Erzielte Ergebnisse

Das Modell eines Vormontageprozesses aus dem Bereich der Halbleiterfertigung (Abb. 1) stellt den Ausgangspunkt für die nachkommend beschriebenen Untersuchungen im Zuge der Rückwärtsmodellierung und -simulation dar.

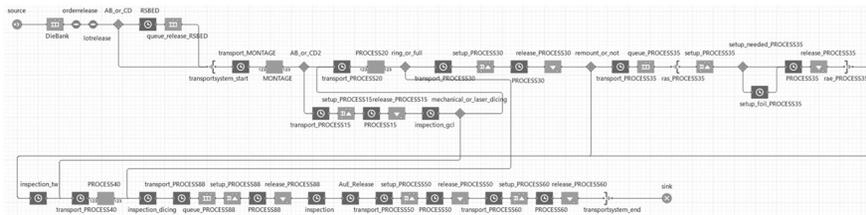


Abbildung 1: Übersicht der Modellstruktur des realen Modellbeispiels

Im ausgewählten Prozessabschnitt kommen verschiedene Maschinenanlagen unterschiedlicher Komplexität zum Einsatz, die im Wesentlichen innerhalb des Hauptmodells MAIN abgebildet wurden. Komplexe Teilprozesse in Form der Maschinenanlagen MONTAGE, PROCESS20 und PROCESS40 wurden in gleichnamigen Teilmodellen abgebildet und in das Hauptmodell eingebettet. Die Teilmodelle (mit jeweils bis zu 26 Instanzen) folgen dabei einem entwickelten Regelwerk, das eine sogenannte Same-Setup-Regel nachempfunden und auf diese Weise die eintreffenden Lose innerhalb der Simulation koordinieren und bearbeiten soll. Die Umrüstung zur Verfügung stehender (freier) Teilmaschinen erfolgt danach gemäß verschiedener Variablen in Abhängigkeit eines losbezogenen Parameters.

Startpunkt des vorwärtsgerichteten Simulationsmodells (Startpunkt der Untersuchung) bilden Auftragsdaten, die in Tabellenform über einen Zeitraum von Dezember 2019 bis April 2020 in das Modell eingeschleust werden. Die Aufträge werden hierbei durch eine Reihe von Parametern definiert (beispielsweise *processgroup*, *basictype* oder *producttype*), um eine Zuordenbarkeit der eintreffenden Lose über den gesamten Simulationslauf hinweg zu gewährleisten. Entsprechend ergeben sich in Abhängigkeit des Produkttypen (Parameter *producttype*) eines Auftrages spezielle Routen. Möglich sind im Zuge dessen beispielsweise diverse bedingungs-basierte Aufteilungen des ausgewählten Prozessabschnittes nach *Product A*, *Product B* und *Product C* sowie *Product D*.

In Folge der Modellierung einer ereignisdiskreten Rückwärtssimulation (Abb. 2) diene das Hauptmodell MAIN als Ausgangspunkt. Die eingebetteten Teilmodelle MONTAGE, PROCESS20 und PROCESS40 werden hierbei genauso wie weniger komplexe Maschinenanlagen (beispielsweise PROCESS50 und PROCESS60) in umgekehrter Reihenfolge betrachtet und modelliert. Die daraus resultierende ereignisdiskrete Rückwärtssimulation hat auftragsspezifische Angaben aus der vorwärtsgerichteten Modellierung übernommen und diene als Grundlage verschiedener umfangreicher Simulationsexperimente, um zeitliche Diskrepanzen im Rahmen der Simulation weitmöglichst ausschließen zu können. Ein Teil der daraus hervorgehenden Ergebnisse soll im Folgenden dargestellt werden.

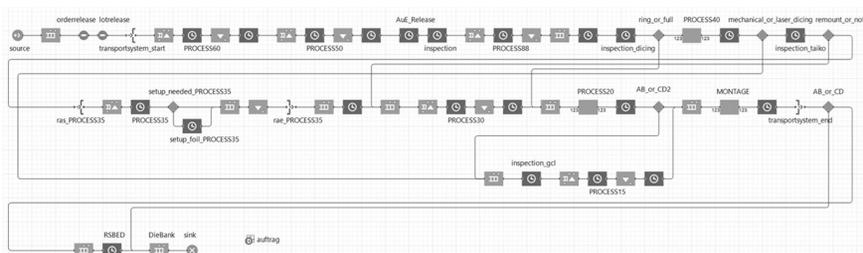


Abbildung 2: Modellstruktur der ereignisdiskreten Rückwärtssimulation

Ausgangspunkt für die experimentellen Versuche bilden Liefertermine (Due Dates), die durch einen vorwärtsgerichteten Simulationslauf generiert wurden (basierend auf den originär hinterlegten Auftragsdaten). Die resultierenden Liefertermine bilden in Folge einer Anpassung die Eingabedaten für die eigentliche Rückwärtssimulation. Die Notwendigkeit einer Anpassung der resultierenden Liefertermine ist darin begründet, dass bestehende Simulationswerkzeuge den zugrundeliegenden Planungszeitraum einer Simulation zeitlich vorwärts betrachten. Eine „simple“ Eingabe der resultierenden Liefertermine in zeitlich absteigender Reihenfolge in das Rückwärts-Pendant des Simulationsmodells bleibt demnach ohne positives Ergebnis, sodass die aus der Vorwärtssimulation generierten Liefertermine zunächst in Bezug auf eine rückwärtsgerichtete Ausführung der Ablaufsimulation anzupassen waren. Die Anpassung der resultierenden Liefertermine, die im Weiteren als „reale“ Liefertermine angenommen werden, hatte gegenüber der Einschleusplanung *VWS0* eine Verzögerung des Starts der Simulationszeit (Dezember 2019 zu April 2020) ausgehend des zuletzt fertiggestellten Auftrages (JobNr. 2119) zur Folge. Für die

weitere Anpassung wurde danach der Liefertermin des zuletzt fertiggestellten Auftrages jeweils mit der Differenz aus dessen *dueDateVWS0* und einem *dueDateVWS0* eines früheren Liefertermins addiert. Die resultierende Einschleusplanung *RWSI* (Abb. 3) wurde anschließend für eine rückwärtsberechnete Ausführung in das Simulationsmodell eingegeben. Im Kontext einer erneuten Anpassung und einer Einschleusplanung *VWSI* (Abb. 4) kann dann eine Beantwortung der Fragestellung erfolgen, wann konkrete Produktionsaufträge einzuplanen sind, um zugesagte Liefertermine hinsichtlich einer zeitlichen Realisierbarkeit und Erfüllung sicherzustellen.



Abbildung 3: Anpassung gemäß einer Einschleusplanung *RWSI*



Abbildung 4: Anpassung gemäß einer Einschleusplanung *VWSI*

Für eine Kategorisierung und Auswertung aus der Rückwärtssimulation und einer anschließenden Vorwärtssimulation (je ein Simulationslauf nach erfolgreicher Anpassung der Eingabedaten im deterministischen Fall) resultierender Liefertermine wurden die Zeitintervalle *viel früher*, *früher*, *pünktlich*, *später* und *viel später* definiert. Diese Zeitintervalle stehen in Abhängigkeit zu den ursprünglichen Due Dates. In das Zeitintervall *viel früher* werden alle Aufträge eingeordnet, welche durch die Einplanung der Rückwärtssimulation eine Verfrühung $t > 6h$ aufweisen. Dagegen kategorisiert das Zeitintervall *viel später* Aufträge, die eine Verspätung $t > 6h$ anzeigen. Die Zeitintervalle *früher* und *später* weisen analog auf eine vorzeitige Ankunft bzw. Verspätung von $t \leq 6h$ sowie $t > 3h$ hin, während das Zeitintervall *pünktlich* alle Aufträge innerhalb von $t \leq 3h$ rund um den „realen“ Liefertermin einschließt.

In einem ersten Schritt und mit dem Ziel die Lösungsgüte der aus der Rückwärtssimulation ermittelten Einschleusplanung *VWSI* mit konkurrierenden Planungsverfahren der kapazitätsbeschränkten Rückwärtsterminierung zu vergleichen, wird die Einschleusplanung aus der Rückwärtssimulation, einer Rückwärtsterminierung *RT1* gemäß der Differenz aus *dueDatesVWS0* und *avCycleTimeVWS0* sowie einer Rückwärtsterminierung *RT2* gemäß der Differenz aus *dueDatesVWS0* und *avCycleTimeVWS0* nach *producttype* gegenübergestellt (Tab.1).

Table 1: Vergleich der Rückwärtssimulation mit Planungsverfahren deterministischer Rückwärtsterminierung

	viel früher	[%]	früher	[%]	pünktlich	[%]	später	[%]	viel später	[%]
VWS1	668	31,49	489	23,06	874	41,21	43	2,03	47	2,22
RT1	1848	87,13	0	0,00	0	0,00	0	0,00	273	12,87
RT2	1987	93,68	38	1,79	42	1,98	7	0,33	47	2,22

Die Ergebnisse der gegenübergestellten methodischen Ansätze zur Generierung einer Einschleusplanung durch Rückwärtssimulation und konkurrierenden Planungsverfahren deterministischer Rückwärtsterminierung verweisen darauf, dass die Methode der Rückwärtssimulation die Einschleusplanung im Zusammenhang mit den resultierenden Lieferterminen sehr viel zuverlässiger gestalten kann. Folglich werden gemäß einer Einschleusplanung *VWS1* 41,21 Prozent aller Aufträge in einem Zeitintervall $t \pm 3h$ pünktlich fertiggestellt, während die Planungsverfahren deterministischer Rückwärtsterminierung hier lediglich einen Wert von 0,00 Prozent gemäß einer Einschleusplanung *RT1* und einen Wert von 1,98 Prozent gemäß einer Einschleusplanung *RT2* aufweisen. Zusätzlich lassen die Ergebnisse, mehr oder weniger, deutliche Vorsprünge auf Seiten einer Einschleusplanung durch Rückwärtssimulation im Hinblick auf ein Zeitintervall $t > 6h$ (vor und nach dem angenommenen „realen“ Liefertermin) erkennen, das heißt, dass die resultierenden Liefertermine sehr viel genauer um die festgesetzten (oder angenommenen) Liefertermine streuen.

Im Weiteren wird die Lösungsgüte der aus der Rückwärtssimulation ermittelten Einschleusplanung *VWS1* einer Einschleusplanung mit stochastischer Rückwärtssimulation *mSTO* (Tab. 2) sowie nach mehreren Versuchsläufen *STO1* bis *STO10* (Tab. 3) gegenübergestellt (*mSTO* ist hier der Mittelwert der ermittelten Einschleustermine). Für die korrespondierenden Simulationsläufe (vor- und rückwärtsorientiert) werden dabei gleiche Seed-Werte verwendet, um eine Vergleichbarkeit der Simulationsläufe zu gewähren. In Folge der Abbildung stochastischer Bearbeitungszeiten und den daraus resultierenden Werten lassen sich sowohl für einzelne Simulationsläufe als auch für deren Mittel tendenziell mehr Verspätungen im Vergleich zu der deterministischen Rückwärtssimulation aufzeigen.

Table 2: Ergebnisse stochastischer Rückwärtssimulation im Mittel

	viel früher	[%]	früher	[%]	pünktlich	[%]	später	[%]	viel später	[%]
VWS1	668	31,49	489	23,06	874	41,21	43	2,03	47	2,22
mSTO	759	27,31	467	22,01	842	39,70	127	5,98	106	5,00

Table 3: Ergebnisse stochastischer Rückwärtssimulation nach mehreren Versuchsläufen (in Prozent, gerundet)

	viel früher	früher	pünktlich	später	viel später
VWS1	31,49	23,06	41,21	2,03	2,22
STO1	25,79	21,83	41,96	5,19	5,23
STO2	27,58	22,30	34,84	9,85	5,42
STO3	27,96	23,24	40,55	3,63	4,62
STO4	27,02	21,59	41,25	5,19	4,95
STO5	26,54	23,57	38,76	5,89	5,23
STO6	28,15	21,17	40,17	5,80	4,71
STO7	28,76	19,94	39,23	7,45	4,62
STO8	26,59	22,21	41,49	4,38	5,33
STO9	26,78	22,11	41,54	4,38	5,19
STO10	28,34	20,56	38,52	8,72	3,87

Allgemein muss beachtet werden, dass die Kategorisierung der resultierenden Liefertermine und die Definition der hier beschriebenen Zeitintervalle gegenüber dem angenommenen originären Liefertermin nicht (immer) optimal ist. Entsprechend ordnen sich die resultierenden Liefertermine der Aufträge sowohl in Bezug auf eine Rückwärtssimulation als auch gemäß konkurrierender Planungsverfahren der kapazitätsbeschränkten Rückwärtsterminierung mitunter nur geringfügig später bzw. früher in das nächstgelegene Zeitintervall ein.

Eine sich daran anknüpfende Untersuchung ermittelt auf Grundlage der Einschleusplanung stochastischer Rückwärtssimulation nach mehreren Versuchsläufen eine zusätzliche Einschleusplanung gemäß einer bestimmten relativen Größe eines Auftrages (Tab. 4).

Table 4: Ergebnisse stochastischer Rückwärtssimulation nach mehreren Versuchsläufen gemäß einer bestimmten relativen Größe eines Auftrages

	viel früher	[%]	früher	[%]	pünktlich	[%]	später	[%]	Viel später	[%]
VWS1	668	31,49	489	23,06	874	41,21	43	2,03	47	2,22
KK3	1545	72,84	233	10,99	192	9,05	46	2,17	105	4,95
KK4	1508	71,10	284	13,39	205	9,67	41	1,93	83	3,91
KK5	1448	68,27	310	14,62	246	11,60	46	2,17	71	3,35
KK6	1406	66,29	300	14,14	283	13,34	40	1,89	92	4,34
KK7	1260	59,41	389	18,34	359	16,93	33	1,56	80	3,77
KK8	1250	58,93	398	18,76	381	17,96	25	1,18	67	3,16
KK9	1194	56,29	365	17,21	427	20,13	50	2,36	85	4,01

Im Rahmen der daraus resultierenden Einschleusplanung und den Simulationsläufen *KK3* bis *KK9* werden hierfür der jeweils über die Versuchsläufe zeitlich gesehen drittfrüheste (*KK3*) bis zweitspäteste (*KK9*) Einschleustermin aus der stochastischen Rückwärtssimulation *STO1* bis *STO10* übernommen.

Die Ergebnisse einer stochastischen Rückwärtssimulation gemäß einer bestimmten relativen Größe eines Auftrages und den Simulationsläufen *KK3* bis *KK9* zeigen, dass sich die Zahl *pünktlich* fertiggestellter Aufträge auf einen Anteil zwischen rund 9,1 Prozent und 20,1 Prozent beläuft. Hinsichtlich eines Zeitfensters von $t \pm 6h$ gegenüber des angenommenen „realen“ Liefertermins werden dagegen zwischen 22,3 Prozent und 39,7 Prozent der Aufträge fertiggestellt. Die 39,7 Prozent beziehen sich hier auf den Simulationslauf, der jeweils den zweitspätesten Einschleustermin eines jeden Auftrages vorsah, während sich die 22,3 Prozent auf den Simulationslauf mit dem drittfrühesten Einschleustermin beziehen. Dies verdeutlicht noch einmal die zuvor herausgestellten Vorsprünge des methodischen Ansatzes der Rückwärtssimulation, indem die ermittelten Liefertermine noch immer genauer um die angenommenen „realen“ Liefertermine als bei den beiden herangezogenen Planungsverfahren deterministischer Rückwärtsterminierung streuen. Darüber hinaus kann auf Grundlage der Simulationsläufe *KK3* bis *KK9* festgehalten werden, dass eine spätere Einschleusplanung (für jeden Auftrag) in Summe zuverlässigere Liefertermine gegenüber der angenommenen „realen“ Liefertermine zur Folge hat.

4 Ausblick

Die auf Basis eines „real-world use-case“ erzeugten Ergebnisse zeigen, dass der methodische Ansatz zur Generierung einer Einschleusplanung durch Rückwärtssimulation unter den Spezifika der Halbleiterfertigung und unter Berücksichtigung stochastischer Einflüsse funktioniert und vielversprechende Ergebnisse liefern kann. Über mehrere Modelle konnte nun nachgewiesen werden, dass die Rückwärtssimulation als ein leistungsfähiges Werkzeug zur Einschleusplanung unter Berücksichtigung stochastischer Einflüsse dienen kann. Ein nächster Schritt im Vorhaben soll sein, den methodischen Ansatz innerhalb des betrieblichen Simulationswerkzeuges (AutoSched AP) beim Praxispartner zu erproben und zu integrieren. Zukünftig wäre es möglich das hier angewendete Verfahren mit verschiedenen Heuristiken zu einer simulationsbasierten Optimierung (rückwärtsgerichtet) weiterzuentwickeln. Anknüpfend daran soll die Anwendbarkeit des Verfahrens an sich als auch solch simulationsbasierter Optimierungskombinationen in anderen Domänen erprobt werden.

Danksagung

Der vorliegende Beitrag wurde teilweise durch das europäische Forschungsvorhaben iDEV40 finanziert, das vom ECSEL Joint Undertaking (JU) unter der Fördervereinbarung Nr. 783163 gefördert wird. Das JU erhält Unterstützung aus dem Forschungs- und Innovationsprogramm Horizon 2020 der Europäischen Union. Es wird von den Konsortiumsmitgliedern sowie Zuschüssen aus Österreich, Deutschland, Belgien, Italien, Spanien und Rumänien mitfinanziert. Die in dieser Veröffentlichung enthaltenen Informationen und Ergebnisse sind die der Autoren und spiegeln nicht zwingend die Meinung des ECSEL Joint Undertaking wider.

Literatur

- Arakawa, M.; Fuyuki, M.; Inoue, I.: A Simulation-based Production Scheduling Method for Minimizing the Due-date-deviation. *International Transactions in Operational Research* 9 (2002) 2, S. 153-167.
- Banks, J.: *Handbook of simulation: Principles, methodology, advances, applications, and practice*. New York: Wiley 1998.
- Breque, M.; De Nul, L.; Petridis, A.: *Industry 5.0 – Towards a sustainable, human-centric and resilient European Industry*. Luxembourg: Publications Office of the European Union 2021.
- Graupner, T. D.; Bornhäuser, M.; Sihn, W.: Backward simulation in food industry for facility planning and daily scheduling. In: Lipovszki, G.; Molnár, I. (Hrsg.): *Proceedings of the 16th ESS, Budapest (Ungarn)*, 17.-20.10.2004.
- Huang, C.; Wang, H.: Backward Simulation with Multiple Objectives Control. *Proceedings of the IMECS Vol. II*, 18.-20.03.2009.
- Jain, S.; Chan, S.: Experiences with Backward Simulation Based Approach for Lot Release Planning. In: Andradóttir, S.; Healy, K. J.; Withers, D. H.; Nelson, B. L. (Hrsg.): *Proceedings of the 29th WSC, Atlanta (USA)*, 7.-10.12.1997, S. 773-780.
- Kuhnle, A.; Röhrig, N.; Lanza, G.: Autonomous order dispatching in the semiconductor industry using reinforcement learning. In: *Proceedings CIRP 79* (2019), S. 391-396.
- Kuhnle, A.; Kaiser, J.-P.; Theiß, F.; Stricker, N.; Lanza, G.: Designing an adaptive production control system using reinforcement learning. In: *Journal of Intelligent Manufacturing* 32 (2021), S. 855-876.
- Laroque, C.; Löffler, C.; Scholl, W.; Schneider, G.: Einsatzmöglichkeiten der Rückwärtssimulation zur Produktionsplanung in der Halbleiterfertigung. In: Deateu, C.; Lückerrath, D.; Ullrich, O.; Durak, U.: *Proceedings ASIM SST 2020*. Wien: RGE SIM Report AR 59, S. 397-401.
- Law, A.; Kelton, D.: *Simulation Modeling and Analysis*. 2nd Edition. New York: McGraw Hill 2000.
- Scholl, W.; Laroque, C.; Weigert, G.: Evaluations on Scheduling in Semiconductor Manufacturing by Backward Simulation. In: Tolk, A.; Diallo, S. Y.; Ryzhov, I. O.; Yilmaz, L.; Buckley, S.; Miller, J. A. (Hrsg.): *Proceedings of the 2014 WSC, Savannah (USA)*, 7.-10.12.2014, S. 2552-2560.
- Schumacher, R.; Wenzel, S.: Der Modellbildungsprozess in der Simulation. In: Wenzel, S. (Hrsg.): *Referenzmodelle für die Simulation in Produktion und Logistik*. Gent: SCS-Europe BVBA 2000, S. 5-11.
- Watson, E. F.; Medeiros, D. J.; Sadowski, R. P.: Generating Component Release Plans with Backward Simulation. In: Evans, G. W.; Mollaghasemi, M.; Russel, E. C.; Biles, W. E. (Hrsg.): *Proceedings of the 25th Winter Simulation Conference (WSC), Los Angeles (USA)*, 12.-15. Dezember 1993, S. 930-938.
- Watson, E. F.; Medeiros, D. J.; Sadowski, R. P.: A simulation-based backward planning approach for order-release. In: Andradóttir, S.; Healy, K. J.; Withers, D. H.; Nelson, B. L. (Hrsg.): *Proceedings of the 29th Winter Simulation Conference (WSC), Atlanta (USA)*, 7.-10. Dezember 1997, S.765-772.
- Ying, C. C.; Clark, G. M.: Order release planning in a job shop using a bidirectional simulation algorithm. In: Tew, J. D.; Manivannan, M. S.; Sadowski, D.A. (Hrsg.): *Proceedings of the 26th WSC, Orlando (USA)*, 11.-14.12.1994, S. 1008-1012.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Scalable cooperative multi-agent-reinforcement-learning for order-controlled on schedule manufacturing in flexible manufacturing systems

Skalierbares kooperatives Multi-Agent-Reinforcement-Learning zur termintreuen, auftragsgesteuerten Fertigungssteuerung in flexiblen Fertigungssystemen

Berend Denkena, Marc-André Dittrich, Silas Fohlmeister,
Daniel Kemp, Gregory Palmer
Leibniz University Hannover, Hannover (Germany),
denkena@ifw.uni-hannover.de, dittrich@ifw.uni-hannover.de,
fohlmeister@ifw.uni-hannover.de, kemp@ifw.uni-hannover.de, gpalmer@l3s.de

Abstract: To operate flexible manufacturing systems efficiently, a robust and optimal production control is crucial. With an increasing number of workpieces being processed in parallel, ensuring guaranteed lead times represents a complex optimization tasks, better known as the *flexible scheduling problem*. Cooperative multi-agent reinforcement learning approaches have recently shown their potential in production control. However, ensuring guaranteed lead times in flexible manufacturing systems with these approaches remains an open problem. In this work, an existing cooperative multi-agent framework for flexible job-shop scheduling is transferred and modified to optimize production control in flexible manufacturing systems. Using a centralized training for decentralized execution multi-agent deep reinforcement learning approach, the goal is to optimize order agents to ensure guaranteed lead times. Furthermore, a comprehensive simulation study investigates the effect of common knowledge on facilitating cooperation, and empirically evaluate the frameworks scalability to a range of challenging scenarios.

1 Introduction

The trend towards an increasing customization of workpieces in the manufacturing industry, while ensuring guaranteed lead times, can be facilitated with flexible manufacturing systems (FMS) (ElMaraghy et al., 2013). FMS typically consist of a group of universal CNC-machine tools, which are interconnected by an automated material handling and storage system for workpieces and cutting tools (Brecher and Weck, 2019). Furthermore, FMS are automatically setup and loaded and thus are able

to attain high flexibility and high productivity without any manual interventions (Manu et al., 2018). However, with an increasing number and range of machine tools and the number of workpieces processed in parallel, a robust production control is crucial for operating a FMS efficiently (Kutin et al., 2018). The allocation of workpieces to dedicated machine tools, e. g. for ensuring granted lead times, is critical, representing a complex optimization tasks, known as the flexible scheduling problem (Sormaz and Patel, 2018).

Cooperative multi-agent systems (MAS) represent a promising approach to achieve a robust and optimized production control in fully automated FMS (Monostori et al., 2016). Their decentralized and cooperative way of solving optimization problems, enable them to cope with complex control tasks (Leitão, 2009). Due to these characteristics MAS are essential for realising an order-controlled production, in which workpieces autonomously coordinate themselves through a manufacturing environment, e.g. a FMS (Bussmann et al., 2004). To enable the intelligent control and steady optimization, deep reinforcement learning (RL) techniques were recently applied to MAS, e.g. (Waschneck et al., 2018b). A major constraint of RL is the dependence on simulation environments (Gallina et al., 2019).

Dittrich and Fohlmeister (2020) proposed a cooperative multi-agent reinforcement learning (MARL) framework, which is integrated in *Tecnomatix Plant Simulation*. Therein, intelligent order-agents are able to solve flexible job-shop scheduling problems efficiently, through learning to cooperate. The authors have shown the approach's potential for successfully reducing mean cycle times in case of production control for job shop production environments. Their approach can also be transferred to the related production control domains, which can be characterized a *flexible scheduling problem* as well. Thus, this article will modify the aforementioned framework for ensuring guaranteed lead times in FMS through expanding the so-called *field-of-view common knowledge*, i. e., the number of agents sharing information between each other. Schroeder de Witt et al. (2018) observed that a large number of cooperative multi-agent tasks benefit from considering it, as its size can increase the likelihood of agents learning to coordinate. To the best of our knowledge, to date no analysis has been performed regarding the impact of field-of-view-size on the performance of MARL agents learning to solve the flexible scheduling problem.

2 State of the Art

As mentioned in the previous section, deep RL-approaches were successfully applied to flexible scheduling problems. Similar approaches mainly differ regarding the underlying optimization goal and their optimization approach. Thus, similar optimization approaches are briefly categorized in Table 1.

Table 1: Key characteristics of similar approaches

Reference	Goal of optimization	Optimization approach
(Dittrich and Fohlmeister, 2020)	Mean cycle times	Cooperative multi-agent approach (CMAA)
(Göppert et al., 2020)	Capacity utilization	Single-agent approach

(Guo et al., 2020)	Tardiness	Single-agent approach
(Han and Yang, 2020)	Total production time	Single-agent approach
(Hofmann et al., 2020)	Run-trough time	Multi-agent approach
(Hu et al., 2020)	Avoidance of downtimes	Multi-agent approach
(Kim et al., 2020)	Tardiness	CMAA*
(Liu et al., 2020)	Total production time	Single-agent approach
(Luo, 2020)	Tardiness	Single-agent approach
(Shiue et al., 2020)	Run-through time	Single-agent approach
(Zhu et al., 2020)	Total production time	Single-agent approach
(Baer et al., 2019)	n. a.	Multi-agent approach
(Qu et al., 2019)	Work-in-Progress	Multi-agent approach
(Silva and Azevedo, 2019)	Lead times	Single-agent approach
(Silva et al., 2019)	Total production time	CMAA
(Qu et al., 2018)	Total production costs	Multi-agent approach
(Wang et al., 2018)	Run-trough time and total production costs	Multi-agent approach
(Waschneck et al., 2018a)	Capacity utilization	CMAA
(Waschneck et al., 2018b)	Run-through time and cycle times	Single-agent approach
(Bouazza et al., 2017)	Waiting time	Multi-agent approach
(Shahrabi et al., 2017)	Total production costs	Single-agent approach
(Qu, 2016)	Total production costs	CMAA
(Qu et al., 2016)	Total production costs	Multi-agent approach

* Scheduling via intelligent machine agents

It can be seen that only few proposed approaches exist, which focus on optimizing the tardiness, i. e., ensuring guaranteed lead times. In addition, most of the approaches do not make any use of common knowledge between agents. None of them were applied to tardiness-optimization in this context. In case cooperative agents were used, no author has comprehensively analysed the size of the field-of-view's effect on the quality of resulting joint-policy, i. e., the extent to which it facilitates cooperation between intelligent agents. However, as stated in Schroeder de Witt et al. (2018), an increased field-of-view size, and thereby increased common knowledge, enables complex decentralised coordination. Therefore, the field-of-view size should be considered.

3 Formulation of the optimization problem

Guarantying lead times or in other words preventing tardiness, is highly relevant in context of production control. The tardiness T is in our case defined as the percentage difference between a guaranteed lead time and the time of completion of an individual workpiece. As any tardiness T is in general unfavourable, the goal of optimization should however not be confused with just minimizing the overall tardiness. A plain minimisation would lead to increasing stock of finished parts and also to increasing

capital commitment. Instead, our goal is to reach an overall mean tardiness \bar{T} of ideally zero percent or slightly below zero percent, while controlling a dynamically generated production program.

Since the underlying flexible scheduling problems can be described as a fully-cooperative multi-agent task, it can be formulated as a decentralized partially observable Markov decision process (Dec-POMDP) (Oliehoek and Amato, 2016). The Dec-POMDP is a tuple (n, X, O, U, P, R) consisting of a state space X , an observation function (see Eq. 1), for each state $x \in X$ a joint action space $U \equiv U^n$, a transition function (see Eq. 2) returning the probability of transitioning from a state x_t to x_{t+1} given an action profile u , and a reward function (see Eq. 3) for each agent a . The agents are optimized using local and global rewards. The global reward function is shared. Finally, we allow *terminal* (absorbing) states at an episode's end.

$$O_i: X \rightarrow \mathbb{R}^d \quad (1)$$

$$P: X_t \times U \times X_{t+1} \rightarrow [0,1] \quad (2)$$

$$R_i: X_t \times U \times X_{t+1} \rightarrow R \quad (3)$$

The presented MARL-framework applies a centralized training for decentralized execution approach using deep Q-learning (DQN) (Mnih et al., 2015). Agents share a multi-layer perceptron, trained to approximate Q-Values for observation-action pairs (see Eq. 4). Network parameters θ are trained using Adam (Kingma and Ba, 2015) on the mean squared Bellman residual with the expectation taken over transitions uniformly sampled from an experience replay memory (see Eq. 5), where Y_t is the defined target (see Eq. 6). Parameters θ'_t belong to a stable *target network*, which is synchronised with the current network every n transitions (Mnih et al., 2015).

$$Q_i: O_i \times A_i \rightarrow \mathbb{R} \quad (4)$$

$$L(\theta_i) = \mathbf{E}_{o,a \sim p(\cdot)} [(Y_t - Q(o, a; \theta_t))^2] \quad (5)$$

$$Y_t \equiv r_{t+1} + \gamma \operatorname{argmax}_{\alpha \in A} Q(o_{t+1}, \alpha; \theta'_t) \quad (6)$$

4 Modifications and hyperparameter settings

The MARL framework (Fig. 1) contains a central DQN module and a decentralized cooperative MAS. Via socket-interfaces the communication to the simulation software *Tecnomatix Plant Simulation* is realised, in which an FMS can be modelled. Workpieces to be processed in the simulated FMS are represented as generic order agents a_i in the cooperative MAS, available machine tools are represented as generic machine agents m_i , respectively. Relevant machine agents $m_{i,t}$ are requested by individual order agents a_i to process the equivalent workpieces. They will propose the request with an expected timeslot for processing. In case no processing is possible, the machine agent $m_{i,t}$ will decline the request. In order to also consider other order agents $\{a_{i+1}, \dots, a_n\}$ in the decision making process, a group of cooperating agents can be requested to provide information about themselves and their current production process. An agent's field-of-view is therefore determined by the number of agents

providing information. After all relevant information is received by the order agent a_i , the information is aggregated and a decision is made through evaluating the latest available local function approximator Q_i . A more detailed description is provided in Dittrich and Fohlmeister (2020).

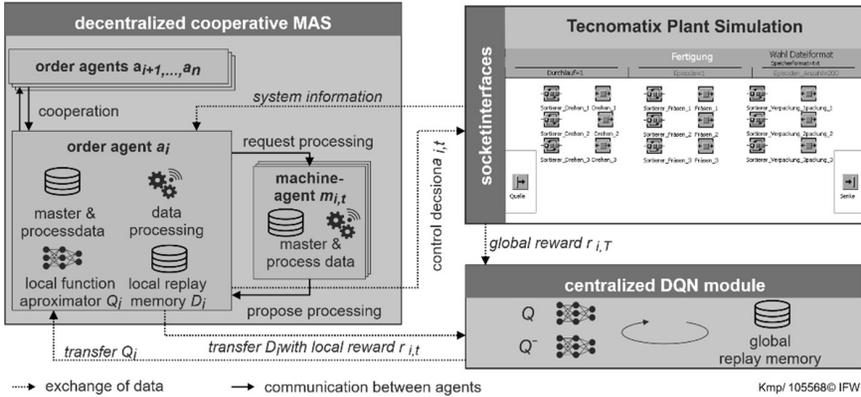


Figure 1: MARL-Framework by Dittrich and Fohlmeister (2020)

The outlined decision-making process can be fully adopted for production control of various FMS. A few modifications have to be considered for ensuring that the DQN module trains the local function approximator Q_i towards an optimal mean tardiness \bar{T} . First and foremost, the global reward function $r_{i,T}$ needs to be modified. For training the network parameters of Q_i regarding an optimal tardiness, preliminary investigations have shown that good results can be achieved with a $r_{i,T}$ according to Equation 7:

$$r_{i,T} = \begin{cases} 0 & \Delta_i < -20\% \\ 3 & -20\% \leq \Delta_i < -10\% \\ 10 & -10\% \leq \Delta_i \leq 0\% \\ -5 & 0\% < \Delta_i. \end{cases} \quad (7)$$

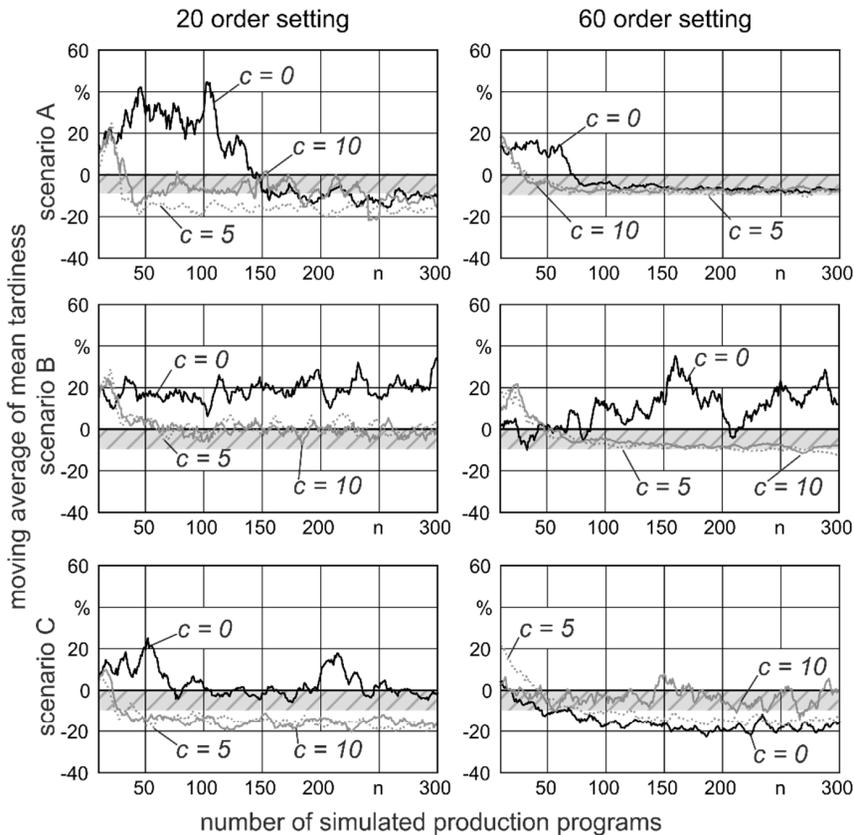
The combination of local and global rewards leads to a comparatively fast convergence towards a nearly optimal tardiness. In comparison to the fundamental framework the use of a larger batch size in the DQN module was implemented. Considering the performance of the simulation studies in section 5, the hyperparameter setting of Table 2 were evaluated as working well by the authors.

Table 2: Hyperparameter settings of the DQN-approach

Hyperparameter settings	
Warmup steps n_{warmup} :	3,000
Capacity D of global replay memory:	20,000
Exploration rate ϵ , decay ϵ_d of ϵ , ϵ_{min} :	1.2, 0.9999, 0.001
Learning rate α , discount factor γ , batch size:	0.001, 0.95, 50

5 Simulation Study

To analyse the effect of an increased common knowledge between cooperating agents, a comprehensive simulation study has been executed. To empirically evaluate the scalability of the outlined approach, a total of eighteen configurations has been composed (Fig 2). Therefore, three FMS were designed and gradually scaled. First, an FMS with three manufacturing processes (MP) and three machines per MP has been modelled (Scenario A). Second, the FMS was scaled by duplicating the number of machine tools per MP (Scenario B). The production program consisted of three, dynamically and randomly created order types.



Details of the flexible manufacturing system

	Scenario A	Scenario B	Scenario C
Number of processes p :	3	3	5
Machine tools per process m :	3	6	6
Order types in production program t :	3	3	6
Size of neural network layers:	8x8x8	20x10x6	70x35x12

Kmp/ 105576 © IFW

Figure 2: Results of the simulation study concerning the common knowledge effect

Third, the number of MP was increased to a total of five, while leaving the number of machine tools per MP constant (scenario C). The production program's complexity was scaled to six order types. All scenarios were simulated for production programs consisting of twenty as well as sixty orders. In each scenario the field-of-view's size c has been varied between zero, five or ten consecutive order agents, which were included into the common knowledge of the decision-making order agent. The framework and all outlined scenarios are provided in detail in Palmer et al. (2021).

Considering the results for scenario A, it can be seen that the MARL approach was able to successfully learn a control-policy in all simulated configurations, which at least ensured a mean tardiness \bar{T} of zero percent or below. Thus, any tardiness was successfully avoided and guaranteed lead times were satisfied. However, without any cooperating agents ($c=0$) it took significantly longer to reach the desired tardiness corridor (hatched area). Each increase of the field-of-view size resulted in an improved performance. As can be observed in the sixty-order-setting, in which the number of executed control decisions tripled compared to the twenty-order-setting, all simulated scenarios reached a similar as well as stable control-policy in the long run.

Scaling up the FMS (scenario B), i. e., increasing the underlying complexity of the *flexible scheduling problem*, lead to a significantly different performance of the $c=0$ configurations. Without any common knowledge the $c=0$ control policies did not converge successfully, meaning that the desired tardiness corridor was never reached. The other field-of-view-configurations were able to reach the desired corridor, with $c=10$ slightly outperforming $c=5$. The sixty-order-setting, lead to an even more stable performance. Yet, due to the higher complexity, these configurations were only able to reach a near optimal \bar{T} of about negative ten percent in the long run.

Considering scenario C, $c=5$ and $c=10$ again outperformed the $c=0$ configuration. In the twenty-order-setting $c=0$ was at no time able to stably avoid positive \bar{T} . For the most complex sixty-order-setting, ten verification runs were conducted. Therefore, average results are presented. It can be observed that $c=0$ reached a reasonable performance regarding its performance average. This could be explained due to no common knowledge of $c=0$, resulting in achieving a fast run-through time triggered by the use of local rewards. Despite this fact, it was not able to reach the desired tardiness corridor and was outperformed by $c=5$ and $c=10$, as well. Hence, the general observation can be derived that considering common knowledge positively affects the approach's performance.

6 Conclusion and Future Outlook

Within FMS guaranteeing lead times, while avoiding tardiness and reducing capital commitment in parallel, represents a complex optimization task. Cooperative MARL appears as a promising solution. However, the factor of common knowledge between cooperating agents is crucial. To evaluate the effect of an increased common knowledge in context of scalability and performance, a comprehensive simulation study has been executed. The study has shown that the *flexible scheduling problem* becomes significantly less manageable without any consideration of cooperating agents. The consideration leads to a faster convergence towards an optimal control policy and results in a better performance. Nevertheless, its exact size needs to be critically evaluated case-by-case, as a broader common knowledge only slightly improved the performance, while increasing the computational complexity. Future

research should therefore consider further simulation studies for determining a rule of thumb effect of common knowledge on multi-agent learning in context of production control. In addition, the presented approach only includes workpieces and machinery so far. Thus, future research should also consider tangential processes, i. e., material and tool supply, as those are of high relevance for actual applications in practice.

Acknowledgements

The authors gratefully acknowledge, that the proposed research is a result of the research project “IIP-Ecosphere”, granted by the German Federal Ministry of Economics and Technology (BMWi) via funding code 01MK20006A.

References

- Baer, S.; Bakakeu, J.; Meyes, R.; Meisen, T.: Multi-Agent Reinforcement Learning for Job Shop Scheduling in Flexible Manufacturing Systems. In: 2nd International Conference on Artificial Intelligence for Industries (AI4I), Laguna Hills, CA, USA, 2019, pp. 22–25.
- Bouazza, W.; Sallez, Y.; Beldjilali, B.: A distributed approach solving partially flexible job-shop scheduling problem with a Q-learning effect. *IFAC-PapersOnLine* 50 (2017) 1, pp. 15890–15895.
- Brecher, C.; Weck, M.: *Werkzeugmaschinen Fertigungssysteme 1*. Berlin, Heidelberg: Springer 2019.
- Bussmann, S.; Jennings, N.R.; Wooldridge, M.: *Multiagent Systems for Manufacturing Control*. Berlin, Heidelberg: Springer Berlin Heidelberg 2004.
- Dittrich, M.-A.; Fohlmeister, S.: Cooperative multi-agent system for production control using reinforcement learning. *CIRP Annals* 69 (2020) 1, pp. 389–392.
- ElMaraghy, H.; Schuh, G.; ElMaraghy, W.; Piller, F.; Schönsleben, P.; Tseng, M.; Bernard, A.: Product variety management. *CIRP Annals* 62 (2013) 2, pp. 629–652.
- Gallina, V.; Lingitz, L.; Karner, M.: A New Perspective of the Cyber-Physical Production Planning System. In: 16th IMEKO TC10 Conference, Berlin, Germany, 2019, 2019, pp. 60–65.
- Göppert, A.; Rachner, J.; Schmitt, R.H.: Automated scenario analysis of reinforcement learning controlled line-less assembly systems. *Procedia CIRP* 93 (2020), pp. 1091–1096.
- Guo, L.; Zhuang, Z.; Huang, Z.; Qin, W.: optimization of dynamic multi-objective non-identical parallel machine scheduling with multi-stage reinforcement learning. In: *IEEE 16th International Conference on Automation Science and Engineering (CASE)*, Hong Kong, Hong Kong, 2020, pp. 1215–1219.
- Han, B.-A.; Yang, J.-J.: Research on Adaptive Job Shop Scheduling Problems Based on Dueling Double DQN. *IEEE Access* 8 (2020), pp. 186474–186495.
- Hofmann, C.; Krahe, C.; Stricker, N.; Lanza, G.: Autonomous production control for matrix production based on deep Q-learning. *Procedia CIRP* 88 (2020), pp. 25–30.

- Hu, L.; Liu, Z.; Hu, W.; Wang, Y.; Tan, J.; Wu, F.: Petri-net-based dynamic scheduling of flexible manufacturing system via deep reinforcement learning with graph convolutional network. *Journal of Manufacturing Systems* 55 (2020), pp. 1–14.
- Kim, Y.G.; Lee, S.; Son, J.; Bae, H.; Chung, B.D.: Multi-agent system and reinforcement learning approach for distributed intelligence in a flexible smart manufacturing system. *Journal of Manufacturing Systems* 57 (2020), pp. 440–450.
- Kingma, D.P.; Ba, J.: Adam: A Method for Stochastic Optimization. In: 3rd International Conference for Learning Representations, San Diego, CA, USA, 2015, pp. 1–15.
- Kutin, A.A.; Dolgov, V.A.; Kabanov, A.A.; Dazuk, I.V.; Podkidyshev, A.A.: Improving the efficiency of CNC machine tools with multi-pallet systems in machine-building manufacturing. *IOP Conference Series: Materials Science and Engineering* 448 (2018), pp. 12010.
- Leitão, P.: Agent-based distributed manufacturing control: A state-of-the-art survey. *Engineering Applications of Artificial Intelligence* 22 (2009) 7, pp. 979–991.
- Liu, C.-L.; Chang, C.-C.; Tseng, C.-J.: Actor-Critic Deep Reinforcement Learning for Solving Job Shop Scheduling Problems. *IEEE Access* 8 (2020), pp. 71752–71762.
- Luo, S.: Dynamic scheduling for flexible job shop with new job insertions by deep reinforcement learning. *Applied Soft Computing* 91 (2020), pp. 106208.
- Manu, G.; Vijay Kumar, M.; Nagesh, H.; Jagadeesh, D.; Gowtham, M.B.: Flexible Manufacturing Systems (FMS): A Review. *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)* 8 (2018) 2, pp. 323–336.
- Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A.A.; Veness, J.; Bellemare, M.G.; Graves, A.; Riedmiller, M.; Fidjeland, A.K.; Ostrovski, G.; Petersen, S.; Beattie, C.; Sadik, A.; Antonoglou, I.; King, H.; Kumaran, D.; Wierstra, D.; Legg, S.; Hassabis, D.: Human-level control through deep reinforcement learning. *Nature* 518 (2015) 7540, pp. 529–533.
- Monostori, L.; Kádár, B.; Bauernhansl, T.; Kondoh, S.; Kumara, S.; Reinhart, G.; Sauer, O.; Schuh, G.; Sihn, W.; Ueda, K.: Cyber-physical systems in manufacturing. *CIRP Annals* 65 (2016) 2, pp. 621–641.
- Oliehoek, F.A.; Amato, C.: *A Concise Introduction to Decentralized POMDPs*. Cham: Springer International Publishing 2016.
- Palmer, G.; Kemp, D.; Fohlmeister, S., 2021: Scalable cooperative Multi-Agent-Reinforcement-Learning for order-controlled on schedule manufacturing in flexible manufacturing systems. https://github.com/gjp1203/fms_marl, accessed May 14th, 2021.
- Qu, S.: Learning Adaptive Dispatching Rules for a Manufacturing Process System by Using Reinforcement Learning Approach. In: 21st International Conference on Emerging Technologies and Factory Automation (EFTA), Berlin, Germany, 2016, pp. 1–8.

- Qu, S.; Wang, J.; Govil, S.; Leckie, J.O.: Optimized Adaptive Scheduling of a Manufacturing Process System with Multi-skill Workforce and Multiple Machine Types: An Ontology-based, Multi-agent Reinforcement Learning Approach. *Procedia CIRP* 57 (2016), pp. 55–60.
- Qu, S.; Wang, J.; Jasperneite, J.: Dynamic scheduling in large-scale stochastic processing networks for demand-driven manufacturing using distributed reinforcement learning. In: *IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA)*, Turin, Italy, 2018, pp. 433–440.
- Qu, S.; Wang, J.; Jasperneite, J.: Dynamic scheduling in modern processing systems using expert-guided distributed reinforcement learning. In: *IEEE 24th International Conference on Emerging Technologies and Factory Automation (ETFA)*, Zaragoza, Spain, 2019, pp. 459–466.
- Schroeder de Witt, C.A.; Foerster, J.N.; Farquhar, G.; Torr, P.H.; Boehmer, W.; Whiteson, S.: Multi-Agent Common Knowledge Reinforcement Learning. In: *32nd Conference on Neural Information Processing Systems (NeurIPS)*, Montréal, Canada, 2018, pp. 1–17.
- Shahrabi, J.; Adibi, M.A.; Mahootchi, M.: A reinforcement learning approach to parameter estimation in dynamic job shop scheduling. *Computers & Industrial Engineering* 110 (2017), pp. 75–82.
- Shiue, Y.-R.; Lee, K.-C.; Su, C.-T.: A Reinforcement Learning Approach to Dynamic Scheduling in a Product-Mix Flexibility Environment. *IEEE Access* 8 (2020), pp. 106542–106553.
- Silva, M.A.; Souza, S.R. de; Freitas Souza, M.J.; Bazzan, A.L.: A reinforcement learning-based multi-agent framework applied for solving routing and scheduling problems. *Expert Systems with Applications* 131 (2019), pp. 148–171.
- Silva, T.; Azevedo, A.: Production flow control through the use of reinforcement learning. *Procedia Manufacturing* 38 (2019), pp. 194–202.
- Sormaz, D.; Patel, C.: Development and evaluation of feature-focused dynamic routing policy. *The International Journal of Advanced Manufacturing Technology* 99 (2018) 1-4, pp. 15–28.
- Wang, M.; Chen, X.; Zhou, J.; Jiang, T.; Cai, W.: Shared Cognition Based Integration Dynamic Scheduling Method. In: *IEEE 2nd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, Xi'an, China, 2018, pp. 1438–1442.
- Waschneck, B.; Reichstaller, A.; Belzner, L.; Altenmüller, T.; Bauernhansl, T.; Knapp, A.; Kyek, A.: Deep reinforcement learning for semiconductor production scheduling. In: *29th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, Saratoga Springs, NY, USA, 2018a, pp. 301–306.
- Waschneck, B.; Reichstaller, A.; Belzner, L.; Altenmüller, T.; Bauernhansl, T.; Knapp, A.; Kyek, A.: Optimization of global production scheduling with deep reinforcement learning. *Procedia CIRP* 72 (2018b), pp. 1264–1269.
- Zhu, J.; Wang, H.; Zhang, T.: A Deep Reinforcement Learning Approach to the Flexible Flowshop Scheduling Problem with Makespan Minimization. In: *IEEE 9th Data Driven Control and Learning Systems Conference (DDCLS)*, Liuzhou, China, 2020, pp. 1220–1225.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

A review on simulation metamodeling for decision support systems using artificial neural networks

Ein Überblick über die Simulationsmetamodellierung für Entscheidungssysteme unter Verwendung künstlicher neuronaler Netze

Harold Billiet, Rainer Stark, Technische Universität Berlin, Berlin (Germany),
harold.billiet@tu-berlin.de, rainer.stark@tu-berlin.de

Abstract: Flexible and agile manufacturing systems can benefit from computer-aided tools that improve and speed-up the decision-making process. The combined use of discrete event simulation together with decision support systems enables manufacturers to test different control alternatives, scheduling and sequencing strategies for a specific planning horizon. Because multiple runs are typically necessary for this, computational effort of simulation models is a key limiting factor that contrasts with the need for quick planning decisions. A popular method for reducing that computational time is the use of Artificial Neural Networks (ANN) metamodels that approximate the simulation in a cheaper-to-compute functional model. This paper aims to review recent cases of simulation metamodeling through ANN for manufacturing and logistics systems in order to understand the current state of the art in simulation metamodeling. A theoretical framework of a possible solution for two identified problems will also be presented.

1 Introduction

Recent years have shown the emergence of new challenges for manufacturers; they need to improve their flexibility, efficiency, and have to adapt to an increasing demand for customizable products (Dalmarco et al., 2019). This is made possible with the use of computer-aided tools and methods. Recognized as a key aspect of the digital factory, Discrete Event Simulation (DES) is widely used as an analysis tool for manufacturing or logistics systems (Bangsow, 2020). It enables a complex analysis of the dynamics of a manufacturing system and the evaluation of different alternatives in order to adopt an appropriated strategy. Mainly used during the planning phase, companies use DES in order to optimize a manufacturing system before transferring these changes reality.

With the emergence of Industrie 4.0, information and communication technologies are enabling companies to manipulate large amounts of real-time data. Thus, DES can

be used during production in order to evaluate new scenarios and changes in the production system in real time (Heilala et al.). These simulation models are often integrated into more or less complex Decision Support Systems (DSS), in order to assist decision makers into taking the best possible decision in a complex environment (Kasie et al., 2017).

Even if recent years have shown a continuous expansion of computational power, DES is still computationally expensive. For users of simulation-based DSS, it is important to reduce this computational time in order to be able to quickly analyse many different decision-making scenarios. For Sobottka et al. (2019), computational effort is a key limitation of simulation-based optimization that contrasts with the need for quick planning decisions. To address this problem, a cheaper-to-compute metamodel of the simulation model can be created. The most promising technique for is the use of Artificial Neural Networks (ANN) (Sobottka et al., 2019).

The goal of this paper is to review cases of DES metamodeling using ANN in literature. Simulation engineers often do not have a machine learning background and could be interested in learning how different metamodels were created in order to reduce computational time and how well they performed.

The first part will address DES in the manufacturing field. The second part will focus on DSS for manufacturing and their limitations. After a small introduction about ANN, a review of current literature about the usage of ANN for the metamodeling of DES will be presented. The sixth part will present current problems for simulation metamodeling and a possible solution approach to them. The last part will focus on the conclusions of the review and how further research can be needed.

2 Discrete Event Simulation for Manufacturing

DES is a simulation method used in order to represent a manufacturing system using a distinct sequence of state changes occurring in time (Omogbai and Salonitis, 2016). Because of its flexibility and simplicity in understanding the dynamics and behaviour of manufacturing systems, DES is one of the most commonly used simulation techniques today (Negahban and Smith, 2014).

The traditional goal of DES is to model and analyse a manufacturing system during its planning phase. Performance, bottlenecks, utilization rate and standby times, can be analysed for different control and planning strategies (Bangsow, 2020).

DES can also be used during the operational phase throughout production. In order to do this, manufacturers need to synchronize real-time data from the manufacturing system with the simulation. In literature, they are referred to as adaptive simulation models (Denkena et al., 2017). Heiko et al. (2008) defined the association of a simulation system and a physical system as a symbiotic simulation system: the simulation model benefits from the physical model by obtaining the necessary data in order to initialize itself and the physical system benefits from the optional control feedback from the simulation. If failures or plan deviations occur, the DES model can be used in order to predict the impact on production and assist decision makers in taking tactical and operational decisions (Heilala et al., 2010; Jahangirian et al., 2010).

3 Decision Support Systems for manufacturing

DSS are interactive computer-based systems that enhance the ability of decision-makers to solve problems (Sprague and Carlson, 1982). They are intended to improve and speed-up the decision making process (Power, 2001). It is possible to classify DSS in five categories: data-driven, communication-driven, document-driven, knowledge-driven and model-driven DSS. Power (2001) provides more information on this topic. Model-driven DSS are widespread in the manufacturing field. They facilitate access and manipulation of a model in order to analyse different decision scenarios and discover the most favourable alternatives under a given situation (Felsberger et al., 2019). The model used in a model-driven DSS can be analytical or a simulation model. In the case of decision making for manufacturing systems, DES is commonly used because of its ability to represent complex and dynamic systems on a global level in a visually interactive model.

Heiko et al. (2008) differentiated decision support systems from decision control systems. For them, a DSS only supports an external decision maker by helping him in taking a decision whereas a decision control system can implement decisions directly. In literature, they can also be referred to as passive and active DSS (Felsberger et al., 2019).

Kunath and Winkler, 2018 defined different use cases for DSS and manufacturing systems. The first use case is dynamic scheduling. Using operational data from the manufacturing system, the DSS can test different control alternatives in order to create scheduling and sequencing strategies for a specific planning horizon. It can also be possible to integrate suppliers as external resources into the dynamic scheduling system. The second use case concerns the offer-making process. For a make-to-order production system, a DSS can be used in order to predict delivery dates and prices by using the configuration of the product and the current state of the manufacturing system as input. This can also be done if the customer wants to change an existing order. In the future, an active DSS could be integrated into the product configurator and automatically suggest prices according to available delivery dates to the customer.

Manufacturers that benefit the most from simulation-based DSS have flexible and highly variable manufacturing systems with dynamic situations and unforeseen events where it is necessary to respond quickly in an appropriate way. (Prajapat et al., 2020)

DSS capable of suggesting optimal scheduling decisions typically require multiple simulation runs in order to find the most appropriate decision. Because DES is a computationally expensive tool, calculation time can often be a limiting factor (van Gelder et al., 2014; Dunke and Nickel, 2020; Kasie et al., 2017). In order to reduce computational time from simulation models, metamodels, also known as surrogate models, can be considered. In this case, a cheaper-to-compute functional model approximates the unknown input-output function of the time-consuming simulation model (van Gelder et al., 2014). Out of all the different metamodeling techniques, the use of Artificial Neural Networks (ANN) has been found to be the most appropriate and popular method (Fonseca et al., 2003; van Gelder et al., 2014; Negahban and Smith, 2014; Sobottka et al., 2019).

4 Artificial Neural Networks

ANN are known as universal function approximators. They are made of different layers of interconnected units called neurons. Cybenko (1989) and Hornik et al. (1989) mathematically proved three decades ago that a neural network with one hidden layer could approximate any continuous function to a reasonable accuracy. Figure 1 presents an example of an ANN with 3 input neurons, 2 hidden layers (5/3 neurons) and 1 output neuron. It approximates an unknown function with 3 inputs and 1 output.

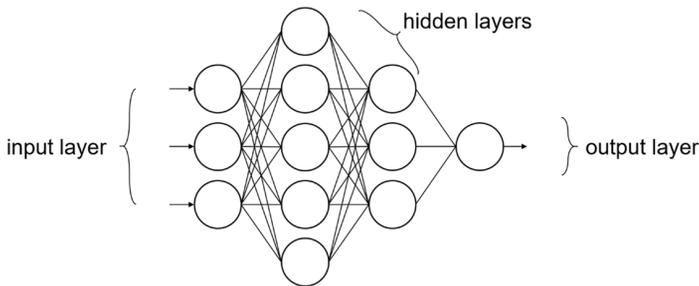


Figure 1: Example of a feed forward Artificial Neural Network

ANN use hyperparameter in order to describe their architecture and behaviour. A hyperparameter can, for example, be a number of hidden layers or a type of activation function. For more understanding about ANN in general and about their architecture, see Nielsen (2015). There are different types of ANN; for the metamodeling of DES models, the most frequently used type of neural network is the feedforward ANN, sometimes referred to as a multilayer perceptron, which is typically used in classic regression problems with tabular data.

After training an ANN, it is possible to measure its accuracy using the coefficient of correlation R . It measures the correlation between output and target values on a scale from 0 to 1 (Nuñez-Piña et al., 2018). It is also possible to use the Mean Absolute Percentage Error (MAPE), which calculates the accuracy of the predictions.

5 Review

A literature review was carried out in order to identify different cases of Simulation metamodeling using ANN. The data collection process was conducted on April 8, 2021, by searching *neural networks "discrete event simulation" manufacturing metamodel* in Google Scholar. Only papers published after 2015 and results from the first 10 pages were shown, except for Senties et al. (2009) and Azadeh et al. (2013). After filtering the results by “relevance”, the authors proceeded to select accessible papers presenting interesting manufacturing or logistics use cases of simulation metamodeling through ANN with enough information in the employed method and results. Every reviewed paper concern DSS capable of finding optimal scheduling parameters for a specific planning horizon: analysing the Buffer Allocation Problem (BAP), finding optimal Priority Dispatching Rules (PDR) and generating an Optimal Production Plan (OPP).

Table 1 gives a quick overview of the reviewed papers, their objective and results.

Table 1: *Reviewed papers*

Authors	Objective	ANN Results
Kłos and Patalas-Maliszewska (2019)	BAP	93% accuracy
Nuñez-Piña et al. (2018)	BAP	R=0,99
Azadeh et al. (2013)	PDR	95% accuracy
Azadeh et al. (2015)	PDR	Satisfying
Xanthopoulos and Koulouriotis (2018)	PDR	R=0,99
Dunke and Nickel (2020)	PDR	95% accuracy
Senties et al. (2009)	OPP	Satisfying
Sobottka et al. (2019)	OPP	R=0,90
Jackson et al. (2019)	OPP	Limited

5.1 Analysing the Buffer Allocation Problem

Kłos and Patalas-Maliszewska (2019) studied the BAP and presented a metamodel of a DES model created in Plant Simulation. Its goal was to study the impact of buffer capacity on the average throughput of a manufacturing system. In order to train an ANN, 100 random permutations of buffer allocations were created using the simulation. The created ANN had two hidden layers (9/3 neurons) and a hyperbolic tangent activation function. The accuracy of the metamodel was around 93%.

An ANN for studying the BAP was also created by Nuñez-Piña et al. (2018), using 360 training samples. The authors compared 780 different ANN architectures and found that the best performing one (R=0,9996) had 4 hidden layers (8/8/10/10 neurons) and a hyperbolic tangent activation function.

5.2 Finding Optimal Priority Dispatching Rules

Azadeh et al. (2013) presented a DSS for the dynamic scheduling of a job shop. Their objective was to find the optimal PDR for each machine. An ANN was trained using data from a DES model. The authors used the following methodology: N dissimilar training samples are created using the DES model. If the desired error percentage after training is not acceptable, the level of N is increased by 10% and the training starts again. An ANN with 3 hidden layers and a log-sigmoid transfer function was trained. By choosing an acceptable MAPE of 5%, it was necessary to generate 200 training samples.

Azadeh et al. (2015) presented another DSS for dynamic scheduling, this time for a production system for different types of canned fruits. The goal was also to calculate every permutation of PDR. A DES model was created in AweSim and generated 100 random permutations, used as training samples for an ANN. For every new planning horizon, every permutation of PDR was calculated through the ANN in order to find the most efficient ones (total of permutations: 2401). The results seemed satisfying for this particular use case.

Xanthopoulos and Koulouriotis (2018) developed an ANN in order to dynamically choose PDR in a flexible manufacturing line. A DES model was created in order to generate 192 training samples. The ANN was made of one hidden layer of 8 neurons and a hyperbolic tangent transfer function. The coefficient of correlation R was 0,99, indication a strong linear relationship between output and target values.

Dunke and Nickel (2020) developed an ANN for dynamically choosing control strategies in an order picking system. Its goal was to determine in real time which order consolidation and picker routing strategies were the most appropriate for the current state of the system. A DES model created using AnyLogic was used to train the ANN. The authors tested different ANN hyperparameter and the best results were obtained with 2 hidden layers (14/10 neurons) and a hyperbolic tangent transfer function. The accuracy of the metamodel was around 95%. The authors suggest using an ANN for a quick first impression on what the outputs might be.

5.3 Generating an Optimal Production Plan

Senties et al. (2009) developed an ANN for the scheduling of a semiconductor wafer fabrication system. Its purpose was to generate planning decisions for the next planning horizon. A DES model of the system was created using MELISSA and used as input, decision variables, that can be optimized and order specific data concerning the next planning horizon. Out of the 2400 generated training samples, 66% were used for training the ANN and 33% for its testing. The authors used an ANN with one hidden layer of 15 neurons and a hyperbolic tangent activation function. The ANN was able to generate planning decisions 100 times faster than the DES model and approximated the simulation with a high accuracy.

Sobottka et al. (2019) studied the feasibility of creating an ANN for the dynamic scheduling of an industrial bakery. The objective was to create an optimal production plan for the next planning horizon. The manufacturing system was modelled using a DES software in order to create training samples for the ANN. If 9 variables needed to be optimized, the authors found that 5000 training samples where enough for a coefficient of correlation R of 0,9. If the number of variables increased to 25, 2 million samples would be necessary.

Jackson et al. (2019) presented a solution for an inventory control problem. Because of the scale and dimensionality of the inventory control problem, optimizing specific decision variables in the simulation model would take too much time. For that reason, the authors trained an ANN using training samples from a DES model. An ANN was created using a rectified linear unit activation function, 3 hidden layers (30/30/10 neurons) and was trained with 5 datasets of 1000 samples. Stochastic noise reduced the accuracy of the model. The authors suggested that increasing the number of simulations runs and changing the architecture of the ANN could improve its accuracy.

6 Solution Approach

The literature review presented in this paper was able to identify problems in the current metamodeling process. A possible theoretical framework for a solution addressing these problems will be presented.

6.1 Research Gaps

The first problem is the technical knowledge needed for the implementation of a simulation metamodel. Simulation engineers are often domain experts, not necessarily computer scientists or programmers. They have deep knowledge of the domain-specific notations used in their domain (Lara et al., 2014). Creating, training, testing and validating an ANN requires complex machine learning skills, which simulation engineers often do not have. Thus, the lack of available qualified resources can be an obstacle for the implementation of such a metamodel.

The second problem concerns the inexistence of a universal methodology for the creation and training of an ANN. The review presented in this paper showed that researchers used almost all the time different ANN hyperparameters: number of hidden layers, number of neurons per hidden layer, activation function and amount of training samples. In some cases, researchers tried different ANN configurations and selected the best ones, see (Nuñez-Piña et al., 2018). Because an ANN is essentially a black-box, it can be difficult to understand and improve the efficiency of the model. Reducing that complexity could make simulation metamodeling more accessible while reducing its average error.

6.2 Framework

A theoretical framework of a possible solution is presented in Figure 2. The goal is to automatically be able to approximate a DES model into a cheaper to compute ANN model. The whole metamodeling process should be automated in a way that no machine learning skills are required from the user. In this case, a specific plugin could be implemented into an existing simulation software in order to simplify the process. This implies that a DES model has already been created. The verification and validation process should be done using the DES model before starting the metamodeling.

The proposed framework enables any simulation user to configure and start an automatic neural network architecture search for the metamodeling of a validated simulation model. The different inputs and outputs of the simulation would first need to be highlighted by the user so that the algorithm can automatically create the necessary training samples for the ANN. Since most of the existing DES software do not include the tools necessary to create a neural network, an external ANN-capable platform will have to be used. An algorithm integrated into the platform will use the generated training samples from the simulation in order to train different ANN models using specific hyperparameters.

When training a metamodel, a small amount of training samples will only be used for its validation. These samples have never been seen by the metamodel. Accuracy can be calculated by analysing the difference between the validation data and the predicted values from the metamodel. If the desired accuracy is achieved, the metamodel is validated.

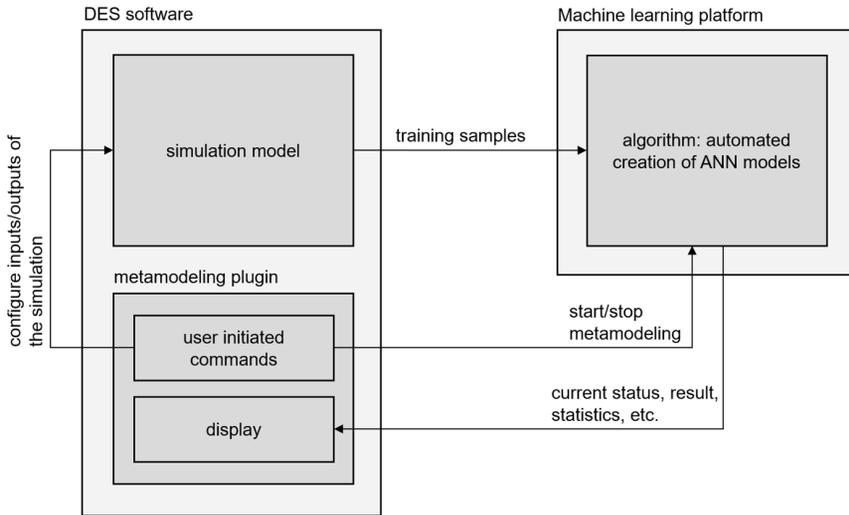


Figure 2: Proposed framework for automating the metamodeling process

6.3 Next Steps

The next research steps for the authors are to implement the theoretical framework in an existing simulation software in order to investigate its feasibility and efficiency. In order to search the best performing ANN architecture, the algorithm will have to train many different ANN. This could lead to high computational time. The algorithm should also be tested with different types of manufacturing systems and different levels of simulation complexity.

7 Conclusion

This paper presented a review of recent literature concerning the metamodeling of DES using ANN for the reduction of computational time. The use cases, methods, metamodeling parameters and results of the reviewed papers have been specified in this paper in order to have a better overview of the current state of the art in simulation metamodeling. The use of Google Scholar for selecting the papers can be criticised, because of the unclear prioritization of the large number of results. This could potentially lead to relevant papers being overlooked. Using a Database like Scopus or the archives of Simulation conferences like the Winter Simulation Conference or the ASIM Dedicated Conference could lead to fewer, but more relevant results.

The review was able to highlight the absence of a clear method for the metamodeling of simulation models and emphasized the complexity of the process. In order to simplify the whole metamodeling process, a theoretical framework was presented. The idea would be to implement an algorithm inside an existing DES software, that would automatically train an optimal ANN for the metamodeling of a specific simulation model inside the DES software. This theoretical framework will need to be implemented, tested and validated using different use cases and levels of complexity.

References

- Azadeh, A.; Maleki-Shoja, B.; Sheikhalishahi, M.; Esmaili, A.; Ziaefar, A.; Moradi, B.: A simulation optimization approach for flow-shop scheduling problem: a canned fruit industry. *The International Journal of Advanced Manufacturing Technology* 77 (2015) 1-4, pp. 751–761.
- Azadeh, A.; Shoja, B.M.; Moghaddam, M.; Asadzadeh, S.M.; Akbari, A.: A neural network meta-model for identification of optimal combination of priority dispatching rules and makespan in a deterministic job shop scheduling problem. *The International Journal of Advanced Manufacturing Technology* 67 (2013) 5-8, pp. 1549–1561.
- Bangso, S.: *Tecnomatix Plant Simulation*. Cham: Springer International Publishing 2020.
- Cybenko, G.: Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems* 2 (1989) 4, pp. 303–314.
- Dalmarco, G.; Ramalho, F.R.; Barros, A.C.; Soares, A.L.: Providing industry 4.0 technologies: The case of a production technology cluster. *The Journal of High Technology Management Research* 30 (2019) 2, pp. 100355.
- Denkena, B.; Wilmsmeier, S.; Hauck, S.: Adaptive Simulationsmodelle - Integration von Maschinendaten in Materialflusssimulationen. *Productivity, Jahrgang 22 (2017) (2017), Heft 3, S. 37 - 39.*
- Dunke, F.; Nickel, S.: Neural networks for the metamodeling of simulation models with online decision making. *Simulation Modelling Practice and Theory* 99 (2020), pp. 102016.
- Felsberger, A.; Oberegger, B.; Reiner, G.: A Review of Decision Support Systems for Manufacturing. *Saml40 workshop at i-KNOW '16 October 18–19, 2016, Graz, Austria (2019).*
- Fonseca, D.J.; Navarrese, D.O.; Moynihan, G.P.: Simulation metamodeling through artificial neural networks. *Engineering Applications of Artificial Intelligence* 16 (2003) 3, pp. 177–183.
- Heiko, A.; Stephen John, T.; Wentong, C.; Malcolm Yoke Hean Low: Symbiotic Simulation Systems: An Extended Definition Motivated by Symbiosis in Biology. *2008 22nd Workshop on Principles 03.06.2008 - 06.06.2008 (2008), pp. 109–116.*
- Heilala, J.; Montonen, J.; Jarvinen, P.; Kivikunnas, S.; Maantila, M.; Sillanpaa, J.; Jokinen, T.: Developing simulation-based Decision Support Systems for customer-driven manufacturing operation planning. *Proceedings of the 2010 Winter Simulation Conference 2010*, pp. 3363–3375.
- Hornik, K.; Stinchcombe, M.; White, H.: Multilayer feedforward networks are universal approximators. *Neural Networks* 2 (1989) 5, pp. 359–366.
- Jackson, I.; Tolujevs, J.; Lang, S.; Kegenbekov, Z.: Metamodelling of Inventory-Control Simulations Based on a Multilayer Perceptron. *Transport and Telecommunication Journal* 20 (2019) 3, pp. 251–259.

- Jahangirian, M.; Eldabi, T.; Naseer, A.; Stergioulas, L.K.; Young, T.: Simulation in manufacturing and business: A review. *European Journal of Operational Research* 203 (2010) 1, pp. 1–13.
- Kasie, F.M.; Bright, G.; Walker, A.: Decision support systems in manufacturing: a survey and future trends. *Journal of Modelling in Management* 12 (2017) 3, pp. 432–454.
- Kłos, S.; Patalas-Maliszewska, J.: An Analysis of Simulation Models in a Discrete Manufacturing System Using Artificial Neural Network (2019).
- Kunath, M.; Winkler, H.: Integrating the Digital Twin of the manufacturing system into a decision support system for improving the order management process. *Procedia CIRP* 72 (2018), pp. 225–231.
- Lara, J. de; Guerra, E.; Boronat, A.; Heckel, R.; Torrini, P.: Domain-specific discrete event modelling and simulation using graph transformation. *Software & Systems Modeling* 13 (2014) 1, pp. 209–238.
- Negahban, A.; Smith, J.S.: Simulation for manufacturing system design and operation: Literature review and analysis. *Journal of Manufacturing Systems* 33 (2014) 2, pp. 241–261.
- Nielsen, M.A.: *Neural Networks and Deep Learning*. Determination Press (2015).
- Nuñez-Piña, F.; Medina-Marin, J.; Seck-Tuoh-Mora, J.C.; Hernandez-Romero, N.; Hernandez-Gress, E.S.: Modeling of Throughput in Production Lines Using Response Surface Methodology and Artificial Neural Networks. *Complexity* 2018 (2018), pp. 1–10.
- Omogbai, O.; Saloniitis, K.: Manufacturing System Lean Improvement Design Using Discrete Event Simulation. *Procedia CIRP* 57 (2016), pp. 195–200.
- Power, D.J.: *Supporting Decision-Makers: An Expanded Framework* (2001).
- Prajapat, N.; Turner, C.; Tiwari, A.; Tiwari, D.; Hutabarat, W.: Real-time discrete event simulation: a framework for an intelligent expert system approach utilising decision trees. *The International Journal of Advanced Manufacturing Technology* 110 (2020) 11-12, pp. 2893–2911.
- Senties, O.B.; Azzaro-Pantel, C.; Pibouleau, L.; Domenech, S.: A Neural Network and a Genetic Algorithm for Multiobjective Scheduling of Semiconductor Manufacturing Plants. *Industrial & Engineering Chemistry Research* 48 (2009) 21, pp. 9546–9555.
- Sobotka, T.; Kamhuber, F.; Faezirad, M.; Sihm, W.: Potential for Machine Learning in Optimized Production Planning with Hybrid Simulation. *Procedia Manufacturing* 39 (2019), pp. 1844–1853.
- Sprague, R.H.; Carlson, E.D.: *Building effective decision support systems*. Englewood Cliffs, N.J.: Prentice-Hall 1982.
- van Gelder, L.; Das, P.; Janssen, H.; Roels, S.: Comparative study of metamodelling techniques in building energy simulation: Guidelines for practitioners. *Simulation Modelling Practice and Theory* 49 (2014), pp. 245–257.
- Xanthopoulos, A.S.; Koulouriotis, D.E.: Cluster analysis and neural network-based metamodelling of priority rules for dynamic sequencing. *Journal of Intelligent Manufacturing* 29 (2018) 1, pp. 69–91.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Towards Production-Ready Reinforcement Learning Scheduling Agents: A Hybrid Two-Step Training Approach Based on Discrete-Event Simulations

Beitrag zu einem produktionsreifen Einsatz von Reinforcement Learning-Planungsagenten: Ein hybrider zweistufiger Trainingsansatz basierend auf ereignisdiskreten Simulationen

Marco Kemmerling, Vladimir Samsonov, Daniel Lütticke, IMA, RWTH Aachen University, Aachen (Germany), marco.kemmerling@ima.rwth-aachen.de, vladimir.samsonov@ima.rwth-aachen.de, daniel.luetticke@ima.rwth-aachen.de

Günther Schuh, Andreas Gützlaff, Matthias Schmidhuber, Tim Janke, Werkzeugmaschinenlabor WZL der RWTH Aachen, Aachen (Germany), g.schuh@wzl.rwth-aachen.de, a.guetzlaff@wzl.rwth-aachen.de, m.schmidhuber@wzl.rwth-aachen.de, t.janke@wzl.rwth-aachen.de

Tobias Meisen, TMDT, University of Wuppertal, Wuppertal (Germany), meisen@uni-wuppertal.de

Abstract: Reinforcement Learning (RL) applications for the tasks of Production Planning and Control (PPC) are gaining the broad attention of the research community. Most of the current research results concentrate on PPC problems such as order release, sequence building, or transportation planning and rely on custom-made production simulations for RL. While this is well suited for conceptual research, it is currently inconceivable in practice due to the complexity of real production environments and the trend towards standardized software applications. Since interfacing with standardized simulations to train RL agents limits the speed at which training can occur, we propose a hybrid RL approach involving pre-training on a fast and flexible python-native production simulation with subsequent training and evaluation on the industrial-grade commercial solution. To enable pre-training in similar conditions as encountered in practice, we further demonstrate how to generate realistic production scenarios using real production data.

1 Introduction

Manufacturing companies are reacting to the increasing demand for customized products by designing their production to be more flexible and by moving towards

job shop production structures with individual workstations and undirected material flows (Schuh et al., 2019; Zijm and Regattieri, 2019). This flexibility comes at the cost of significantly increased complexity in Production Planning and Control (PPC) (Schlegel et al., 2019). In order to fulfil the general production targets and keep logistical costs low, advanced order release and sequencing methods are required. Multiple exact methods, meta-heuristics and dispatching rules already exist to address these problems (Xie et al., 2019). However, finding a complete scheduling solution upfront, which may then become obsolete shortly afterwards, unnecessarily consumes computational time. For this reason, continuous online scheduling, where only parts of a solution are computed as they are required based on information of the system's current state, might be preferable in practice (Lang et al., 2019). However, there is no universal method suitable for online scheduling capable of finding close-to-optimal solutions in a limited time.

Multiple recent studies investigate the applicability of Reinforcement Learning (RL) for order release control as well as sequence building in production (Kuhnle et al., 2020; Lang et al., 2020; Waschneck et al., 2018) and demonstrate that the application of RL can yield competitive solutions in this domain. Especially job shop production systems – further gaining in importance for customized industrial goods – are difficult to optimize with present methods. Training of RL agents requires trial-and-error and is thus not advisable within a real production system, because an untrained agent will perform many deficient decisions. To avoid unnecessary costs such agents are typically trained on simulation environments.

Once trained, an agent can be transferred from the simulation to the real counterpart. These simulations need to closely depict the complexity of a real production system while being performant, scalable and offering the ability to interface with RL agents at the same time. Industrial-grade commercial solutions such as Automode, AnyLogic or Plant Simulation fulfill the complexity requirements, but lack defined communication interfaces with RL agents and possibilities in parallelization and thus limit the feasibility of fast training. Currently, as far as the literature research has shown, no research is available on how to combine the flexibility and parallelization of custom-made simulation solutions with the error-proofed realism of commercial solutions. In this study, we investigate the aspect of combining an RL agent, fast python-native discrete-event simulations and further detailed commercial simulation models (using Plant Simulation) for order release tasks in a job shop production, which offers a novel strategy to mitigate current shortcomings and brings RL-based solutions one step closer to deployment.

2 State-of-the-Art

Considerable uncertainty levels and constantly changing conditions in production environments require adaptive online planning. Simulation methods offer high flexibility in dealing with uncertainty and finding planning solutions that satisfy the production goals at hand. However, the number of parameters to consider in complex production setups makes the search space too large for an exhaustive search. It is common to apply hybrid methods combining production simulation with metaheuristic, Mixed-Integer Programming (MIP) or Constraint Programming (CP) solvers. Nevertheless, such hybrid solutions result in high computational times for

the considered planning period and leave little space for online scheduling applications. (Xie and Allen, 2015)

One promising research direction is the application of learning-based methods such as Reinforcement Learning for PPC activities. The main advantage of learning-based methods is the possibility to anticipate good solution strategies for given production conditions and goals with little computation time involved. The option to continuously learn and refine solution strategies without human supervision gives the learning-based methods a considerable advantage over common heuristic approaches (Bello et al., 2016).

Multiple studies formulate PPC as a Markov Decision Problem (MDP) (Howard, 1960) with subsequent use of RL. Building on the MDP process formulation, Qu et al. (2015) adopt a single Q-learning RL agent to control a multi-stage production system with three machines and the goal of minimizing the overall waiting time. Evaluation against a First-In-First-Out (FIFO) priority rule, a custom-made greedy heuristic and multi-agent Q-learning demonstrates better performance of the proposed scheduling RL agent. Training and evaluation are conducted in a custom-made simulation environment implemented in python.

The study of Waschneck et al. (2018) demonstrates one of the first deep RL applications on scheduling tasks for the example of semiconductor wafer processing. A custom-made production model implemented in MATLAB is used for training and evaluation. Three Deep Q-Network (DQN) agents independently control three machines and minimize the cycle time spread for three product groups. An alternative scenario used for comparative evaluation relies on a production setup with each machine being controlled according to an individual heuristic. The authors demonstrate a lower variation of production cycle times by using RL agents.

Thomas et al. (2018) adopt a single-agent DQN RL setup for job shop scheduling. The overall goal is the maximization of the production system throughput. Additionally, a neural network is introduced for continuous bottleneck detection. Recognized bottlenecks are mitigated via a temporal increase of throughput for involved machines. The production model is implemented as a discrete event simulation in the AnyLogic software.

Schneckenreither and Haeussler (2019) propose a DQN-based order-release mechanism superior to the backward infinite loading heuristic. The considered application covers a two-stage flow shop with three machines. An RL agent monitors arriving orders and determines when a batch of accumulated orders should be released in the production system. Implementation details of the production simulation are not provided.

Gannouni et al. (2020) address the problem of sequence-dependent setup waste minimization for a single-stage production process. A single RL agent builds production sequences for three machines and delivers good solution results in a fraction of the time required by Google OR-Tools and Gurobi solvers. The adopted neural architecture is a modification of the REINFORCE RL with Pointer Network (Nazari et al., 2018). The material flow and production processes simulation is built in python and uses Machine Learning (ML) models trained on real production data.

Scheduling of flexible job shop production is addressed in the work of Lang et al. (2020). Two independent DQN agents with Q-values approximated by recurrent

neural networks are responsible for the selection of orders and machines. Schedules generated by RL agents for production layouts with up to eight machines result in lower total makespan and tardiness compared to the randomized adaptive search (GRASP) metaheuristic, which is selected as a baseline. The production simulation is implemented using the open-source Salabim library in python.

Samsonov et al. (2021) propose a DQN-based RL solution for job shop scheduling problems for makespan minimization. The developed approach features an action space agnostic solution design and an exponential reward shape facilitating learning of near-optimal solutions. Evaluation is conducted on a production layout with six machines in a custom-made simulation implemented in python.

A problem of dynamic route guidance for material handling systems in automated production sites is investigated in the work of Hwang and Jang (2020). The authors propose a new RL algorithm based on the $Q(\lambda)$ learning approach with low computation overhead. The production and transport system simulation is conducted in the commercial software AutoMode. The achieved average delivery time and the total amount of deliveries suggest a good performance with little computation time compared to several common static and dynamic routing algorithms.

The study of Kuhnle et al. (2020) investigates several action spaces and reward designs for a learning-based production control system. The work focuses on the control of order dispatching and transportation between machines using a multi-agent RL setup and a production simulation implemented in python with the open-source library SimPy. The considered job shop production consists of eight machines unified in three machine groups. An evaluation against rule-based heuristics such as nearest job first (NJF) and FIFO demonstrates that the proposed RL setup can be successfully used for control for various production scenarios.

As Figure 1 shows, most of the RL-based solutions for PPC tasks which provide competitive results for online scheduling tasks rely on custom discrete-event simulations. Tailored solutions offer almost unlimited possibilities for customization and fast distributed training making it a well-suited choice for conceptual studies.

	Task				Process flow				Simulation				Real data	
	Order release	Order scheduling	Material routing		Job shop	Flow shop	Batch process	Continuous process	Python (custom)	MATLAB (custom)	AnyLogic	AutoMode	yes	no
Qu et al. (2015)	○	●	○		○	○	●	○	●	○	○	○	○	●
Waschneck et al. (2018)	●	●	○		●	○	○	○	○	●	○	○	○	●
Thomas et al. (2018)	○	●	○		●	○	○	○	○	○	●	○	○	●
Schneckenreither and Haeussler (2019)	●	○	○		○	●	○	○	○	○	○	○	○	●
Gannouni et al. (2020)	○	●	○		○	○	○	●	●	○	○	○	●	○
Lang et al. (2020).	○	●	○		●	○	○	○	●	○	○	○	○	●
Samsonov et al. (2021)	●	○	○		●	○	○	○	○	○	○	○	○	●
Hwang and Jang (2020).	○	○	●		○	○	○	○	○	○	○	●	○	○
Kuhnle et al. (2020)	●	○	●		●	○	○	○	●	○	○	○	○	●

Figure 1: Overview of the considered approaches

However, for widespread use of such methods in manufacturing companies, the applications must be robustly deployable and based on simulation tools already in use to be an alternative to current PPC software. The call for implementation in real production applications will gradually turn the spotlight on the questions of governance and certifications of production simulations used for training and production deployment of RL agents. Current approaches mostly focus on a job shop production showing a promising alternative for heuristics used in the PPC, because of the high complexity of today's manufacturing systems, which can't be controlled down to the smallest detail. In addition to the lack of practical relevance of some approaches in terms of the simulation tools used, the basis of purely synthetic data is also a shortcoming of the approaches used to date and calls for a more practical solution.

3 Approach and Methodology

Our experimental setup aims to demonstrate the general workflow for training and evaluation of RL agents for the task of order release under conditions close to practical applications. We use historical production data and a job shop based production structure implemented in Siemens Plant Simulation as a standard simulation software in the production domain. We design a suitable RL solution and devise an approach for deployment using a combination of two simulations as described in the following.

3.1 Job Shop Scheduling with Reinforcement Learning

Before an RL agent can be successfully applied to a given problem, it needs to go through a training period in which it learns how to solve the problem at hand. Once trained, an agent can make decisions very quickly, but the training period is often lengthy and requires many interactions with the given environment. Minimizing the time required for each individual interaction with the simulation environment is thus of high importance. In practice, such an environment is usually represented by a simulation of the process at hand. Rather than working with a static dataset collected upfront, a simulation allows for the possibility of exploration, i.e. the agent can influence the data it receives by performing actions and receives realistic subsequent state data.

In the production domain, simulations are often implemented with the help of commercial software such as Plant Simulation. If an RL agent is to be of any practical use, it needs to be interoperable with simulation tools already used as a standard by the domain experts. Enabling this interoperability, however, requires overcoming a set of challenges, because the existing commercial simulations are generally not explicitly designed to be used in the RL context. The central problem is that the state-of-the-art RL software packages are not supported by standard simulation software used in the production domain. Therefore, a communication layer must be created between the RL agent and Plant Simulation. While such an interface is well suited for the use with an already trained agent, it is not the ideal solution for training an agent due to the additional time delay introduced by the communication layer.

Hence, we propose a hybrid approach consisting of three stages as depicted in Figure 2: (1) Training of the agent in a custom-made simulation in SimPy closely imitating the desired target simulation and designed to minimize the training time. (2) Transfer learning, i.e. further training of the already trained agent in Plant Simulation, to adapt to minor differences between the two simulations. The number of training steps in this stage is thus kept to a minimum. If the difference between the two simulations is sufficiently small, the second step may be skipped entirely. (3) Application of the trained agent and running simulation in Plant Simulation for dynamic order release in a real production environment.

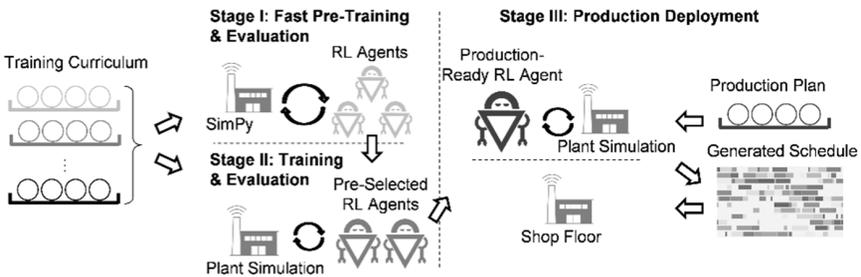


Figure 2: Hybrid training and evaluation scheme

To demonstrate the general feasibility of the methodology described above, we use the agent design, state space, and reward function introduced by Samsonov et al. (2021) and apply it to the more limited problem of order release, in which our agent does not schedule every operation, but only the first operation of every order. The remaining operations of each order follow the production plan and are scheduled on each machine according to the FIFO principle.

Due to the more limited scope of the problem, the state space of the agent contains only the backlog and remaining processing time on each machine, the duration of the first operation for each order, the total processing time required for each order and the type of machine required for the first order as well as the current simulation time. The agent is rewarded based on the comparison of the makespan it achieves for the given problem instance and the optimal makespan for this instance, as described in Samsonov et al. (2021).

3.2 Production Data and Simulation in Plant Simulation

The focus on order release in a real production environment necessitates the use of simulation software commonly used in real production plants. Moreover, validation against real historical production data and typical heuristics for order release is suggested, e.g. constant work in progress (CONWIP) or workload control. Therefore, a representative job shop production setup with originally 200 machines and a total of 15,000 depicted production orders is used for this purpose instead of creating an imaginary simulation model well suited for the used methodology but with no further regard to the industry’s reality.

For the final setup of dynamic order release at the appropriate time and rapid adaptability of the experiments, we use Siemens Plant Simulation as it is widely

used for discrete-event simulation of real production plants, closely modelling a wide range of production characteristics such as layout specific transition times.

Further research on the state space representation is needed to enable the RL methods used in this study to work with production scenarios consisting of hundreds of machines and thousands of orders. To test the hybrid training and evaluation concept proposed in this work, we hence downscale the model and the data set of the real production in two steps, such that the original model and the downscaled model deliver similar results. For this purpose, a simulation of the entire model is carried out and only those orders are considered that occupy the machines during the defined period. Then, the machines to be included are reduced by considering only those that have the highest utilization in the time interval under consideration. In this process, the processing times on missing work centers are represented by additional transition times. This results in a new data set with 160 orders produced on 10 machines in three days, which is an appropriate size for the methods examined here and will serve as the basis for the present study. To ensure comparability, the results of both the original and reduced simulations are evaluated and show a maximum deviation of 3% for the workstations' utilization and 10% for the throughput times.

3.3 Synthetic Training Data and Simulation in SimPy

To enable the first training stage of our approach, we implement a discrete-event production simulation for order release using the SimPy open-source library in python following the OpenAI Gym API, which is specifically designed to be a fast and flexible tool for RL training.

The production structure and program can change considerably over time. Such major changes might require retraining of the order release RL agent while having a limited amount of suitable historical data available. Starting with a limited amount of data consisting of the 20 production scenarios with different available machines and orders to produce (see Section 3.2), we generate a larger amount of synthetic data by mimicking the characteristics of this limited initial data set. Investigation of the operation time histogram and the Q-Q plot (see Fig. 3) of the log-transformed data demonstrates that the production data follows a lognormal distribution. Further investigations of transportation times and operation numbers per order demonstrate similarities to lognormal distributions.

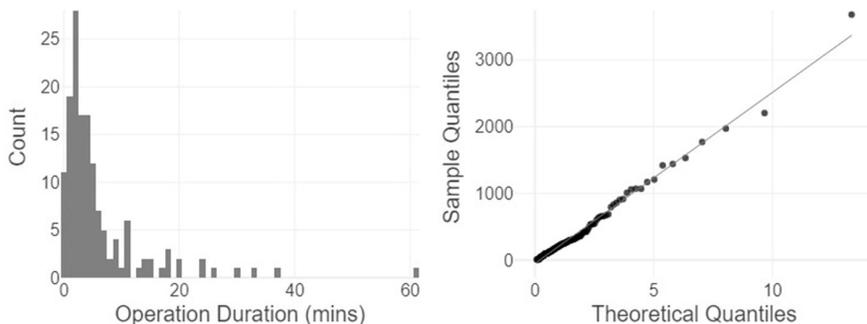


Figure 3: Distribution of operation durations & Q-Q plot of log transformed operation durations

Using the 20 available production scenarios, we estimate the probability density of the lognormal distributions for operation and transportation times, as well as the operation numbers per order, in terms of the mean and standard deviation. Those parameters are utilised to generate realistic synthetic production data and ensure the exposure of the RL agent to a wide range of production situations during training.

To evaluate the proposed hybrid training and evaluation scheme, several validation runs demonstrate identical behaviour with the production model implemented in Plant Simulation, yet it potentially lacks the level of implementation reliability provided by Plant Simulation. The python-based simulation is used primarily for fast training at the first stage of the proposed hybrid approach.

3.4 Connecting Plant Simulation and RL Agent

To enable communication between the RL agent and Plant Simulation, an interface to exchange messages between the two components is implemented via network sockets. The proposed methodology requires a bidirectional information flow, with the simulation sending the current state to the agent and the agent in return deciding which orders to release based on this state. This decision is then transferred back to the simulation, which continues to run until another selection is pending. In the resulting message exchange, each component alternately sends a message and subsequently waits for a response from the other component.

While messages from the agent to the simulation merely contain the order numbers to be released, the structure of messages from the simulation to the agent is more complex and contains several sections described in Table 1.

Table 1: Sections of a state message as sent from the simulation to the agent

Section	Content
Header	Starting positions of the remaining sections in the message
Available Orders	List of orders currently available for release
KPIs	Number of orders that are late, simulation time to date in seconds, total slack of all orders that are not late, delay (negative slack) of late orders
Machine Information	List containing a triplet (Machine name, remaining processing time on machine, machining and setup time in machine queue) for each machine
Finished orders	Orders which have been finished between the previous message and the current one

The state information exchanged through the interface contains only information that may change from step to step. Additional, static information required by the agent is loaded upfront from a file describing the current scheduling instance in order to avoid unnecessarily long messages, which negatively affect the speed of communication. This static information includes the processing times for individual operations, and on which machine each operation needs to be processed.

Knowing the length of a message in advance enables each component to ascertain whether a particular message has been received fully and thus ensures a more robust

communication between the two components. Since the state message as communicated by the simulation varies in size from step to step, the simulation will communicate the length of the incoming message before sending its contents.

4 Experimental Results

To validate the approach presented here, we first evaluate the performance of our agent by investigating whether the agent's performance improves during training and how the trained agent fares in comparison with commonly used order release rules in Plant Simulation. As Figure 4 shows, the agent's performance, as measured by the achieved makespan in a problem instance, improves during training, indicating that the agent is learning how to solve the problem.

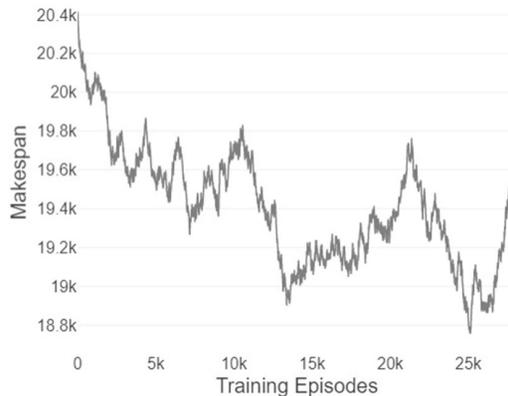


Figure 4: RL agent performance during the training process (total makespan)

To investigate the performance of the trained agent, we apply it to each of the instances described in Section 3.2. As a comparison, we further apply the CONWIP rule (Lödding, 2016) with 15, 30 and 50 concurrent orders in the production simulation. Each of these four approaches is then evaluated on two metrics, the utilization rate and the average waiting time, which are deliberately chosen instead of the makespan. The makespan is well suited for training because each training instance consists of a limited number of orders to be released and the agent can then learn to minimise the time it takes for the system to process all orders. In practice, however, there typically is no limit on the number of orders, but rather a continuous influx of new orders, such that the task is never “done”. We replicate this behaviour and simply cut the simulation off at a predefined point in time to arrive at an evaluation period. Since the overall makespan is now no longer a useful metric anymore, we turn to the two commonly used metrics described above.

The results are illustrated in Figure 5, where each data point represents the solution of one problem instance by one particular method. Each different method forms a clear cluster with different trade-offs between the utilization rate and the average waiting time. The RL agent, however, achieves both comparatively high utilization rates and low average waiting times.

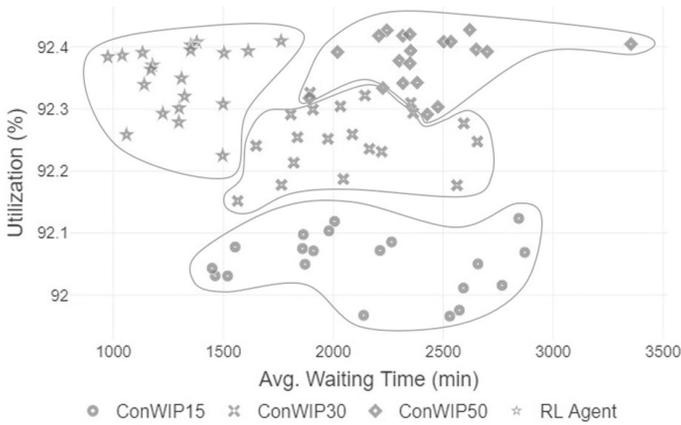


Figure 5: Quality of the solutions obtained by the agent and CONWIP

To quantify and emphasize the necessity of the hybrid simulation approach presented here, we investigate the time individual simulation steps require in an RL setting using both simulations. As Figure 6 elucidates, there is a marked difference between the two simulations, with the average step in Plant Simulation requiring 37.6 ms while the average step in the native python simulation only requires 1.4 ms. Although both values may seem fairly small at first glance, the sheer number of steps required in RL, especially in more complex problem settings can lead to hours or days of training time. A simulation that is slower by a factor of ~ 27 , as in this case, can make an already lengthy training process impractically long. This difference in speed is not necessarily because Plant Simulation is inherently slower, but because communication through the socket interface incurs extra time costs.

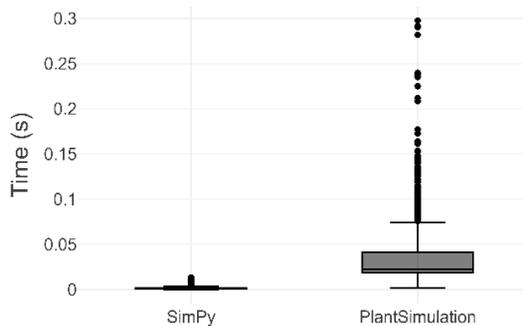


Figure 6: Time required for one simulation step

5 Conclusion

In this work, we argue that while RL is evolving into a promising approach to solve common PPC tasks, integrating it into the existing simulation ecosystem used in the production domain remains non-trivial. To enable the usage of RL in conjunction with commercial production simulations, we propose a hybrid methodology and

demonstrate its feasibility on the example of order release tasks. Since communication between the commercial production simulation and the RL agent is costly in terms of time, the training becomes impractical. Instead, we propose to perform the bulk of the training in a native python simulation, which closely resembles the original model, and then transferring the trained agent to the commercial simulation. If necessary, the agent can be retrained for a much more limited duration to compensate for any potential differences between the two simulations.

We show that an RL agent trained in the native python simulation on synthetic data designed to resemble the characteristics of real production data can be successfully transferred to a Plant Simulation production environment. Not only does the communication protocol allow the RL agent to control this simulation, but the initially learned dispatching strategies lead to satisfactory outcomes even without additional retraining of the RL agent. We measure these outcomes by the average waiting times and utilization rates across 20 problem instances. As the results in Section 4 show, the RL agent achieves higher utilization rates on average and lower average waiting times compared to the commonly used CONWIP heuristic.

While we demonstrate the general validity of our approach, extending our work to more complex PPC tasks such as sequence building is the main direction for future work. Although further training of the agent in the commercial simulation was not necessary in this example, increasing both the complexity of the tasks performed and the simulation scenarios may lead to greater differences between the two simulations in use. The second training step should therefore remain a part of the presented approach and its necessity should be evaluated on an individual basis.

Acknowledgement

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612

Bibliography

- Bello, I.; Pham, H.; Le, Q.V.; Norouzi, M.; Bengio, S.: Neural combinatorial optimization with reinforcement learning. arXiv:1611.09940 (2016).
- Gannouni, A.; Samsonov, V.; Behery, M.; Meisen, T.; Lakemeyer, G.: Neural Combinatorial Optimization for Production Scheduling with Sequence-Dependent Setup Waste. In: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2020, pp. 2640–2647.
- Howard, R.A.: Dynamic programming and markov processes (1960).
- Hwang, I.; Jang, Y.J.: $Q(\lambda)$ learning-based dynamic route guidance algorithm for overhead hoist transport systems in semiconductor fabs. *International Journal of Production Research* 58 (2020) 4, pp. 1199–1221.
- Kuhnle, A.; Kaiser, J.-P.; Theiß, F.; Stricker, N.; Lanza, G.: Designing an adaptive production control system using reinforcement learning. *Journal of Intelligent Manufacturing* 54 (2020) 4, pp. 1196.

- Lang, S.; Behrendt, F.; Lanzerath, N.; Reggelin, T.; Müller, M.: Integration of Deep Reinforcement Learning and Discrete-Event Simulation for Real-Time Scheduling of a Flexible Job Shop Production. In: Bae, K.-H.; Feng, B.; Kim, S.; Lazarova-Molnar, S.; Zheng, Z.; Roeder, T.; Thiesing, R. (Hrsg.): Proceedings of the 2020 Winter Simulation Conference. Orlando, USA 2020, pp. 3057–3068.
- Lang, S.; Schenk, M.; Reggelin, T.: Towards Learning- and Knowledge-Based Methods of Artificial Intelligence for Short-Term Operative Planning Tasks in Production and Logistics: Research Idea and Framework. IFAC-PapersOnLine 52 (2019) 13, pp. 2716–2721.
- Lödding, H.: Verfahren der Fertigungssteuerung: Grundlagen, Beschreibung, Konfiguration. Berlin, Heidelberg: Springer Vieweg 2016.
- Nazari, M.; Oroojlooy, A.; Snyder, L.V.; Takáč, M.: Reinforcement Learning for Solving the Vehicle Routing Problem, 2018,
- Qu, S.; Chu, T.; Wang, J.; Leckie, J.; Jian, W.: A centralized reinforcement learning approach for proactive scheduling in manufacturing. In: Proceeding of the 20th Conference on Emerging Technologies & Factory Automation 2015, pp. 1–8.
- Samsonov, V.; Kemmerling, M.; Paegert, M.; Lütticke, D.; Sauermann, F.; Gützlaff, A.; Schuh, G.; Meisen, T.: Manufacturing Control in Job Shop Environments with Reinforcement Learning. In: International Conference on Agents and Artificial Intelligence (ICAART), 2021, pp. 589–597.
- Schlegel, T.; Siegert, J.; Bauernhansl, T.: Metrological Production Control for Ultra-flexible Factories. Procedia CIRP 81 (2019) 3, pp. 1313–1318.
- Schneckenreither, M.; Haeussler, S.: Reinforcement Learning Methods for Operations Research Applications: The Order Release Problem. In: Nicosia, G.; Pardalos, P.; Giuffrida, G.; Umeton, R.; Sciacca, V. (Hrsg.): Machine Learning, Optimization, and Data Science. Cham: Springer 2019, pp. 545–559.
- Schuh, G.; Prote, J.-P.; Sauermann, F.; Franzkoch, B.: Databased prediction of order-specific transition times. CIRP Annals 68 (2019) 1, pp. 467–470.
- Thomas, T.E.; Koo, J.; Chaterji, S.; Bagchi, S.: Minerva: A reinforcement learning-based technique for optimal scheduling and bottleneck detection in distributed factory operations. In: 2018 10th International Conference on Communication Systems & Networks (COMSNETS), Bengaluru, 1/3 - 1/7, 2018, pp. 129–136.
- Waschneck, B.; Reichstaller, A.; Belzner, L.; Altenmüller, T.; Bauernhansl, T.; Knapp, A.; Kyek, A.: Deep reinforcement learning for semiconductor production scheduling. In: Proceedings of the 29th Annual 2018, pp. 301–306.
- Xie, C.; Allen, T.T.: Simulation and experimental design methods for job shop scheduling with material handling: a survey. The International Journal of Advanced Manufacturing Technology 80 (2015) 1-4, pp. 233–243.
- Xie, J.; Gao, L.; Peng, K.; Li, X.; Li, H.: Review on flexible job shop scheduling. IET Collaborative Intelligent Manufacturing 1 (2019) 3, pp. 67–77.
- Zijm, H.; Regattieri, A.: Manufacturing Planning and Control Systems. In: Zijm, H.; Klumpp, M.; Regattieri, A.; Heragu, S. (Hrsg.): Operations, Logistics and Supply Chain Management. Cham: Springer International Publishing 2019, pp. 251–271.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Entscheidungsbäume und bestärkendes Lernen zur dynamischen Auswahl von Reihenfolgeregeln in einem flexiblen Produktionssystem

Using Decision Trees and Reinforcement Learning for the Dynamic Selection of Sequencing Rules in a Flexible Manufacturing System

Jens Heger, Mazhar Zein El Abdine, Sughanthan Sekar, Thomas Voß,
Leuphana Universität Lüneburg, Lüneburg (Germany),
heger@leuphana.de, abdine@leuphana.de, sughanthan.sekar@stud.leuphana.de,
voss@leuphana.de

Abstract: The integration of machine learning methods can improve the performance of a manufacturing system by adjusting priority rules dynamically. In this paper, three methods for adjusting the k-factors of the ATCS sequencing rule were trained: neural networks, decision trees and reinforcement learning. They are evaluated in a static and a dynamic scenario. The required dataset was synthetically generated using a discrete event simulation, where product mix and system utilization were varied within a flow shop environment. Across all scenarios, it is shown that all three methods can improve the performance. On par, RL and NN can reduce the mean tardiness by 15% and compensate for unplanned product mix changes.

1 Einleitung und Problemstellung

Die Reihenfolgebildung von Vorgängen an einer Maschine kann unter sich ändernden Bedingungen, wie zum Beispiel Maschinenausfall schwierig sein. Da die Anwendung von zentralen und statischen Lösungsverfahren in komplexen und unsicheren Szenarien nicht geeignet ist, stellen dezentrale Reihenfolge-Regeln eine praktikable Option dar. Diese nutzen lokal verfügbare Informationen zur schnellen Entscheidungsfindung. Es existiert jedoch keine Regel, die alle anderen bei unterschiedlicher Systemleistung übertrifft.

Aus diesem Grund wird eine Hyperheuristik entwickelt, um Gewichtungsfaktoren eine Prioritätsregel zur Reihenfolgebildung dynamisch, auf Basis des Systemstatus, auszuwählen und anzupassen. Basierend auf mehreren Trainingsszenarien unter Berücksichtigung stochastischer Einflüsse, wie z.B. variierende Ankunftszeiten oder wechselnde Anteile im Produktmix, werden die Vorteile von dynamischer Anpassung der Regel vorgestellt. Für die dafür benötigte Schätzung des

Systemverhaltens zur Auswahl und Anpassung der Gewichte der Regel, passend zum aktuellen Zustand, haben unterschiedlichste Modelle des maschinellen Lernens sehr gute Ergebnisse geliefert (Mouelhi-Chibani und Pierreval, 2010; Heger, 2014; Shiue et al., 2018). Bei der Verwendung dieser Methoden stellt sich aber nicht zuletzt die Frage nach der Menge an Trainingspunkten, dem Nachlernen von Wissen in neuen Szenarien und der Tauglichkeit, Generalisierbarkeit und Nachvollziehbarkeit der verwendeten Methoden (Priore et al., 2006; Priore et al., 2018; Usuga Cadavid et al., 2020). Das dabei generierte Wissen und Verständnis der Handlungen wird anstelle von einfacher Vorhersage und Blackbox Optimierung zunehmend bevorzugt (Nunes und Jannach, 2017; Rehse et al., 2019). An dieser Stelle kann der verwendete Ansatz der Hyperheuristik zur Auswahl unterschiedlicher Reihenfolgeregeln in Kombination mit nachvollziehbaren Lernmethoden (z.B. Entscheidungsbäumen) seine Tauglichkeit beweisen.

2 Stand der Technik

Auf Grund ihrer einfachen Nachvollziehbarkeit und der sehr kurzen Rechendauer ist die Verwendung von Prioritätsregeln zur Reihenfolgebildung, d.h. zur Auswahl des nächsten Auftrags zur Bearbeitung durch die Maschinen, in der Industrie sehr beliebt. Dabei ist zu beachten, dass über 100 Regeln bekannt sind, die je nach Szenario unterschiedliche Leistungen bringen (Panwalkar und Iskander, 1977). Im Verlauf der Jahre wurden kombinierte Prioritätsregeln entwickelt, die mehrere Auftragsattribute gleichzeitig betrachten, um die Leistung des Systems zu verbessern. So auch die Regel „Apparent Tardiness Cost“, die neben der gewichteten Prozesszeit auch den geplanten Fertigstellungszeitpunkt und einen Gewichtungsfaktor (k_1) betrachtet (Vepsalainen und Morton, 1987). Mit Blick auf die Rüstzeiten wurde die Regel dann um einen Rüstzeiterm erweitert und wird seither als „Apparent Tardiness Cost with Setups“ (ATCS) beschrieben. Der zusätzliche Term bezeichnet das Verhältnis zwischen reihenfolgeabhängiger Rüstzeit und durchschnittlicher Rüstzeit, multipliziert mit dem zweiten Gewichtungsfaktor (k_2). Die Regel wird in der in Gleichung (1) gezeigten Form für die folgende Studie verwendet. Die Kombination von drei Attributen und die Verwendung von zwei Gewichtungsfaktoren ermöglichen es, bei einer korrekten Dimensionierung über eine breite Masse an Szenarien hinweg gute Leistung zu erzielen (Lee et al., 2002).

$$Z_i^t = \frac{w_i}{p_i} \exp\left(-\frac{(d_i - t - p_i)^+}{k_1 \bar{p}}\right) \exp\left(-\frac{s_{i,l}}{k_2 \bar{s}}\right) \quad (1)$$

Wissend, dass die Systemleistung stark von der korrekten Auswahl der k-Faktoren passend zur Systemauslastung abhängt, werden diese dynamisch an die Situation in der Fertigung angepasst. Es handelt sich bei der dynamischen Anpassung folglich um eine Hyperheuristik. Zum Aufbau der Wissensbasis über den Zusammenhang der k-Faktoren und der resultierenden Leistung müssten allerdings alle möglichen Kombinationen aus k-Faktoren, Produktmix und Systemzustand simuliert werden. Da auf Grund der Komplexität in realen Systemen nicht alle möglichen

Kombinationen an Einflussfaktoren simuliert werden können, werden häufig definierte Kombinationen aus Systemzuständen simuliert und die unbekannt Situationen durch ein Regressionsverfahren geschätzt.

Dabei stellt die Verwendung von Neuronalen Netzen (NN) die Basiskennlinie dar. Die Kombination von NN zur Vorhersage der Systemleistung wurde ausführlich hinsichtlich der Abhängigkeit von den k-Faktoren und der Systemauslastung betrachtet (Mönch et al., 2006; Mouelhi-Chibani und Pierreval, 2010; Heger, 2014; Heger et al., 2016). Trotz guter Ergebnisse ist hierbei jedoch festzustellen, dass die NN grundsätzlich als Blackbox verwendet werden und keine Aussage über den Einfluss bestimmter Faktoren ermöglichen. Aus diesem Grund wird die Verwendung von NN als Basislinie mit der Verwendung von Entscheidungsbäumen und bestärkendem Lernen verglichen (Rai, 2020).

Die Verwendung von Entscheidungsbäumen erhielt im Kontext von erklärbaren Modellen (engl.: Explainable AI) des maschinellen Lernens deutliche Aufmerksamkeit (Rai, 2020; Puiutta; Veith, 2020).

Im Gegensatz zu komplexen Modellen, wie tiefe NNs, die nicht-interpretierbare Blackbox-Modelle erzeugen, sind Entscheidungsbäume regelbasierte Modelle, die dem Benutzer eine intuitive Darstellung von Regeln und Abläufen bieten. An jedem Knoten eines Entscheidungsbaums wird eine bestimmte Zielfunktion geprüft. Das Ergebnis liefert den Pfad zum neuen Knoten. Die Struktur wiederholt sich, bis eine bestimmte Voraussetzung getroffen wird.

Regeln können von Pfaden durch den Entscheidungsbaum abgeleitet und Entscheidungsbäume sowohl für Klassifikations- als auch für Regressionsaufgaben verwendet werden. Somit eignen sie sich beispielsweise zur dynamischen Auswahl von Prioritätsregeln, jedoch mit relativ niedriger Leistung in unbekannte Szenarien (Shahzad und Mebarki, 2016). Zudem sind sie, zusammen mit weiteren baumbasierten Verfahren, wie Random Forest und XGBoost, in deren Anwendungen weit verbreitet. Diese Verfahren, die auf Entscheidungsbäume basieren, haben allerdings ein gewisses Maß an Interpretierbarkeit eingebüßt, um bessere Schätzungen und Generalisierbarkeit dieser zu erreichen. Dennoch werden sie immer häufiger mit weitere Funktionalitäten ausgestattet, um die Interpretation zu verbessern (Lundberg et al., 2020).

Die Verwendung von bestärkendem Lernen hat als Hyperheuristik bei der dynamischen Auswahl von Reihenfolgeregeln bereits gute Ergebnisse erzielt. Studien zeigen, dass die inhärenten Vorteile des bestärkenden Lernens, im Gegensatz zu den Methoden des überwachten Lernens, bei der direkten Interaktion mit dem System liegen. Der Agent lernt die richtigen Verhaltensweisen basierend auf dem beobachteten Verhalten und der erhaltenen Rückmeldung. Konkret handelt es sich bei dem Anwendungsfall um die Auswahl von Prioritätsregeln für alle Maschinen im System (Heger und Voss, 2020, 2021). Die Autoren zeigen, dass ausgehend von der beobachteten Systemauslastung und Warteschlangen die Leistung durch dynamische Auswahl von Reihenfolgeregeln verbessert werden kann. Analog dazu zeigen andere Autoren in ihrer Studie, dass die dynamische Anpassung von maschinenspezifischen Regeln in verschiedenen Szenarien signifikante Leistungsverbesserungen ermöglicht (Shiue et al., 2018, 2020).

In diesem Beitrag wird geprüft, inwieweit die drei beschriebenen Verfahren zur Unterstützung der Auswahl von geeigneten k-Faktoren taugen. Dabei ist zu prüfen,

inwieweit die Verwendung von nachvollziehbaren Handlungen zu einer Reduktion der Leistung führt. Das Zusammenspiel zwischen Leistung und Erklärbarkeit wird im Kontext des vorgestellten Szenarios genauer untersucht. Zusätzlich ist zu prüfen, unter Verwendung der trainierten Hyperheuristiken, ob diese in einem unbekanntem Szenario immer noch in der Lage sind, die k-Faktoren entsprechend dem Systemstatus auszuwählen und dynamisch zu ändern, und dadurch eine robuste Leistung zu erzielen.

3 Simulation und Szenario

Die Studie wird im Rahmen eines Fertigungssystems mit reihenfolgeabhängigen Rüstzeiten eingesetzt und durchgeführt. Dabei werden ungeplante und unbekannte Änderungen, wie Produktmix-Wechsel und Auslastungsschwankungen, hinzugefügt, um ein Verhalten der verschiedenen Methoden in unbekanntem Szenarien betrachten zu können. Im Folgenden ist das Szenario detailliert beschrieben:

Tabelle 1: Detaillierte Szenariobeschreibung

System	Maschinen: 10 Maschinengruppen: 5 Organisationsform: Fließfertigung Produktfamilien: 4 Produktfamilien Verteilung: je nach Produktmix Operationen pro Auftrag: 10
Auftragsparameter	Verteilung der Zwischenankunftszeit: Poisson Prozessbearbeitungszeit: 1 – 99 Verteilung der Prozesszeit: gleichverteilt Fälligkeitstermin: TWK Methode
Simulation	Aufwärmphase: 2500 Aufträge Simulationsdauer: 12500 Aufträge
KPIs	Durchschnittliche Verspätung Durchschnittliche Durchlaufzeit

Der besondere Fokus liegt dabei auf den reihenfolgeabhängigen Rüstzeiten, die im Folgenden dargestellt sind. Je nach Produktmix sind die Anteile der vier Produkttypen unterschiedlich, was zu mehr oder weniger großen Rüstzeitanteilen führt. So ist davon auszugehen, dass ein Produktmix mit den ersten drei Produktfamilien deutlich weniger Rüstzeit benötigt als ein Produktmix, der alle vier Produktfamilien enthält. Dabei kann die Matrix in (2) wie folgt gelesen werden: die Rüstzeit von Familie 1 auf Familie 2 beträgt 5 Minuten; die Rüstzeit von Familie 1 auf Familie 4 beträgt 25 Minuten. Bei einer durchschnittlichen Bearbeitungszeit von 50 Minuten kann die Rüstzeit massiven Einfluss auf die Leistung des Systems haben.

$$\begin{pmatrix} 0 & 5 & 10 & 25 \\ 5 & 0 & 10 & 25 \\ 5 & 5 & 0 & 25 \\ 5 & 5 & 10 & 0 \end{pmatrix} \quad (2)$$

Zur Messung der Leistung werden die durchschnittliche Verspätung sowie die durchschnittliche Durchlaufzeit dokumentiert. Die durchschnittliche Verspätung ergibt sich aus dem geplanten Fertigstellungszeitpunkt und der tatsächlichen Fertigstellung. Dabei ist zu beachten, dass zu früh gefertigte Aufträge mit einer Verspätung von 0 bewertet werden. Der Fertigstellungszeitpunkt berechnet sich aus der Summe der Startzeit und der durchschnittlichen geplanten Bearbeitungszeit multipliziert mit einem Verspätungsfaktor (engl. Due Date Factor). Dieser ist so gewählt, dass eine gewisse Rüst-, Wart- und Transportzeit zwischen den Maschinen akzeptabel ist.

Für das Erstellen von Regressionsmodellen werden in einer umfangreichen Parameterstudie Trainingsdaten mit Hilfe von ereignisdiskreter Simulation generiert. Die Verwendung von Simulation ermöglicht es, die Länge und Breite des Datensatzes selbst zu bestimmen. Diese generierten Daten bilden die Grundlage für die wissensbasierten Ansätze der dynamischen Anpassung. Das Simulationsmodell steht in diesem Fall als Trainingsumgebung für das bestärkende Lernen zur Verfügung.

In der Abbildung 1 ist das Verfahren beschrieben. Beginnend unten in der Mitte steht das Simulationsmodell. Durch die Parameterstudie zur Betrachtung der Leistung über die unterschiedlichen Systemzustandskombinationen ist es möglich, Trainingsdaten für die unterschiedlichen Regressionsmodelle zu erstellen.

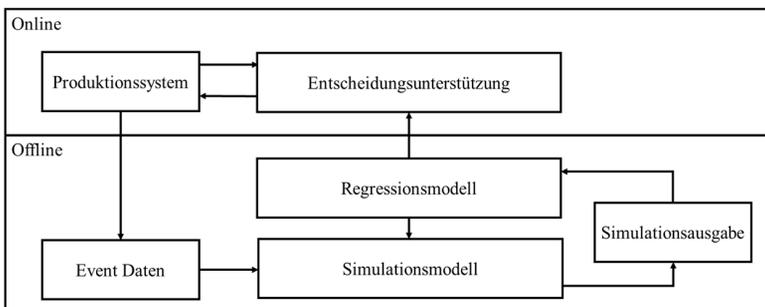


Abbildung 1: Konzept zur Verwendung von Regressionsmodellen zur Entscheidungsunterstützung

Diese werden dann als Entscheidungsunterstützung im realen Produktionssystem genutzt. Ausgehend von dem aktuellen Systemzustand (z.B. Maschinenauslastung und Produktmix) schlägt die Entscheidungsunterstützung einen neuen k -Werte-Paar vor. Eine Simulation aller k -Werte würde in der Onlineanwendung viel Zeit in Anspruch nehmen.

In der Abbildung 2 ist das Verfahren zur dynamischen Anpassung schematisch dargestellt. Dabei wird die Leistung des Systems über die Zeit und unter sich verändernden Zuständen sowie einer Auswahl an k -Werten betrachtet. Je nach Ausprägung des Leistungsindicators gilt es, den Wert zu minimieren oder zu maximieren; in diesem Beitrag wird die Minimierung der durchschnittlichen Verspätung betrachtet. Zu definierten Zeitpunkten wird, unabhängig vom Verfahren, durch die Entscheidungsunterstützung eine Aussage über die für die Situation am

besten geeigneten k -Werte zur Verbesserung der Leistung getätigt, welche dann für einen definierten Zeitraum konstant gehalten werden. Es ist zu evaluieren, ob und bis zu welchem Grad die Schwankung durch die inhärenten stochastischen Unsicherheiten des Simulationsmodells oder durch die Auswahl der k -Faktoren geschieht. Im Rahmen der Simulationsstudie kann weiterhin evaluiert werden, wie die Häufigkeit der Anpassung Einfluss auf die Leistung nimmt.

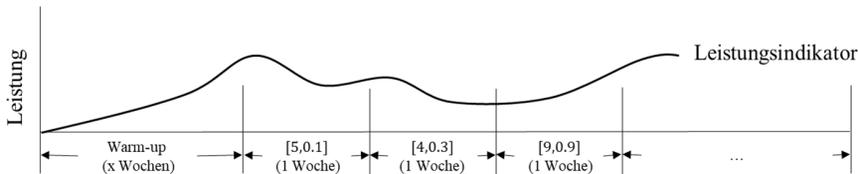


Abbildung 2: Systematische Abbildung des Vorgehens über den Simulationsverlauf

Eine Parameteranalyse wird in dem Zusammenhang durchgeführt, um herauszufinden, welche Beobachtungen signifikanten Einfluss haben und welche nicht. Die Datenvorverarbeitung wie Standardisierung, One-Hot-Encoding und einer Kombination daraus werden unabhängig von der verwendeten Methode durchgeführt, auf Grund des Simulationsfokus werden die Auswirkungen allerdings nicht detailliert evaluiert.

4 Auswertung

Als Trainingsdatensatz wurde eine Parameterstudie mit allen möglichen Kombinationen aus k_1 -Werten von 1 bis 10 und k_2 -Werten von 0.01 bis 1.01 unter 7 verschiedenen Einlastungen von 85 % bis 95 % ebenso wie 12 verschiedenen Produktmischen mit unterschiedlichen Rüstanteilen durchgeführt. Für jede der 9240 individuellen Parameterkombinationen wurden 5 Replikationen durchgeführt. Der verwendete Datensatz umfasst damit 46 200 Samples. Die Beobachtungen aus dem System waren die durchschnittliche Verspätung, die durchschnittliche Durchlaufzeit sowie die durchschnittliche Maschinenauslastung des Szenarios. Dabei sind die k -Faktoren, die Auslastung sowie die Leistungsindikatoren kontinuierliche Variablen. Der Produktmix wird als kategorische Variable betrachtet.

In den Rohdaten ist klar zu erkennen, dass unterschiedliche Produktmische bei derselben Einlastung und der Verwendung von statischen k -Werten unterschiedliche Leistungen bringen. Exemplarisch ist dies für zwei Produktmische in Abbildung 3 zu erkennen. Ein kleiner k_2 -Wert, welcher vorteilhaft für Produktmix 1 ist, würde bei Produktmix 6 zu einer Verschlechterung der Leistung um 30 % führen. Weiterhin zeigen die Rohdaten, dass auf Grund der unterschiedlichen Rüstzeitanteile pro Produktmix die Auslastung des Systems bei konstanter Einlastung bis zu 5 % variieren kann. In dieser Studie ist eine Einlastung von 85 % vorgesehen. Es ist festzuhalten, dass eine zu geringe Einlastung dazu führt, dass kein Verbesserungspotenzial möglich ist. Diesen Sachverhalt gilt es über mehrere Produktmische zu erkennen und zur Verbesserung der Leistung zu nutzen.

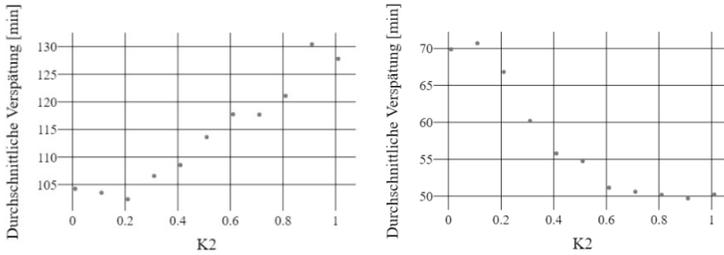


Abbildung 3: Leistung von Produktmix 1 (links) und Produktmix 6 (rechts) bei 85 % Einlastung unter unterschiedlichen k_2 -Werten.

Zusätzlich zum ersten Datensatz mit 46200 Datenpunkte wird ein zweiter Datensatz mit 13860 Datenpunkte (entspricht 30 % des ersten Datensatzes) generiert. Im Anschluss wurden die NN ebenso wie die Entscheidungsbäume (DT) auf beiden Sätzen trainiert, um eine Aussage über die Verbesserung der Präzision mit mehr Datenpunkten machen zu können. Eine gute Konfiguration für die NN sowie die DT wurde mit Hilfe eines vorgelagerten Grid-Search-Verfahrens ermittelt. Das NN wurde mit scikit learn als Multi Layer Perceptron in Python implementiert. Die Parametrierung wurde mit Hilfe eines Grid-Search-Verfahren durchgeführt. Das resultierende, zweilagige Netz mit 10 Neuronen in der ersten und 30 Neuronen in der zweiten Schicht hatte die Aktivierungsfunktion „relu“. In Kombination mit dem Solver „adam“ zeigte eine Minibatchgröße von 500 Samples gute Ergebnisse. Die initiale Lernrate wurde mit 0.01 festgelegt. Eine L2-Regulierung wurde durchgeführt.

Die DT wurden mit scikit learn als Entscheidungsbaum Regressor in Python implementiert. Die endgültige Konfiguration wurde mit Hilfe eines Grid-Search-Verfahren ermittelt. In diesem Rahmen zeigte sich, dass die Verwendung einer maximalen Tiefe von 5, mit mindestens 4 Samples pro Blatt zu guten Ergebnisse führte. Der RL-Agent wurde mit Hilfe der Pathmind Software-as-a-Service-Plattform mit 12000 Simulationsläufen trainiert. In dem Zuge wurde das ereignisdiskrete Simulationsmodell als alleinstehendes Java-File exportiert und auf der Plattform, unabhängig von lokalen Ressourcen, über 12 Stunden trainiert. Dabei wurden verschiedene Hyperparameterkonfigurationen im Rahmen von populationsbasiertem Training automatisch evaluiert und die beste Konfiguration für das Szenario gefunden. Pathmind verwendet Ray und RLlib für das Training des Agenten. Die Strategie des Agenten wurde durch Proximal Policy Optimization trainiert.

Im Rahmen der Evaluation werden die drei Methoden in der Onlineverwendung getestet. Dabei wird für ein bekanntes Szenario bei statischer Einlastung und bekanntem Produktmix die Leistung dokumentiert. Im Folgenden handelt es sich um eine Einlastung von 85 % und den Produktmix 1 von oben. In der Abbildung 4 ist exemplarisch das Verhalten des Agenten, trainiert mit bestärkendem Lernen, abgebildet. Wie bereits oben zu sehen war, ist für den Produktmix 1 ein niedriger k_2 -Wert von Vorteil. Es ist klar zu erkennen, dass bei Schwankungen der durchschnittlichen Verspätung (linke Y-Achse) im System über die Zeit (X-Achse) die verwendeten k_2 -Werte der ATCS-Regel (rechte Y-Achse) angepasst wird.

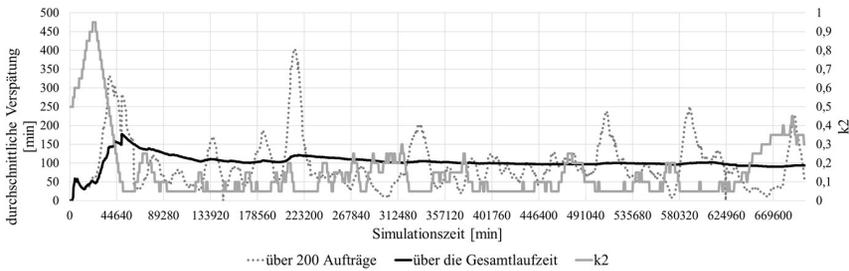


Abbildung 4: Änderung der k_2 -Werte durch den RL-Agenten zur dynamischen Anpassung an die Systemauslastung

Über 30 Replikationen zeigt sich, dass bei einer statischen Einlastung und bei einem bekannten Produktmix die dynamische Anpassung der k -Werte mit RL einen positiven Einfluss hat, sich aber nicht signifikant zur statischen Auswahl von k -Werten unterscheidet. Besonders interessant (siehe Abbildung 5) ist an dieser Stelle die schlechte Leistung der NNs; es ist davon auszugehen, dass das dynamische Umschalten und Anpassen durch das NN einen negativen Einfluss auf die Leistung im statischen Szenario hat (Priore et al., 2006).

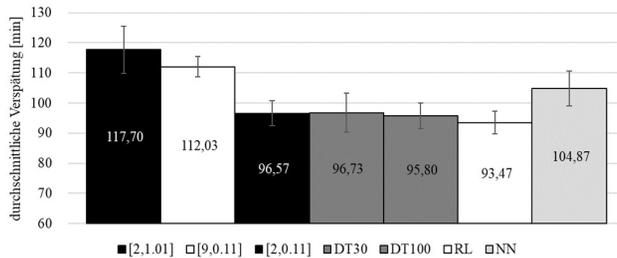


Abbildung 5: Vergleich der Leistung bei statischen k -Werten und der dynamischen Anpassung in einem bekannten Szenario

Im zweiten Evaluationsszenario wird die dynamische Anpassung bei einer statischen Einlastung unter wechselnden Produktmixen getestet. Im Szenario wird über $\frac{1}{4}$ der Simulationszeit ein neuer Produktmix (in diesem Fall Produktmix 6) im System betrachtet. Dieses Verhalten ist exemplarisch an das Weihnachtsgeschäft angelehnt. Es zeigt sich in der Abbildung 6, dass die Auswahl von guten statischen k -Werten bereits zu guter Leistung führt. Im Vergleich dazu bringen die DTs, die NNs ebenso wie der RL-Agent eine zusätzliche signifikante Verbesserung von bis zu 15 %. Dabei ist der RL-Agent noch 3 % besser als die Verwendung der DTs. Im Gegensatz zum statischen Szenario kann das NN seine Vorteile bzgl. der Generalisierbarkeit von Verhalten zeigen. Die vergleichbare Leistung von RL und NN ist nachvollziehbar, da RL zur Schätzung der Belohnung NNs verwendet.

Über die Evaluation in beiden Szenarien zeigt sich, dass DTs bekanntes Systemverhalten sehr gut wiedergeben können und bis zu einem gewissen Grad dynamische Verhalten beschreiben können. Die Verwendung von NN und RL ist besonders in Szenarien mit unbekanntem Verhalten im Vorteil und kann zu einer Verbesserung der Leistung von bis zu 15 % führen.

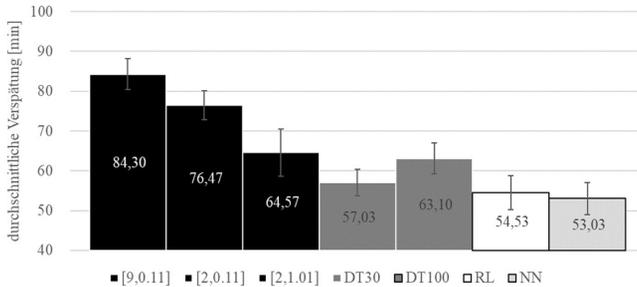


Abbildung 6: Vergleich der Leistung bei statischen k-Werten und der dynamischen Anpassung in einem unbekanntem Szenario

5 Zusammenfassung und Ausblick

Die dynamische Anpassung von Prioritätsregeln mit Hilfe von verschiedenen Methoden des maschinellen Lernens kann zu einer Verbesserung der Leistung führen. Im Rahmen des Beitrags wurden drei Verfahren zur Anpassung der k-Faktoren der ATCS-Regel trainiert und über zwei Szenarien evaluiert. Am Beispiel einer Fließfertigung wurde mit Hilfe einer ausführlichen Parameterstudie ein Datensatz über die Zusammenhänge zwischen Produktmix, k-Faktoren und Systemauslastung erstellt. Dieses wurde anschließend als Trainingsgrundlage für DT und NN genutzt. Das ereignisdiskrete Modell wurde als Trainingsumgebung für den RL-Agenten verwendet. Der Vergleich im statischen Szenario zeigt, dass DT und RL in der Lage sind, die Leistung der statischen k-Faktoren abzubilden. Im Rahmen des Trainings zeigte sich, dass die Verwendung von DTs bei qualitativen Aussagen helfen kann. Im dynamischen Szenario zeigt sich, dass alle drei Verfahren die Leistung verbessern können. Gleichauf können RL und NN die durchschnittliche Verspätung um 15 % reduzieren und ungeplante Produktmixwechsel kompensieren. Im nächsten Schritt soll eine tiefe und detaillierte Analyse der dynamischen Anpassung über mehrere Produktmixe und unbekannte Szenarien durchgeführt werden.

Literaturverzeichnis

- Heger, J.: Dynamische Regelselektion in der Reihenfolgeplanung. Wiesbaden: Springer Fachmedien Wiesbaden 2014.
- Heger, J.; Branke, J.; Hildebrandt, T.; Scholz-Reiter, B.: Dynamic adjustment of dispatching rule parameters in flow shops with sequence-dependent set-up times. *International Journal of Production Research* 54 (2016) 22, S. 6812–6824.
- Heger, J.; Voss, T.: Dynamically Changing Sequencing Rules with Reinforcement Learning in a Job Shop System with Stochastic Influences. In: Bae, K.-H.; Feng, B.; Kim, S.; Lazarova-Molnar, S.; Zheng, Z.; Roeder, T.; Thiesing, R. (Hrsg.): *Proceedings of the Winter Simulation Conference 2020*, S. 1608–1618.
- Heger, J.; Voss, T.: Dynamically adjusting the k-values of the ATCS rule in a flexible flow shop scenario with reinforcement learning. *International Journal of Production Research* (2021).

- Lee, Y.H.; Jeong, C.S.; Moon, C.: Advanced planning and scheduling with outsourcing in manufacturing supply chain. *Computers & Industrial Engineering* 43 (2002) 1, S. 351–374.
- Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I.: From Local Explanations to Global Understanding with Explainable AI for Trees. *Nature machine intelligence* 2 (2020) 1, S. 56–67.
- Mönch, L.; Zimmermann, J.; Otto, P.: Machine learning techniques for scheduling jobs with incompatible families and unequal ready times on parallel batch machines. *Engineering Applications of Artificial Intelligence* 19 (2006) 3, S. 235–245.
- Mouelhi-Chibani, W.; Pierreval, H.: Training a neural network to select dispatching rules in real time. *Computers & Industrial Engineering* 58 (2010) 2, S. 249–256.
- Nunes, I.; Jannach, D.: A systematic review and taxonomy of explanations in decision support and recommender systems. *User Modeling and User-Adapted Interaction* 27 (2017) 3-5, S. 393–444.
- Panwalkar, S.S.; Iskander, W.: A Survey of Scheduling Rules. *Operations Research* 25 (1977) 1, S. 45–61.
- Priore, P.; La Fuente, D. de; Puente, J.; Parreno, J.: A comparison of machine-learning algorithms for dynamic scheduling of flexible manufacturing systems. *Engineering Applications of Artificial Intelligence* 19 (2006) 3, S. 247–255.
- Priore, P.; Ponte, B.; Puente, J.; Gómez, A.: Learning-based scheduling of flexible manufacturing systems using ensemble methods. *Computers & Industrial Engineering* 126 (2018), S. 282–291.
- Puiutta, E.; Veith, E.M.: *Explainable Reinforcement Learning: A Survey*, 2020.
- Rai, A.: Explainable AI: from black box to glass box. *Journal of the Academy of Marketing Science* 48 (2020) 1, S. 137–141.
- Rehse, J.-R.; Mehdiyev, N.; Fettke, P.: Towards Explainable Process Predictions for Industry 4.0 in the DFKI-Smart-Lego-Factory. *KI - Künstliche Intelligenz* 33 (2019) 2, S. 181–187.
- Shahzad, A.; Mebarki, N.: Learning Dispatching Rules for Scheduling: A Synergistic View Comprising Decision Trees, Tabu Search and Simulation. *Computers* 5 (2016) 1, S. 3.
- Shiue, Y.-R.; Lee, K.-C.; Su, C.-T.: Real-time scheduling for a smart factory using a reinforcement learning approach. *Computers & Industrial Engineering* 125 (2018), S. 604–614.
- Shiue, Y.-R.; Lee, K.-C.; Su, C.-T.: A Reinforcement Learning Approach to Dynamic Scheduling in a Product-Mix Flexibility Environment. *IEEE Access* 8 (2020), S. 106542–106553.
- Usga Cadavid, J.P.; Lamouri, S.; Grabot, B.; Pellerin, R.; Fortin, A.: Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0. *Journal of Intelligent Manufacturing* 31 (2020) 6, S. 1531–1558.
- Vepsalainen, A.P.; Morton, T.E.: Priority rules for job shops with weighted tardiness costs. *Management Science* 33 (1987) 8, S. 1035–1047.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

An Approach for Deep Reinforcement Learning for Production Program Planning in Value Streams

Ein Nutzungsansatz für Deep Reinforcement Learning zur Produktionsprogrammplanung in Wertströmen

Nikolai West, Florian Hoffmann, Lukas Schulte, TU Dortmund, Dortmund
(Germany), nikolai.west@ips.tu-dortmund.de, florian.hoffmann@ips.tu-
dortmund.de, lukas.schulte@ips.tu-dortmund.de

Victor Hernandez Moreno, University of Technology, Sydney (Australia),
victor.hernandezmoreno@student.uts.edu.au

Jochen Deuse, TU Dortmund, Dortmund (Germany) and University of Technology,
Sydney (Australia), jochen.deuse@ips.tu-dortmund.de,

Abstract: The application of Reinforcement Learning (RL) methods offers a potential for improvement in operational Production Program Planning. Numerous influences and domain-specific practices characterize the multi-dimensional planning paradigm. RL can support human planning personnel in the determination of optimal production parameters. This requires a suitable abstraction of the overall system by means of simulation and subsequent optimization by a self-learning agent. In this paper, the authors present an application example for sequence planning using RL. The case study includes a discrete-event simulation built with *SimPy* that is trained by a Dueling Deep-Q-Network implemented in *PyTorch*. Finally, the suitability of two reward functions is discussed. The authors fully provide the case study via GitHub.

1 Introduction

During *Production Program Planning* (PPP), a company must define a market-driven and future-oriented selection of products to manufacture in a certain period according to type and quantity. While in long- and mid-term PPP a general program outline of products is defined, short-term PPP defines specific production quantities and variant sequences. A major goal is to find optimal product sequences and batch sizes for a satisfactory production program. To enable optimal and short-term PPP, it is no longer sufficient under increasing competitive factors to operate purely based on real-world planning data. Due to variability-influenced behaviour, anticipation of a real-world production system is a challenging task that requires in depth domain and system knowledge. Computer-based approaches to support planning production capacities often include more complex solutions, such as event-discrete simulation models or

mathematical replications. *Machine Learning* (ML) represents an increasingly popular approach to achieve targeted results in such systems. In this paper, we focus on the usage of *Reinforcement Learning* (RL), a sub-domain of ML, together with a simulation model to identify beneficial scheduling plans for mixed product lines. Our goal is to develop and compare two different reward functions for demonstration.

2 Fundamentals and related-work

2.1 Short-term production program planning

Within the manufacturing domain, there are numerous tasks to solve with regard to PPP, starting from aggregated PPP, line configuration, followed by PPP and the subsequent sequencing for companies. Line configuration represents the deterministic basis for the new or reconfiguration of a production line with regard to its capacity. PPP carries out the scheduling of customer orders. Both planning problems are subject to a medium-term time horizon, which is why planning processes here have a direct influence for achieving on-time delivery in a production system. Sequence planning, on the other hand, serves a short-term time horizon, whereby corresponding planning errors show rapid effects on production, which can lead to undesirable developments with regard to on-time delivery or temporary capacity requirements (Volling 2009). The efficiency of production is thus significantly influenced by the choice of sequence planning (Dörmer and Günther 2013). As of yet, the choice of a planning approach is highly dependent on the respective target variable selected for optimization. Boysen et al. (2009b) and Gujjula et al. (2011) differentiate existing scheduling approaches according to the three groups of *car sequencing*, *mixed-models* and *level scheduling*. Car sequencing was formulated by Parrello et al. (1986) and, similar to mixed-model sequencing that was formulated by Kilbridge (1963), it focuses on the necessary working time requirements of orders. The level scheduling of the Toyota production system, on the other hand, is designed to achieve a uniform material requirement (Boysen et al. 2009a). Existing approaches can also be assigned with regard to the distinction according to logistical target variables. With regard to on-time delivery, classic methods such as *First-in-Frist-Out*, short FIFO, (Wiendahl 1997) and *minimum residual slack* (Lödding 2016) can be mentioned. On the other hand, methods such as *setup time-optimizing sequencing* (Lödding 2016) and the *extended work in next queue* method (Conway 2012) focus on optimizing the performance target. According to Conway (2012) the *shortest operation time rule*, there are additional possibilities to reduce the inventory, the average lead time, the average delivery delay and to increase the delivery reliability.

Summarizing, PPP can benefit from a multitude of industrial methods that guide and support the determination of optimized sequencing with regard to different target conditions. In addition, numerous methods simplify the management of potential solution spaces. A mayor task is to narrow down the problem in terms of the targeted solution and solve it in an appropriate duration. Due to model abstraction, the risk of simplifying disturbing influences of the solution parameters remains. In this paper, we target the use of automated algorithms based on RL for it offers a promising approach to enable computer-aided sequencing in industrial value streams. RL enables experience-based planning recommendations through computer-based learning without the need for deep subject or domain knowledge. However, while other

approaches such as evolution strategies or genetic algorithms offer promising solutions as well, we deliberately chose to implement a RL model for the case study.

2.2 Reinforcement learning in production program planning

The research area concerning RL and the related *Deep Reinforcement Learning* (DRL) is rapidly evolving. It finds applications in numerous areas, such as the gaming industry, robotics, natural language processing, computer vision or system control (Li 2018). Along with *Supervised Learning* and *Unsupervised Learning*, RL is the third branch of ML, which provides automated methods to detect patterns in data and to use them to fulfil organizational tasks (Bishop 2006). In the domain of ML, RL is concerned with the task of learning how agents ought to take sequences of actions in an environment in order to maximize cumulative rewards (François-Lavet et al. 2018). A network for DRL is characterized by a succession of multiple processing layers, where each layer consists of a non-linear transformation and the sequence of these transformations leads to learning different levels of abstraction (Erhan et al. 2009). According to their learning policy, methods for DRL can be divided into three groups: *Value-based methods*, e.g. Q-Learning or Deep Q-Networks, *policy gradient methods*, e.g. Stochastic Policy Gradient or Actor-Critics-Methods, and *model-based methods*, e.g. purely model-based or hybrid methods (François-Lavet et al. 2018).

Using a virtual replication of the real world, RL helps to find solutions to complex problems, where a problem's solution does not need to be known a priori. Due to its widespread application, there are numerous examples of successful industrial usage. For example, Freitag and Hildebrandt (2016) address multi-objective optimisation in order to improve dispatching rules for production planning and control. Scholz-Reiter and Hamann (2008) optimise the material flow to significantly reduce inventories, lead times and its variation coefficient. Wang and Usher (2005) address the agent-based finding of a dispatching rule for single machine jobs. Zhang et al. (2011) develop a multi-step RL algorithm to solve a scheduling problem with a throughput-based objective. Qu et al. (2016) adapt RL to optimally schedule a multi-step manufacturing system for multiple product types of diverse machines and multi-skilled workforce. Wang (2020) considers a job-shop scheduling problem solved weighted Q-learning algorithm. The selected cases are exemplary, for we do not aim to provide a comprehensive review of RL application. For this, we refer to Shyalika et al. (2020) who provide a detailed review of existing approaches for RL techniques in scheduling tasks. Due to the technical challenge and the economic potential, we derive a need for research that led to the following approach.

3 Reinforcement learning approach

The principle of RL is based on the training of an ML algorithm, hereafter referred to as an agent, that uses constant feedback to interact with an environment (François-Lavet et al. 2018). We display the general principle of RL systems in *Figure 1*. An agent acts independently during the learning phase and it tries to successively optimise its own behaviour by having its actions rewarded or punished by the environment. Starting from a discrete point in time ($t = 0, 1, 2, \dots, n$), the agent is in a observable state s_t . Then, the agent can select an action a_t from predefined set of actions to perform in an environment. Each action may influence the environment. Through this action, the environment enters a new state s_{t+1} at the following time $t + 1$. In case

the state is considered beneficial, the agent receives the reward r_t to reinforce the selected action (Sutton and Barto 1998). The changes of the environment are related to the selected actions of the agent. This allows the training of reward maximizing strategies, also called a policy in order to maximise its total numerical reward (Sim et al. 2003). Through repeated agent-environment interactions, the agent learns a useful strategy that leads to more rewarding states as they are reinforced by the environment.

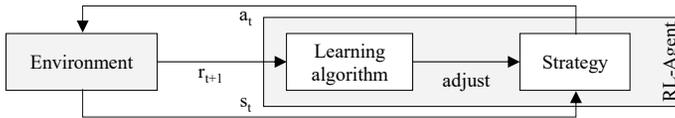


Figure 1: Operating principle of Reinforcement Learning (Sutton and Barto 1998)

For application in operational PPP, the main task of the system is to determine an ideal sequence of products for the PPP. From a production engineering perspective, the main task lies in the correct mapping of the overall system by means of simulation and the definition of state and reward. From a data analytics perspective, the main task is to select an appropriate policy and to specify a function to determine the reward. We explain and demonstrate the approach to DLR in PPP in practice with a case study.

4 Case study

In the following section, we present an exemplary application of DRL in operational PPP using a case study. This section is divided into three parts: First, we present the design of the use case. Second, we explain the structure of the chosen DRL model. Third, we show and discuss the results obtained from two different reward functions.

4.1 Design and implementation of the environment

To demonstrate the potential of DRL in operational PPP, we designed a production system with an implied sequence scheduling trade-off (see Figure 2). The exemplary system manufactures two products A and B on a flow line with five machines M_i . While the process time $pt_{M_i}^A$ for manufacturing product A decreases further down in the value stream, the process time $pt_{M_i}^B$ for manufacturing product B increases. All stations are linked via a buffer b_i , with an additional buffer r_1 with raw material at the start of the line. As the last buffer in the value stream, b_5 acts as the transfer point to the customer. No inventory limits are set for the buffers and the system starts empty. All machines operate according to the FIFO principle. For the software implementation, we use the package SimPy (V. 4.0.1) for Event Discrete Simulation in Python (Matloff 2008). In the code documentation of this paper (West 2021), the simulation of the described factory system is implemented as *factory_simulation.py*.

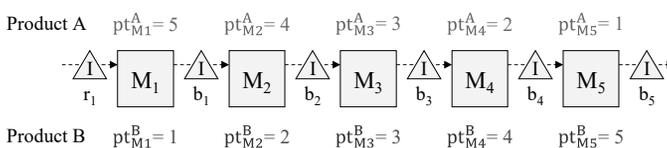


Figure 2: Structural layout of the designed use case for Reinforcement Learning

The chosen study design aims at solving the following planning task: The total output of products A and B has to be maximized for a given time period. If the PPP forces a sequence of one product type, either A or B, we expect an output of one product in five time steps. This is associated with availability losses, as the potential overall output decreases due to the respective bottleneck, either $pt_{M_1}^A$ for a monotonous production of A, or $pt_{M_5}^B$ for the same case of B. Alternating the production of A and B can balance the varying production times, and thus reduce the downtimes of the machines M_i . This strategy allows for the production two products in six time steps. We show the expected ideal state as the overall goal in *Figure 3*. After an initial settling period, we expect an alternating output of product A and B in the optimal PPP case. For manageability, we limit the simulation duration t to a total of 60 time steps.

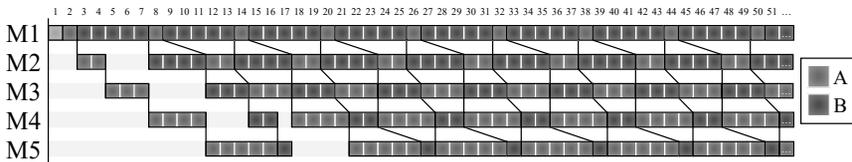


Figure 3: Expected sequence for efficient product sequencing in the use case

4.2 Model selection and implementation for job scheduling

The next task is to select and implement a model. Each of the three groups of DRL provides a potential approach. For the case study, we implement a value-based method, notwithstanding the fact that policy-gradient or model-based methods offer equally appropriate ways to approach the planning task (Beitelspacher et al. 2006). As a model-free approach, value-based techniques allow us to leverage the simulation model without the need to learn an accurate model of the real-world system behaviour.

Due its success in recent years, we select a *Deep-Q-Learning* (DQN) based approach for the case study. *DQN* represents a further development of so-called *Q-Learning* with an extension using deep neural networks and a method referred to as experience replay (Mnih et al. 2015). For our case study, we chose to implement a variant of DQN with a duelling network architecture (DDQN), 48 neurons per layer and rectified linear unit as activation function. We refrain from an in-depth description of the DDQN and instead refer to François-Lavet et al. (2018) for an overview of DQN and to Wang et al. (2015) for an explanation of the duelling architecture. For software implementation, we use the package `PyTorch` (V. 1.8.1) as ML framework in Python. We modelled the interaction between agent and factory according to Allen and Monks (2020), and follow the structure of the established *OpenAI Gym* environments. In the code of this paper (West 2021), the DDQN and the training program for the agent-factory-interaction is implemented as `factory_agent.py`.

As illustrated in *Figure 1*, we need to define an action space \vec{a} , an observation state \vec{s} and a reward function rf for the interaction between the simulation model and the DDQN agent. At each time step of the simulation, the agent has to choose between one of three actions to influence the environment: It either triggers the production of product A, or the production of B. Additionally, it may remain in an idle state and thus manufacture no product. The process runs via virtual release of raw material for the two products A and B for the raw material buffer r_1 . We define $\vec{a} = [0,1,2]$, with

- $a = 0$: The DDQN agent choses the production of product A ('A'),
- $a = 1$: The DDQN agent choses the production of product B ('B'), and
- $a = 2$: The DDQN agent chose to be idle and produce no product ('0').

In principle, all variable parameters in the simulation model are possible observations for \vec{s} , with buffer levels b_i being the most promising from a production engineering point of view. Other influencing variables such as machine states, e.g. running, blocked or starved, or production key figures, e.g. average cycle time or throughput, would also be plausible observations. Since we refrain from influencing the process times pt_{M_i} by a variability factor in the use case, the demand characteristic for semi-finished goods is deterministic. Thus, it is sufficient to consider the initial buffer r_1 in \vec{s} and we avoid additional influences by the buffer levels b_1 to b_5 . To better distinguish between demands for product A and product B, we track the available amount of raw material for A and B in R1 with two separate variables. We define $\vec{s} = [r_1^A, r_1^B]$, with

- $r_1^A = 0$: Amount of raw material for Product A in buffer r_1 (initially set to zero),
- $r_1^B = 0$: Amount of raw material for Product B in buffer r_1 (initially set to zero).

Selecting the action 0 ('A') releases one unit of raw material r_1^A for A in r_1 and enables production. Analogous, action 1 ('B') releases raw material for B in r_1 . Technically, we implemented the separation of material types using sets of dual container in *SimPy*.

In the section 4.3, we use and compare two different reward functions rf_1 and rf_2 . Both functions enable the determination of a time-dependent reward r during the training runs and they operate on the same production cost factors within the environment. We set material cost for a raw material release at 35 *reward units* (RU). We also assume an inventory holding cost of 1 RU per product and buffer. Material costs are applied in the moment the agent chooses to release raw material through the appropriate actions, i.e. 0 or 1. Storage costs are applied once per time step. According to the use case, machine M_5 acts as transfer point to the customer. Thus, after completion of a product, either A or B, rf_1 returns a fixed reward of 300 RU, analogous to a product sale. We build rf_2 to reflect a customer preference for changing product deliveries. The function returns a reward at the same stage, but the value of the reward is based on the prior deliveries. Starting at 200 RU and set in the limit of 100 to 400 RU, rf_2 repeatedly applies a factor of 1,5 for each consecutive switch between products, as well as a factor of 0,75 for each consecutive production of the same product type, either A or B. Both reward functions should be able to identify profitable production sequences, with rf_2 specifically designed to induce the sequencing behaviour that we displayed in *Figure 3*.

4.3 Training results using two reward functions

Next, we show the training results of the DDQN agent in the designed case study after 1000 training runs. We consider the results of both reward functions separately.

4.3.1 Simulation and DDQN agent using reward function rf_1

Figure 4 shows the exploration rate and total reward of 1000 training runs of the DDQN agent in the simulation environment with reward function rf_1 . Each run represents a full simulation over the set training duration of $t = 60$. Following an *Epsilon-greedy* exploration strategy, after 10 random runs, we decrease the exploration rate using a decay of 0,9995 to a minimum rate of 0,01. After initial

fluctuations due to exploration, the overall reward reaches a relatively stable plateau at circa 850 RU. Besides random downward outliers, the reward curve also shows recurrent solutions around 1000 RU. It is therefore apparent that the agent has found a reliable strategy to achieve a positive reward, but it may not be the optimal strategy.

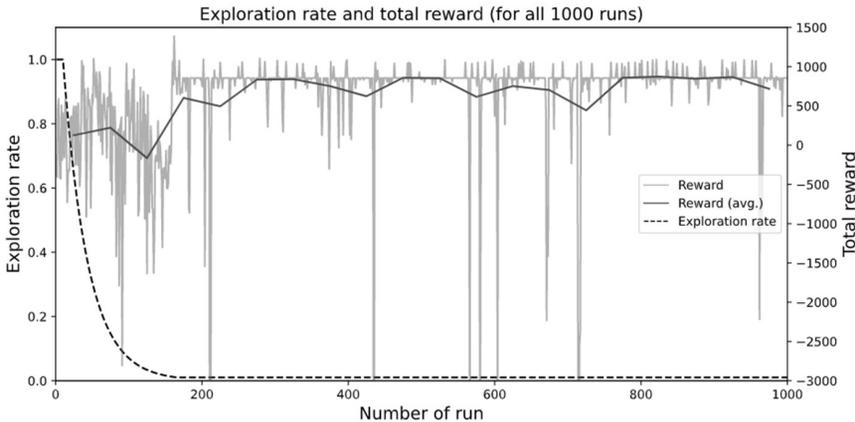


Figure 4: Exploration rate and total reward of run 1000 with reward function rf_2

The plot of an exemplary run after 1000 training runs in Figure 5 confirms this assumption. After initial costs, products A are continuously completed after about 15 time steps. The running costs become minimal the accumulated reward curve takes on a staircase-like shape. The agent uses a strategy of exclusively producing product A, to reliably achieve a positive reward. It has learned the systems output behaviour and the agent thus triggers the production of product A in time intervals that result in minimal inventories while finishing A in the shortest possible succession.

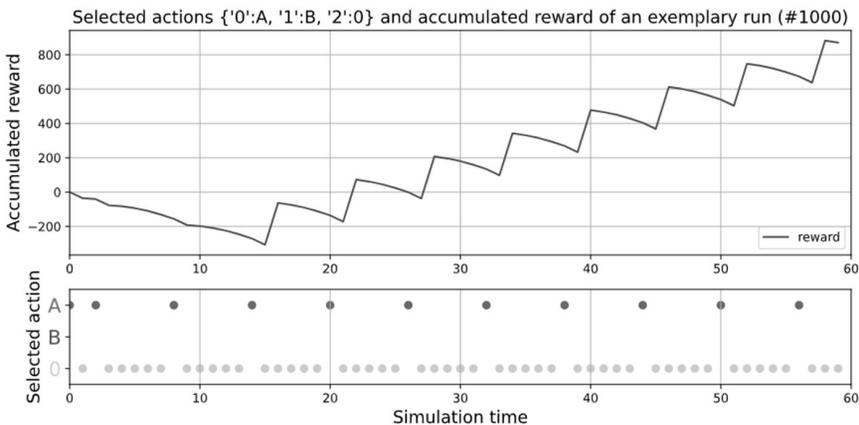


Figure 5: Exemplary results after 1000 training runs, displaying all selected actions and the accumulated reward with reward function rf_1

Due to our preliminary considerations, we know a more optimal solution of the planning task. Furthermore, the upward deviations in Figure 5 indicate better

solutions. For full disclosure, we emphasize that such solutions might also be found using rf_1 by further tuning the training parameters, e.g. by using a lower exploration decay or by increasing the number of training runs. Additionally, from the number of alternative DRL models, some might also provide a more targeted selection of actions.

4.3.2 Simulation and DDQN agent using reward function rf_2

For comparability with rf_1 , we again use the same exploration strategy and present exploration rate and absolute rewards in *Figure 6*. Again, there are initially expected fluctuations in reward due to exploration. However, unlike before, the total reward reaches a stable range much later, about 150 training runs later. Downward outliers also occur much more frequently, whereas there are almost no spikes above a recognizable maximum value of around 2700 RU. A higher average reward for rf_2 was to be expected due to the design of the reward function, however, the achieved threshold indicates the identification globally optimal strategy.

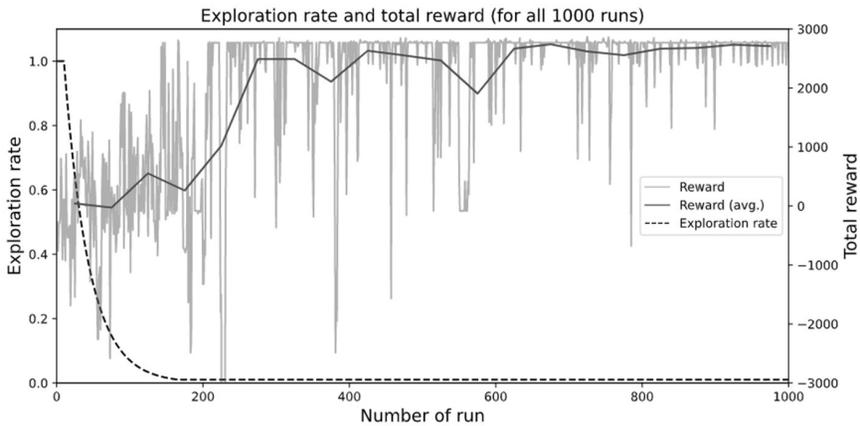


Figure 6: Exploration rate and total reward of run 1000 with reward function rf_2

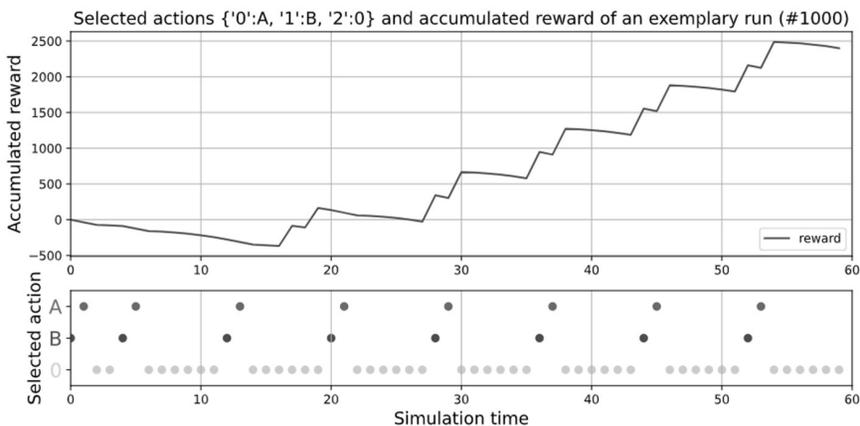


Figure 7: Exemplary results after 1000 training runs, displaying all selected actions and the accumulated reward with reward function rf_2

In *Figure 7*, we show a visualization of an exemplary run after 1000 training steps. The plot of the accumulated reward again shows a phase with initial costs up to time step 15. The subsequent reward course has a double staircase-like structure with two successive ascents after another descending phase. The double increases are due to alternate completion of product B and A. In the consideration of chosen actions, it becomes apparent that the agent finds a seemingly ideal sequence of possible actions. The chosen sequence corresponds to the expectation laid out in *Figure 3*. In addition to the ideal sequence, using rf_2 , the DDQN agent additionally succeeds in releasing orders at cost-maximizing time points.

5 Conclusion and outlook

In this paper, we demonstrated an exemplary approach to use DRL for tasks of PPP. Thereby, this contribution serves to affirm the potential benefits of combining DRL with PPP tasks, as stated at the beginning. From both a scientific and an economic point of view, there is a need for further development. The technical realization of a generally valid reward function for a target-oriented and variability-proof determination of ideal planning states is required. In addition, the approaches that have to date only been realized as theoretical implementation examples should be confirmed in real-world scenarios with economically measurable benefit potential.

In addition to its postulated use as a supporting tool in product sequence planning during PPP, the demonstrated DRL approach offers additional usefulness for advanced control methods that follow an inventory-oriented order release strategy.

References

- Allen, M.; Monks, T. (2020): Integrating Deep Reinforcement Learning Networks with Health System Simulations. In *arXiv:2008.07434*, pp. 1–6.
- Beitelspacher, J.; Fager, J.; Henriques, G.; McGovern, Amy (2006): Policy gradient vs. value function approximation. A reinforcement learning shootout. In *Technical Report of the School of Computer Science, University of Oklahoma* (06-001).
- Bishop, C. (2006): Pattern recognition and machine learning. New York: Springer.
- Boysen, N.; Fliedner, M.; Scholl, A. (2009a): Level scheduling for batched JIT supply. In *Flexible Services and Manufacturing Journal* 21 (1-2), pp. 31–50.
- Boysen, N.; Fliedner, M.; Scholl, A. (2009b): Sequencing mixed-model assembly lines. Survey, classification and model critique. In *EJOR* 192 (2), pp. 349–373.
- Conway, R. W. (2012): Theory of scheduling: Dover Publications.
- Dörmer, J.; Günther, H. (2013): Production program planning with multi-variant flow production. Dissertation. Wiesbaden: Springer Gabler.
- Erhan, D.; Bengio, Y.; Courville, A.; Vincent, P. (2009): Visualizing higher-layer features of a deep network. In *University of Montreal* 1341 (3).
- François-Lavet, V.; Henderson, P.; Islam, R.; Bellemare, M.; Pineau, J. (2018): An introduction to deep reinforcement learning. In *Foundations and Trends in Machine Learning* 11 (3-4), pp. 1–102.
- Freitag, M.; Hildebrandt, T. (2016): Automatic design of scheduling rules for complex manufacturing systems by multi-objective simulation-based optimization. In *CIRP Annals* 65 (1), pp. 433–436.

- Gujjula, R.; Werk, S.; Günther, H.-O. (2011): A heuristic based on Vogel's approximation method for sequencing mixed-model assembly lines. In *International Journal of Production Research* 49 (21), pp. 6451–6468.
- Kilbridge, M. (1963): The assembly line model-mix sequencing problem. In *Proceedings of the International Conference on Operations Research* 3.
- Li, Y. (2018): Deep reinforcement learning. In *arXiv:1810.06339*, pp. 1–150.
- Lödding, H. (2016): Production control method. Basics, description, configuration. 3. Edition. Berlin, Heidelberg: Springer Vieweg.
- Matloff, N. (2008): Introduction to discrete-event simulation and the SimPy language. Davis, CA: UC Davis.
- Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A.; Veness, J.; Bellemare, M. et al. (2015): Human-level control through deep reinforcement learning. In *Nature* 518 (7540), pp. 529–533.
- Parrello, B.; Kabat, W.; Wos, L. (1986): Job-shop scheduling using automated reasoning. A case study of the car-sequencing problem. In *J. o. A. Reasoning* 2 (1).
- Qu, S.; Wang, J.; Govil, S.; Leckie, J. O. (2016): Optimized adaptive scheduling of a manufacturing process system with multi-skill workforce and multiple machine types. In *Procedia CIRP* 57, pp. 55–60.
- Scholz-Reiter, B.; Hamann, T. (2008): The behaviour of learning production control. In *CIRP Annals* 57 (1), pp. 459–462.
- Shyalika, C.; Silva, T.; Karunananda, A. (2020): Reinforcement learning in dynamic task scheduling. A review. In *SN Computer Science* 306 (1), pp. 1–17.
- Sim, S.; Ong, K.; Seet, G. (2003): A foundation for robot learning. In *IEEE International Conference on Control and Automation* 4, pp. 649–653.
- Sutton, R.; Barto, A. (1998): Reinforcement learning. An introduction. Cambridge: MIT Press.
- Volling, T. (2009): Order-related planning for multi-variant series production. Zugl.: Braunschweig, Techn. Univ., Diss., 2008. 1. Edition. Wiesbaden: Gabler (Gabler Edition Science Production and Logistics).
- Wang, Y.-F. (2020): Adaptive job shop scheduling strategy based on weighted Q-learning algorithm. In *Journal of Intelligent Manufacturing* 31 (2), pp. 417–432.
- Wang, Yi-Chi; Usher, John M. (2005): Application of reinforcement learning for agent-based production scheduling. In *Engineering Applications of Artificial Intelligence* 18 (1), pp. 73–82.
- Wang, Z.; Schaul, T.; Hessel, M.; van Hasselt, H.; Lancot, M.; Freitag, N. de (2015): Dueling Network Architectures for Deep Reinforcement Learning. In *arXiv:1511.06581*, pp. 1–15.
- West, N. (2021): Code to Case Study. Available online at github.com/nikolaiwest/2021-reinforcement-learning-asim.
- Wiendahl, H. P. (1997): Production control and Logistical control of manufacturing processes based on the Trichtermodell. München: Hanser.
- Zhang, Zhicong; Zheng, Li; Hou, Forest; Li, Na (2011): Semiconductor final test scheduling with Sarsa(λ, k) algorithm. In *European Journal of Operational Research* 215 (2), pp. 446–458.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Simulative Dispatching Optimization of Maintenance Resources in a Semiconductor Use- Case Using Reinforcement Learning

Simulative Optimierung der Planung von Wartungsressourcen in einem Halbleiter-Anwendungsfall mithilfe von Reinforcement Learning

Clara Hoffmann, Thomas Altenmüller, Infineon Technologies AG, Neubiberg
(Germany), clara.hoffmann@infineon.com, thomas.altenmueller@infineon.com

Marvin Carl May, Andreas Kuhnle, Gisela Lanza, Karlsruher Institut für
Technologie, Karlsruhe (Germany), marvin.may@kit.edu, andreas.kuhnle@kit.edu,
gisela.lanza@kit.edu

Abstract: Semiconductors are high-tech products imposing strong requirements on manufacturing processes. To meet these high quality and accuracy standards, semiconductor manufacturing requires complex and reliable processes on high-tech equipment. Due to the high investment costs for machines and thus associated high machine downtime costs and high process dynamics, challenges arise in handling machine maintenance and finding the optimal maintenance resource dispatching for machines requiring repair or preventive maintenance. The interdependency between production control and maintenance resource dispatching can be modelled using a complex manufacturing system simulation. Deep reinforcement learning is promising in handling the ever-increasing complexity in modern production systems and the associated optimization of maintenance resource dispatching.

1 Introduction

Handling the maintenance of machines cannot be studied appropriately without the context of the underlying manufacturing system. The interdependency between production control and maintenance resource dispatching can be modeled using a simulation of a complex manufacturing system. Reinforcement learning (RL), in particular Deep RL, offers promising opportunities to handle ever-increasing complexities in modern manufacturing systems and optimization of maintenance resource dispatching. Recent progress, e.g. RL beating the best human players in complex strategic games, stress this promising development. Therefore, this study investigates whether the application of RL to the problem of maintenance resource dispatching can improve the performance of semiconductor manufacturing systems.

2 Fundamentals and Literature Review

2.1 Production and Maintenance Planning

Production planning and control focuses on organizing and optimizing the internal processes in a manufacturing system (Eversheim & Wiendahl 2000). Maintenance planning is viewed as a special type of production planning (Ben-Daya et al. 2009). Thus, maintenance planning methods are developed in line with production planning methodologies. However, the two systems differ in various aspects. Maintenance jobs have more variability among themselves and maintenance planning requires the coordination with other functional units. In addition, in semiconductor manufacturing the highly stochastic failure pattern of machines lead to a higher variability of demand for maintenance work than for production.

Production planning aims at increasing machine utilization whereas maintenance planning targets guaranteeing high long-term machine readiness levels. Machine interruptions cause delays in production schedules and vice versa. This leads to conflicts in timing, creating negative interactions between the two functional units. To overcome these conflicts, a coordinated planning process and integrated optimization tools for planning and scheduling is proposed. One solution is to prioritize either function and to use the output plan as an input for the second function, i.e. the input from the other function is then taken as constraint (Al-Turki et al. 2014). This is the case when the maintenance schedule in maintenance planning is adjusted to the given WIP schedule in production planning.

2.2 Maintenance Resource and Task Assignment

In maintenance planning, maintenance resource dispatching can be split up into two categories: (1) maintenance resource assignment, where an available resource pulls work items from a common pool of work items, and (2) task assignment, pushing tasks to resources to queue for handling (Kumar et al. 2002). The decision problems coping with these assignments are either deterministic problems, with a static environment, or stochastic problems, solving assignment problems with tasks arriving in a dynamic, real-time manner. Deterministic problems are mostly solved with classical, mathematical programming techniques, considering different objectives: maximizing a pre-defined score (Kuhn 1995, Martello & Toth 1995), minimizing the job completion time (Arora & Puri 1998, Chauvet, Proth & Soumare 2000), or balancing the resource's workload (Karsu & Azizoglu 2012, Chen et al. 2017), as categorized in Table 1. Deterministic assignment problems can be transformed between categories. In stochastic environments, however, resource assignment is either solved rule-based (Iravani & Krishnamurthy 2007, Arias et al. 2018) or history-based (Liu et al. 2008, Millán-Ruiz & Hidalgo 2010). In comparison, for task assignment mainly rule-based solutions are implemented, either with fixed rules (Stecke & Aronson 1985, Iravani & Krishnamurthy 2007), or adapted rules (Mosley, Teyner & Uzsoy 1998, Langer et al. 2010). Besides, machine learning (ML) is recently used to solve resource and task assignment, using natural language processing (Mo et al. 2020), neural networks (Mao et al. 2016), and RL (Naik, Negi & Sastry 2015). However, so far the focus is on machine reliability and planning or scheduling maintenance service for a single equipment or multiple in a manufacturing system and only limited work was devoted to study the problem in a system-wide

environment, integrating maintenance activities and considering machine and maintainer interaction.

Table 1: Structure of existing literature in this domain

	Deterministic	Stochastic	
		Resource assignment	Task assignment
Without machine learning	Kuhn (1995), Arora & Puri (1998), Martello & Toth (1995), Chauvet, Proth & Soumare (2000), Karsu & Azizoglu (2012), Chen et al. (2017)	Iravani & Krishnamurthy (2007), Arias et al. (2018), Liu et al. (2008), Millán-Ruiz & Hidalgo (2010)	Stecke & Aronson (1985), Mosley, Teyner & Uzsoy (1998), Iravani & Krishnamurthy (2007), Langer et al. (2010)
With machine learning	Mural, Puri & Prabhakaran (2010)	Mo et al. (2020), Mao et al. (2016)	Naik, Negi & Sastry (2015)

2.3 Basics of Reinforcement Learning

RL is a ML algorithm that can solve problems in dynamic environments. A so-called *RL agent* constantly adapts its strategy in a known or unknown environment through receiving feedback. Based on the state of the environment $s_t \in S$ perceived at each time step t , it selects an action $a_t \in A$. The response of the environment includes the resulting state s_{t+1} and a feedback in form of reward r_t . A Markov decision process describes this iteration formally. The ultimate goal of the RL agent is to optimize its strategy to maximize the cumulated, discounted reward. (Altenmüller et al. 2020)

3 Use Case and Reinforcement Learning Modelling

This section introduces the use case oriented on a real-world wafer fab. The literature analysis findings are incorporated in the simulation set-up. The RL modelling regarding state and action space, reward function and the used optimization algorithm is introduced, answering the following research questions:

1. Can a single-agent-based RL approach learn to optimize global key performance indicators through maintenance resource dispatching?
2. Can the proposed RL agent outperform competitive standard heuristics?
3. Does task assignment perform better than resource assignment regarding defined key performance indicators?

3.1 Description of the Wafer-Fab Use Case in the Simulation

A process-based discrete-event simulation, based on Altenmüller et al. (2020), for production planning and control, is adapted for maintenance resource dispatching in order to train an intelligent and autonomous RL agent for decision making in a complex job shop environment with corrective and preventive maintenance, aiming

at outperforming existing approaches. The discrete-event simulation represents the characteristics of typical complex semiconductor-like job shops.

The two approaches, resource and task assignment, are implemented in two independent views, the machine and maintenance view, as shown in Figure 1. The main difference between the two views is, that the maintainer view waiting time results from waiting in the maintenance machine shop, in contrast to the machine view, where the waiting time results from waiting inside the maintainer skill queue.

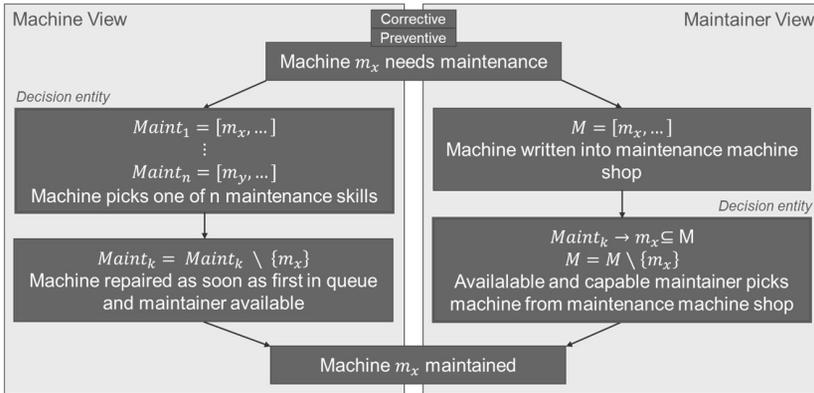


Figure 1: Process structure of the machine view compared to the maintainer view

The considered use case is based on the eWLB line (Altenmüller et al. 2020) and represents a production process with 40 specialized machines organized in five machine groups. Every machine group has a buffer stock with a fixed number of buffer slots. Challenges arise from machines requesting corrective or preventive maintenance which need to be handled by given, sparse maintenance resources. Hereby, the negative influence on the cycle time (CT) of the products being processed within the system shall be reduced to increase speed and lower costs. Furthermore, order flow is re-entrant in earlier visited machine groups. Finally, machines require a product-specific set-up. For the resulting product flow of the used, simplified eWLB line and the parameters used see Altenmüller et al. (2020).

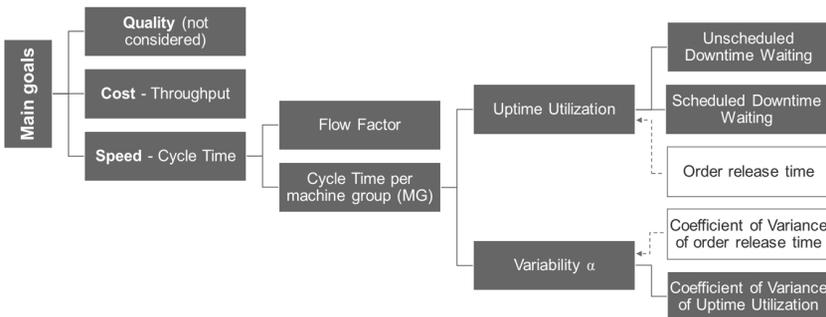


Figure 2: Hierarchy of the tracked KPIs in the simulation

3.2 Performance Evaluation

Key performance indicators (KPIs), oriented at SEMI E10-0312 (2012), are used for performance evaluation and, thus, the fulfillment of target performances. The KPIs evaluated in this paper are hierarchically ordered as displayed in Figure 2, showing which KPIs on the lower level (right), influence KPIs on the upper level (left). The overall goals are derived from the overall equipment performance (OEE).

3.3 State Space Modelling

The basis for the decision-making by the RL agent is the information in the observation space passed to it about the current status of the system, shown in Table 2. Resource-related information contain information about the decision requesting entity whereas maintenance-related information is made up of information about the decision to take as an action. Production-related information are more global, covering the state of the final products in the exit buffer.

Table 2: Observation space for the RL agent in the machine and maintainer view

Information	Machine view	Maintainer view
Resource-related	Active machine	Active maintainer
	Machine group	Maintainer skills
Maintenance-related	Capable maintainer	Possible machines
	Maintainer shift count	Possible machines for scheduled maintenance
	Mean time to repair (MTTR) in queue	Waiting time per machine
Production-related		Waiting time of orders in stock
		Processing time of orders in stock

3.4 Action Space Modelling

The action space is discrete, as the RL agent selects a distinct action from a finite action set that depends on the view. The action space of the machine view is smaller, as it contains seven options, one for each maintenance skill, whereas the maintainer view is made up of 40 options, one for each machine. Selecting a maintainer skill not capable of repairing the requesting machine in the machine view is considered as an invalid action, as well as selecting a machine not requesting maintenance in the maintainer view. Invalid actions are not executed (i.e. the simulation state is not changed) but used as feedback for learning: the agent is penalized with a negative reward and requested to select another action.

3.5 Reward Function Modelling

The reward function r is the key element that leads to the RL agent learning a desired behavior. As introduced above, invalid actions are in general rewarded with $r = -1$. A distinction is made between local and global rewards. Local rewards consider the current situation at the machine respectively maintainer requesting a decision, whereas global rewards regard the situation in the exit buffer for finished orders for computing the reward, as a proxy for the entire manufacturing system. The local

reward in the machine view is based on a ranking of the length of the sum of the MTTR of machines waiting in the queue of the chosen maintainer skill normalized with the maintainer skill available. In the maintainer view, the ranking is built over the waiting times of the machines requesting maintenance.

The global reward targets the highest through maintenance resource dispatching manipulable level of the hierarchy of KPIs, namely the overall cycle time. For each selected valid action a , the RL agent is rewarded based on the cycle time of orders in the exit buffer. As single finished orders in the exit buffer cannot be linked to a certain maintenance resource dispatching decision taken, the cycle time of the last n orders is set in ratio to the cycle time of $n * 3$ last orders. This is done by z-score normalization, where the value is then inserted in an exponential function such that the RL agent receives exponentially more reward for more desirable behavior than for bad behavior (Bonsai 2017). The resulting episodic global CT reward is made up of a sparse, global reward and a modeled reward, focusing on (in)valid actions.

3.6 Reinforcement Learning Algorithm

A Proximal Policy Optimization (PPO) agent provided by the library Stable Baselines is used as it outperformed Deep Q-Network and Trust Region Policy Optimization agents in several earlier investigations. In general, PPO is a model-free, online, on-policy, policy gradient reinforcement learning method and thus a type of policy gradient training that alternates between sampling data through environmental interaction and optimizing a clipped surrogate objective function with stochastic gradient descent. The clipped surrogate objective function improves training stability by limiting the policy change per step (Schulman et al. 2017).

3.7 Benchmark Heuristics

The benchmark heuristics are compared to the performance of the RL agent in the two views. In machine view, two heuristics are developed. First, the valid heuristic picks maintenance resources that have the skill to handle the machine randomly. Second, the local heuristic picks the maintenance skill with the shortest normalized waiting queue, such that the queues of the maintenance skills are smoothed according to the availability and capability of the maintenance resources, resulting in a smoothing of waiting times between the machines. In the maintainer view, the first in first out (FIFO), longest queue (LQ), longest queue shortest repair time (LQSRT), most wafer (MW) and least remaining capacity (LRC) heuristics from Mosley, Teyner & Uzsoy (1998) are adapted. The difference between these heuristics lies in the way of choosing machines by maintenance resources to maintain next.

4 Results

In this chapter, the results of the best performing benchmark heuristic are compared with results of the above explained local and global reward in both views.

4.1 Behavior of the Reinforcement Learning Agent

The agent's learning already converges after a relatively small number of steps. This is shown through the truncation point marked in grey in Figure 3 on the right, evaluating the development of the CT on the left, through using the MSER-5

algorithm, an efficient and effective truncation heuristic. Thus, a stopping criterion is implemented after the learning converged, disrupting the learning of the agent, at a point where the policy shows a good performance.

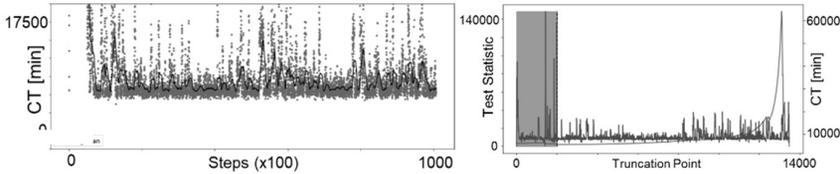


Figure 3: Instability of the CT during learning and Convergence with MSER-5

4.2 Comparison of Benchmark Heuristics and Reinforcement Learning Agent

Table 3: Mean KPI values for machine and maintainer view (high machine failures)

	TP	CT	CT MG0	CT MG1	CT MG2	CT MG3	CT MG4	UDW	SDW
Machine view									
	Heuristic – local								
	3.79	12700	179	195	4633	501	690	26.27%	4.64%
	Local reward								
	3.72	12744	180	193	4631	495	721	26.45%	4.55%
	Episodic global CT reward								
	3.18	7206	3126	137	375	283	505	33.51%	4.63%
Maintainer view									
	Heuristic – LQ								
	4.36	7748	1887	294	797	700	704	28.95%	1.87%
	Local reward								
	3.76	11995	170	242	4594	412	419	23.07%	5.94%
	Episodic global CT reward								
	4.09	6146	2426	167	381	291	357	22.05%	3.83%

Table 3 presents the results of both the machine and maintainer view with comparably high machine failures. Generally, minimizing the waiting times through local heuristics and rewards for maintenance actions does not lead to the shortest CT, as 90% of the machines are not bottleneck machines, such that waiting for maintenance has a small influence on the total CT. The CT is further reduced using the RL agent with episodic global CT reward compared to the benchmark heuristics. Through the learnt RL policy, the bottleneck is shifted to the first machine group. Thus, the order release into the production system is regulated by high waiting times for maintenance in the first machine group and hence a reduction of availability of machines in the first machine group, also referred to as trumpet planning. The maintainer resources are mainly used to maintain downstream machines. This leads to a high availability of these machines such that orders pass downstream machines as soon as they are released from the first machine group. However, the reduction of the CT happens at the expense of the Throughput (scheduling dilemma). The main mechanism is the

implemented simulation logic, reducing order release into the system as soon as buffer and overflow buffer are overloaded. The RL agent utilizes this by filling the buffers in front of the first machine group by disproportionately increased waiting times for maintenance, referred to as reward hacking.

Comparing these results to results with realistic, smaller machine failures in Figure 4 shows, that the throughput is not influenced negatively by the CT reduction, as the agent cannot influence the order release to the previous extent. However, in the maintainer view, the CT performance through using the episodic global CT reward is not further improved in comparison to the benchmark heuristic.

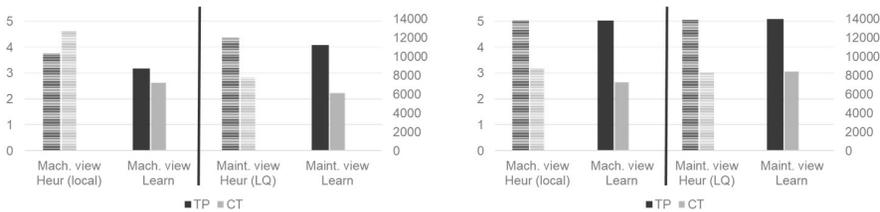


Figure 4: Result comparison for large (right) and small (left) machine failure rates

4.3 Discussion

The observation that for higher time to failures (TTFs) the PPO agent in the machine view is performing worse regarding the CT than in the maintainer view can be explained with the different action space size. More differentiated decision-making is possible in the maintainer view due to the bigger action space size. Furthermore, in the machine view, a queue of machines requests in the respective maintenance skill is built up. As the queue is quite long and based on the MTTR of the respective machines, the decision is fraught with uncertainty and not changeable after the initial assignment. In addition, the effect of the decision is more delayed and uncorrelated than within the maintainer view, where the decision is immediately executed.

In general, the large action space is reduced to valid machines, respectively those requesting maintenance. As the total number of machines requesting maintenance is reduced in comparison to lower TTF, the decision options are reduced significantly, such that a differentiated decision might have less influence on the request waiting times. Thus, ad hoc decisions in the maintainer view lack the benefits of planning incorporated in the machine view by filling queues of maintenance resources due to a more short-term view, in a similar setup shown by Millán-Ruiz & Hidalgo (2010).

The PPO agent with episodic global CT reward within the machine view could be used to build up a decision-making according to 'safe AI', which are in general methods working towards a higher safety of AI. The decisions then are short term, with the decision being executed not immediately. Through this, a human control of the resulting maintenance resource dispatching from the PPO agent is possible, increasing the safety aspect and thus leading to a safer AI application in practice.

5 Summary

Semiconductor manufacturing in times of pervasive digitalization opens up new opportunities and challenges, leading to a new era of operations management. The use

of RL enhances intelligent and autonomous control approaches. This paper contributes to research that concerns the performance of maintenance resource dispatching in the context of a complex flexible job shop. The paper presents RL utilized to regulate maintenance resource dispatching to achieve CT minimization.

The results of multiple simulation runs show that different RL agents can autonomously develop a policy for solving the maintenance resource dispatching problem. The resulting KPIs, especially the target CT performance of the best PPO agent is superior to the underlying benchmark. If machines request maintenance less often, short-term planning through assigning machines to maintenance resources is outperforming maintenance resource dispatching with ad hoc decisions regarding the CT performance, whereas for a larger number of requests ad hoc decisions excel.

6 Outlook

Based on this results, further research should be conducted on using Deep RL for maintenance resource dispatching, i.e. focusing on further improving the performance of Deep RL agents. Instead of using a single decision-making RL agent, multiple RL agents, respectively a multi-agent system for dispatching, maintenance scheduling and maintenance resource dispatching should be considered. The implemented system architecture offers the potential to conduct more detailed analyses by extending the maintainer skill matrix through further splitting up skills in preventive and corrective skills, as well as incorporate re-prioritization and thus maintenance action disruptions. The parameters for setting up the simulation are oriented at realistic data but adapted and simplified. Through process mining, gaining realistic simulation input data is possible. Thus, simulating production lines in factories realistically becomes possible.

References

- Altenmüller, T., Stüker, T., Waschneck, B., Kuhnle, A., & Lanza, G. (2020). Reinforcement learning for an intelligent and autonomous production control of complex job-shops under time constraints. *Production Engineering*, 14, 319-328.
- Al-Turki, U. M., Ayar, T., Yilbas, B. S. & Sahin, A. Z. (2014), *Integrated maintenance planning in manufacturing systems*, Springer, New York.
- Arias, M., Munoz-Gama, J., Sepúlveda, M., & Miranda, J. C. (2018). Human resource allocation or recommendation based on multi-factor criteria in on-demand and batch scenarios. *European Journal of Industrial Engineering*, 12(3), 364-404.
- Arora, S., & Puri, M. C. (1998). A variant of time minimizing assignment problem. *European Journal of Operational Research*, 110(2), 314-325.
- Ben-Daya, M., Duffuaa, S. O., Raouf, A., Knezevic, J. & Ait-Kadi, D. (2009), *Handbook of maintenance management and engineering*, vol. 7, Springer.
- Bonsai (2017), *Deep Reinforcement Learning Models: Tips and Tricks for Writing Reward Functions*. <https://medium.com/@BonsaiAI/deep-reinforcement-learning-models-tipstricks-for-writing-reward-functions-a84fe525e8e0>.
- Chauvet, F., Proth, J. M., & Soumare, A. (2000). The simple and multiple job assignment problems. *International journal of production research*, 38(14), 3165-3179.
- Chen, G., He, W., Leung, L. C., Lan, T., & Han, Y. (2017). Assigning licenced technicians to maintenance tasks at aircraft maintenance base: a bi-objective

- approach and a Chinese airline application. *International Journal of Production Research*, 55(19), 5550-5563.
- Eversheim W., Wiendahl HP. (eds) (2000). *Wörterbuch der PPS—Dictionary of PPC: German–English/English–German*. Springer, Berlin
- Iravani, S. M., & Krishnamurthy, V. (2007). Workforce agility in repair and maintenance environments. *Manufacturing & Service Operations Management*, 9(2), 168-184.
- Karsu, Ö., & Azizoğlu, M. (2012). The multi-resource agent bottleneck generalised assignment problem. *International Journal of Production Research*, 50(2), 309-324.
- Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2), 83-97.
- Kumar, A., Van Der Aalst, W. M., & Verbeek, E. M. (2002). Dynamic work distribution in workflow management systems: How to balance quality and performance. *Journal of Management Information Systems*, 18(3), 157-193.
- Langer, R., Li, J., Biller, S., Chang, Q., Huang, N., & Xiao, G. (2010). Simulation study of a bottleneck-based dispatching policy for a maintenance workforce. *International Journal of Production Research*, 48(6), 1745-1763.
- Liu, Y., Wang, J., Yang, Y., & Sun, J. (2008). A semi-automatic approach for workflow staff assignment. *Computers in Industry*, 59(5), 463-476.
- Mao, H., Alizadeh, M., Menache, I., & Kandula, S. (2016, November). Resource management with deep reinforcement learning. In *Proceedings of the 15th ACM workshop on hot topics in networks* (pp. 50-56).
- Martello, S., & Toth, P. (1995). The bottleneck generalized assignment problem. *European journal of operational research*, 83(3), 621-638.
- Millán-Ruiz, D., & Hidalgo, J. I. (2010, April). A memetic algorithm for workforce distribution in dynamic multi-skill call centres. In *European Conference on Evolutionary Computation in Combinatorial Optimization* (pp. 178-189). Springer, Berlin, Heidelberg.
- Mo, Y., Zhao, D., Du, J., Syal, M., Aziz, A., & Li, H. (2020). Automated staff assignment for building maintenance using natural language processing. *Automation in Construction*, 113, 103150.
- Mosley, S. A., Teyner, T., & Uzsoy, R. M. (1998). Maintenance scheduling and staffing policies in a wafer fabrication facility. *IEEE Transactions on Semiconductor Manufacturing*, 11(2), 316-323.
- Mural, R. V., Puri, A. B., & Prabhakaran, G. (2010). Artificial neural networks based predictive model for worker assignment into virtual cells. *International Journal of Engineering, Science and Technology*, 2(1), 163-174.
- Naik, N. S., Negi, A., & Sastry, V. N. (2015). Performance improvement of MapReduce framework in heterogeneous context using reinforcement learning. *Procedia Computer Science*, 50, 169-175.
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A. & Klimov, O. (2017), „Proximal policy optimization algorithms“, in: arXiv preprint arXiv:1707.06347.
- SEMI E10-0312 (2012), Specification for definition and measurement of equipment reliability, availability, and maintainability (RAM) and utilization.
- Stecke, K. E., & Aronson, J. E. (1985). Review of operator/machine interference models. *International Journal of Production Research*, 23(1), 129-151.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Automatisierung im Prozess der Wissensentdeckung in Simulationsdaten – Charakterisierung der Ergebnisdaten

Automation within the process of knowledge discovery in simulation data – Characterization of the result data

Jonas Genath, Sören Bergmann, Niclas Feldkamp, Steffen Straßburger,
TU Ilmenau, Ilmenau (Germany),
jonas.genath@tu-ilmenau.de, soeren.bergmann@tu-ilmenau.de,
niclas.feldkamp@tu-ilmenau.de, steffen.strassburger@tu-ilmenau.de

Abstract: The traditional application of simulation in production and logistics is usually aimed at changing certain parameters in order to answer clearly defined objectives or questions. In contrast to this approach, the method of knowledge discovery in simulation data (KDS) uses a simulation model as a data generator (data farming). Subsequently using data mining methods, hidden, previously unknown and potentially useful cause-effect relationships can be uncovered. So far, however, there is a lack of guidelines and automatization-tools for non-experts or novices in KDS, which leads to a more difficult use in industrial applications and prevents a broader utilization. This paper presents a concept for automating the first step of the KDS, which is the process of characterization of the result data, using meta learning and validates it on small case study.

1 Einleitung

Die traditionelle Anwendung der Simulation in Produktion und Logistik zielt meist auf die Veränderung bestimmter Parameter ab, um im Vorfeld klar definierte Zielstellungen bzw. Fragestellungen zu beantworten (März et al. 2011; Law 2013). Die Richtlinie 3633 des VDI nennt als typische zu bearbeitende Zielstellungen beispielsweise das Finden von Kapazitätsengpässen und die Bewertung von unterschiedlichen Maschinenbelegungen in einem Prozess. Bei diesem Vorgehen hängen die Ergebnisse stark von dem Expertenwissen der für die Durchführung der Simulationsstudie verantwortlichen Personen ab, da diese auf der Basis von eigener Erfahrung die Spanne und die zu variierenden Parameter festlegen. Die resultierenden Simulationsergebnisse werden anschließend analysiert und es können entsprechende Rückschlüsse über Maßnahmen für das reale System erarbeitet

werden. Obwohl diese Vorgehensweise zur Beantwortung von im Vorhinein festgelegten Fragestellungen als Best-Practice-Ansatz gilt, könnte, da der abgedeckte Modellraum nur sehr gering ist, über das grundsätzliche Modellverhalten nur wenig Wissen generiert werden (Feldkamp et al. 2015). Sind die Zielstellungen der Simulationsstudie weniger konkret, wie z. B. „Welche Parameter besitzen den größten Einfluss auf eine Kenngröße des Systems?“ oder besteht das Interesse an einem verbesserten Gesamtverständnis über das Systemverhalten sowie den Beziehungen zwischen Parametern, muss bereits im Experimentdesign für einen umfassenderen und besser abgedeckten Modellraum gesorgt werden (Law 2013; Kleijnen et al. 2005). Aus dieser Idee, die Simulation für mehr als nur die Beantwortung bestimmter Fragestellungen eines Systems zu nutzen, ergibt sich ein neues Anwendungsgebiet: Das von Brandstein und Horne für die Gefechtssimulation entwickelte Data Farming (Brandstein und Horne 1998).

Im Data Farming wird das Simulationsmodell als Datengenerator genutzt, indem im Experimentdesign ein umfangreicher Wertebereich für die Eingangsparameter gewählt und dadurch der möglichst vollständige Ergebnisraum des Modells abgedeckt wird (Sanchez 2014). Hiermit können zur Unterstützung der Entscheidungsfindung positive und negative Wirkzusammenhänge entdeckt werden, welche unter Betrachtung eines spärlich abgedeckten Modellraums verborgen geblieben sind (Horne und Meyer 2004). Unter anderem erkannten Feldkamp et al. (2017) die ungenutzten Analysemöglichkeiten des Data Farming für die Simulation in der Produktion und Logistik und entwickelten das Konzept der Wissensentdeckung in Simulationsdaten.

Dieser Beitrag widmet sich der Forschungsfrage, inwieweit die Charakterisierung der Ergebnisdaten in der Wissensentdeckung in Simulationsdaten unterstützt bzw. automatisiert werden kann. Der Schwerpunkt der Analyse liegt hier auf der Filterung von relevanten Ergebnisparametern, welche zu finden für den unerfahrenen Anwender entweder hohen Aufwand bedeuten oder ihm sogar möglicherweise verborgen bleiben. In diesem Kontext werden unter relevanten Ergebnisparametern, die verstanden, welche statistische Auffälligkeiten besitzen und dadurch im Verlauf anschließender Analysen zur einem Informationsgewinn über das Modell führen können.

Das Ziel ist, dass auch Anwender, welche mit dem Prozess der Wissensentdeckung nicht vertraut sind, oder erfahrene Anwender bei dem Erkennen von relevanten Ergebnisparametern einer Data-Farming-Studie unterstützt werden und diese somit für weitere Analyseschritte nutzen können.

Der Aufbau des Beitrags gliedert sich entsprechend wie folgt: Zuerst werden die theoretischen Grundlagen des Prozesses der Wissensentdeckung in Simulationsdaten und der Methode des Meta Learning vorgestellt. Im Anschluss folgt das konzeptionelle Kapitel, welches das vorgeschlagene Konzept zur Automatisierung der Charakterisierung der Ergebnisdaten vorstellt und sich mit der Entwicklung auseinandersetzt. Abschließend werden die Ergebnisse der Konzeptentwicklung anhand eines Realweltszenarios evaluiert. Der Beitrag wird mit einem Fazit und einem Ausblick abgeschlossen.

2 Grundlagen der behandelten Methodiken

In den folgenden zwei Unterkapiteln werden die Grundlagen der in diesem Beitrag verwendeten Methoden vorgestellt, um dem Leser ein tieferes Verständnis über das vorgeschlagene Konzept zu ermöglichen.

2.1 Wissensentdeckung in Simulationsdaten

Die Wissensentdeckung in Simulationsdaten hat ihren Ursprung im Prozess der Knowledge Discovery in Databases zur Auswertung von großen Datenmengen nach Fayyad et al. (1996). Feldkamp et al. (2017) adaptierten diesen Prozess an das Data Farming, da sich die generierten Daten auf Grund ihres Umfangs nicht mehr manuell auswerten lassen. Die in der Wissensentdeckung in Simulationsdaten zur Datenanalyse verwendeten Methoden stammen aus dem Spektrum des Data Mining und wurden bisher im Kontext von Simulationsdaten nur wenig genutzt. Die anschließende Ergebnis- und Datenrepräsentation erfolgt im Sinne des eng mit dem Data Mining verzahnten Konzepts der Visual Analytics. Unter Visual Analytics wird die interaktive Visualisierung von großen und komplexen Datensätzen zur Unterstützung der Mustererkennung sowie Entscheidungsfindung verstanden (Keim et al. 2008). Die Wissensentdeckung in Simulationsdaten lässt sich in drei Prozessschritte aufteilen. Bei diesen handelt es sich um die Charakterisierung der Ergebnisdaten, die Mustererkennung und Klassenbildung sowie die Untersuchung der Beziehung zwischen Eingangs- und Ergebnisdaten.

Die Hauptaufgabe in der Charakterisierung der Ergebnisdaten besteht in der Identifizierung von interessanten Ergebnisparametern. Diese bilden entweder die Basis oder die Eingangsparameter für die darauffolgenden Data-Mining-Verfahren, welche für die Suche nach Wissen über das Modellverhalten genutzt werden. Das Vorgehen erfolgt aktuell manuell und Anwender müssen die statistischen Lagemaße sowie dessen Visualisierungen individuell auswerten, um die relevanten Ergebnisparameter zu identifizieren. Bei der im Data Farming üblichen großen Menge an Ergebnisparametern kann diese Aufgabe zu einem hohen zeitlichen Aufwand führen. Dieser Beitrag konzentriert sich daher auf die Automatisierung dieses Vorgehens. Eine Betrachtung der anschließenden Prozessschritte in der Wissensentdeckung in Simulationsdaten ist aktuelles Forschungsgebiet.

2.2 Meta Learning

Damit eine automatisierte Identifikation der statistisch auffälligen Ergebnisparameter erreicht werden kann, wird die Methodik des Meta Learning vorgeschlagen. Für das Meta Learning gibt es in der Wissenschaft noch keine einheitliche Definition. Dieser Beitrag orientiert sich daher an der Interpretation nach Brazdil et al. (2009), die Meta Learning als Methode definiert, welche Domänenwissen und Metafeatures benutzt, um die Effizienz von Prozessen mittels der Adaption von Machine-Learning-Modellen oder Data-Mining-Verfahren zu steigern. Bei Metafeatures handelt es sich um Charakteristika eines Datensatzes, Algorithmus oder Machine-Learning-Modells (Brazdil et al. 2009). Das Domänenwissen wird für das Meta Learning mittels spezifischer definierter Kennzahlen repräsentiert. Beide Aspekte lassen sich unter dem Begriff Metadaten zusammenfassen. Grundlegend kann Meta Learning als das Anwenden von Machine Learning zur Verbesserung von Machine-Learning-Modellen beschrieben werden.

Im Bereich der prognostizierenden Modelle oder Hyperparameteroptimierung wurde das Konzept des Meta Learning bereits mit guter Erfolgsquote umgesetzt (Ren et al. 2018). Diese Anwendungen behandeln jedoch nur akademische Klassifikations- und Regressionsprobleme, bei denen es eindeutige Ergebnisse gibt. Bei dem vorgestellten Prozessschritt handelt es sich jedoch um einen komplexen Ablauf mit einem starken explorativen Charakter, welcher von dem Simulationsmodell und dem Experimentdesign abhängig ist. Daher sind die Ergebnisse nicht immer eindeutig „richtig oder „falsch“.

Das typische Vorgehen bei der Entwicklung eines Meta-Learning-Modells lässt sich in drei Abschnitte einteilen. Dabei handelt es sich als Erstes um das Extrahieren der Metadaten. Dieser Schritt besitzt eine hohe Relevanz, da nur mit einer geeigneten Repräsentation und Auswahl an Metadaten das Meta-Learning-System in der Lage ist, sinnvolle und hilfreiche Unterstützung zu liefern. Vanschoren (2011) und Brazdil et al. (2009) empfehlen für die Extraktion die Entwicklung eines Metadaten-Generators. Der nächste Schritt ist die Auswahl eines geeigneten Algorithmus für das Meta Learning. Grundsätzlich sind die meisten traditionellen Machine-Learning-Algorithmen auch für die Verwendung als Meta Learner geeignet. Für die verfügbaren Trainingsdaten des Meta-Learning-Algorithmus empfiehlt Lemke et al. (2015) eine erweiterbaren Metadatenbank. Die Erweiterbarkeit ist dahingehend wichtig, dass ein kontinuierlicher Verbesserungsprozess der Prädikationen des trainierten Modells realisiert werden kann.

3 Konzeptentwicklung

In dem folgenden Abschnitt werden das Konzept und dessen Entwicklungsschritte zur Lösung der genannten Probleme erläutert. Das Konzept wurde nach den drei von Lemke et al. (2015) empfohlenen Abschnitten erarbeitet und ist in Abbildung 1 dargestellt. Dabei wird das Konzept zusätzlich in zwei Phasen aufgeteilt. In der Offline-Phase werden aus den Ergebnissen bereits durchgeführter Data-Farming-Studien mittels Metadaten-Generators die festgelegten Metadaten extrahiert. Diese dienen als Trainingsdatensatz für das Meta-Learning-Modell, welches anschließend der Online-Phase zur Verfügung gestellt wird. In der Online-Phase werden neue Data-Farming-Studien ausgewertet. Zuerst extrahiert der Metadaten-Generator die definierten Metadaten aus den Simulationsdaten und stellt diese dem trainierten Modell zur Verfügung. Nur die simulationspezifischen Metadaten müssen von dem Anwender mittels einer einfachen Maske abgefragt werden. Mit Hilfe der extrahierten Metadaten inferiert das trainierte Modell eine Rangliste über die prognostizierte Relevanz der Ergebnisparameter. Diese liefert dem Anwender Informationen über mögliche statistische Auffälligkeiten und eine Basis für die folgenden Data-Mining-Verfahren.

Die drei Abschnitte der Konzeptentwicklung werden in den folgenden Ausführungen detailliert erläutert.

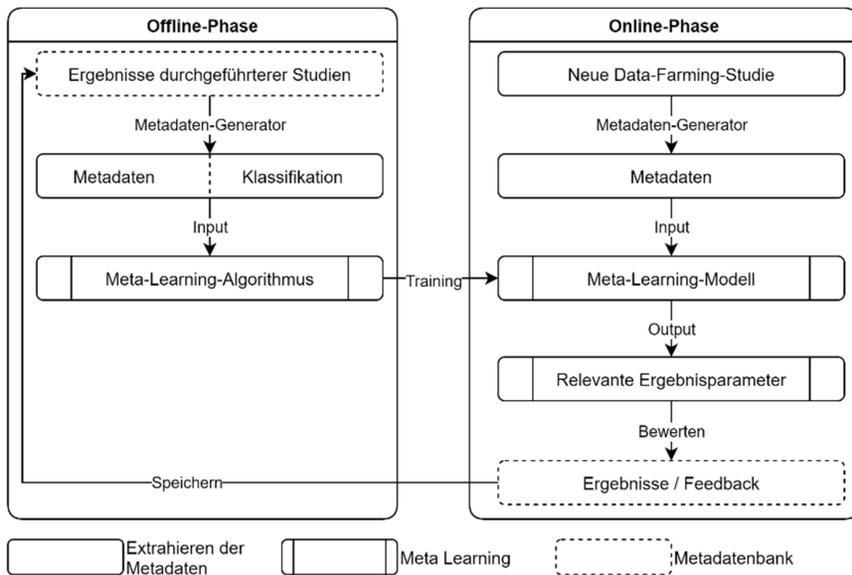


Abbildung 1: Konzeptionelles Vorgehen zur automatischen Charakterisierung der Ergebnisdaten.

Extrahieren der Metadaten

Die Schwierigkeit beim Meta Learning besteht in der Definition der beschreibenden Domäneninformationen und wie diese dem für das Meta Learning verwendeten Machine-Learning-Modell zugänglich gemacht werden. Hierbei gilt es vor allem zu beachten, dass ein Machine-Learning-Modell keine Informationen darüber hat, welchen Kontext der Ergebnisparameter beschreibt. Beispielsweise wird eine hohe Auslastung einer Maschine meist als positiv erachtet aber die Rüstzeiten in einem System sollten eher niedrig ausfallen. Ähnliches gilt für die Unterscheidung zwischen einer zeitlichen und mengenbezogenen Kennzahl. Einzig anhand des numerischen Wertes ist ein Machine-Learning-Modell zu dieser Interpretation nicht in der Lage.

Rivoli et al. (2018) definieren über 100 standardisierte Metafeatures und teilen diese in sechs Gruppen ein. Im Kontext des Ziels und der verfügbaren Daten eignen sich hierfür nur die Metafeatures der statistischen und allgemeinen Gruppierung. Dazu zählen unter anderem verschiedene Quartile, Min- bzw. Max-Werte und die statistische Schiefe. Die domänenspezifischen Informationen müssen je nach Anwendungsfall experimentell bestimmt werden. Hierfür wurde eine Reihe von Experimenten mit simulationsspezifischen Informationen durchgeführt. Die besten Resultate wurden mit Metadaten über den oben beschriebenen Kontext der Kennzahlen erreicht. Zu diesem Zweck wurde eine Kategorisierung der Kennzahlen nach der Richtlinie 3633 des VDI vorgenommen. Die Einteilung ist über den Typ (Zeit, Menge, Quotient, Prozent) sowie die Gruppierung (Produktivität, Nutzungsgrad, Systemlastentitäten) des Parameters definiert. In den Eingangsdaten für das Modell werden diese Eigenschaften über binäre Einträge repräsentiert. Die Metadaten werden dem Meta-Learning-Algorithmus in Tabellenform übergeben. Tabelle 1 liefert einen beispielhaften Ausschnitt der bereitgestellten Daten.

Table 1: Ausschnitt ausgewählter Metadaten.

ID	Ergebnisdaten	Min	Max	Mean	Menge	Zeit	Systemlastentität	Nutzungsgrad
1	Durchlaufzeit	12,3	22	17,35	0	1	1	0
2	Auslastung	0,78	0,99	0,80	0	0	0	1
3	Rüstvorgänge	5	12	6	1	0	0	1

Meta Learning

Der verwendete Meta-Learning-Algorithmus muss verschiedene Anforderungen erfüllen, um für den vorliegenden Kontext geeignet zu sein. Zu diesen zählen zum einen, dass eine Rangordnung der Ergebnisse feststellbar sein muss. Das heißt die Zuordnung muss quantifizierbar sein. Weiterhin müssen bereits wenige Trainingsdaten klassifizierbar und Ergebnisse präzisierbar sein können.

In dem vorliegenden Beitrag wurde sich für einen *k-nearest neighbors classifier* entschieden. Dieser erfüllt die Anforderungen und hat in den durchgeführten Experimenten zu guten Ergebnissen geführt. Es sind aber auch andere Machine-Learning-Algorithmen wie z. B. Support-Vector-Maschinen geeignet. Als Zielgröße der Klassifizierung wurden qualitative Klassen („Sehr relevant“, „Nähere Betrachtung sinnvoll“, „Nicht relevant“, ...) gebildet, welche die Relevanz des Ergebnisparameters wiedergeben. Dafür wurden die verwendeten Trainingsdaten manuell ausgewertet und in die entsprechenden Klassen eingeteilt. Die verwendeten Trainingsdaten stammen aus bereits durchgeführten Data-Farming-Studien.

Metadatenbank

In der Kombination mit den extrahierten Metadaten werden die Zielklassen in der Metadatenbank gespeichert und dem Algorithmus zur Verfügung gestellt. Im Falle einer neuen Data-Farming-Studie arbeitet der Anwender mit den prädierten Ergebnissen, welche er auf Basis der Rangliste auswählt. Dadurch erfolgt im weiteren Verlauf der Wissensentdeckung automatisch eine Evaluation der ausgewählten Ergebnisparameter. Bei Bedarf können zum Abschluss der Studie die prädierten Klassen und Metadaten der evaluierten Ergebnisparameter als Erweiterung der Trainingsdaten in der Metadatenbank gespeichert werden.

4 TestszENARIO

Im Folgenden wird das trainierte Meta-Learning-Modell an einem TestszENARIO evaluiert und erste Ergebnisse der Data-Farming-Studie werden diskutiert.

4.1 Simulationsmodell

Zur Evaluation des trainierten Modells wird ein reales Szenario aus dem Logistikbereich herangezogen. Hierbei handelt es sich um eine Umladestation aus dem Automobilzuliefererbereich. Konkret werden Lastkraftwagen (LKW) an zwei Ladedocks mit jeweils unterschiedlicher Beladung (*Produkt A/Produkt B*) durch einen Gabelstapler entladen und mit dem produktneutralen Leergut wieder beladen. Für diese Aufgabe stehen für das entladene Gut sowie das Leergut mehrere Puffer zur Verfügung. Abbildung 2 zeigt einen Screenshot des Simulationsmodells aus Siemens Plant Simulation.

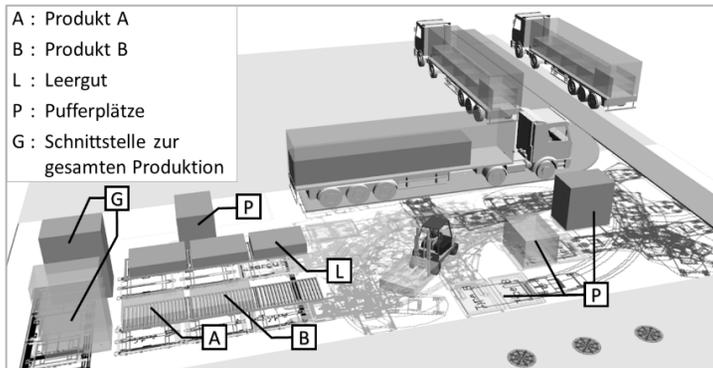


Abbildung 2: Screenshot des 3D-Simulationsmodell.

Zu den variierbaren Eingangsfaktoren des Simulationsmodells gehören die Anteile der Produkte A (*ProduktionAnteilA*) und B (*ProduktionAnteilB*) an der gesamten Produktion. Die Darstellung der gesamten Produktion ist für den Anwendungsfall nicht weiter relevant und ist daher in dem Modell nicht abgebildet. Weitere Faktoren des Simulationsmodells sind das gesamte Produktionsvolumen, die Anlieferzyklen der LKWs an jeweiligen Docks (*ZykluszeitA*, *ZykluszeitB*) und die Größe des Puffers für Vollgut (*Pufferplätze_Vollgut*). Im Gegensatz zu der geringen Anzahl an Eingangsparametern stehen 24 Ergebnisparameter zur Analyse des Szenarios zur Verfügung. Die Ergebnisparameter erfassen unter anderem Daten über den Gabelstapler (*Auslastung*, *Aufträge*, *Blockierzeit*), die LKWs (*Min-/Max-/Durchschnitts-Beladung*, *Backlog*, *Zykluszeit überschritten*), die Puffer und weitere allgemeine Eigenschaften (*Verbrauch Produktion*, *Stillstand Produktion*, *DLZ*). Die Anzahl der in dem vorliegenden Testszenario durchgeführten Experimente der Data-Farming-Studie beläuft sich auf 5000. Um eine gute Raumabdeckung durch die Experimente zu erzielen, wurde ein Latin-Hyper-Cube-Sampling durchgeführt. Die Grenzwerte der variierbaren Faktoren sind in Tabelle 2 dargestellt.

Tabelle 2: Experimentdesign.

Eingangsfaktor	Min-Wert	Max-Wert	Einheit
Produktionsvolumen	25	33	Einheiten/Stunde
ProduktionAnteilA	4,75	50	Prozent
ProduktionAnteilB	0,25	50	Prozent
ZykluszeitA	45	120	Minuten
ZykluszeitB	45	120	Minuten
PufferplätzeVollgut	3	10	Einheiten

4.2 Ergebnisse

Auf die akkumulierten Simulationsdaten der Data-Farming-Studie wurde das trainierte Meta-Learning-Modell zur Filterung der Ergebnisdaten angewendet. Die Resultate der Prädikation der 24 Ergebnisparameter werden in Tabellenform ausgegeben. Einen ersten Überblick über die interessantesten drei Ergebnisparameter liefert Tabelle 3.

Tabelle 3: Ausschnitt der Prädiktionen des Meta-Learning-Modells.

ID Parameter	Klasse	Distanzmaß	Rank
13 Zykluszeit überschritten	Sehr relevant	0,914	1
18 Verbrauch Produktion	Sehr relevant	0,815	2
3 Durchschnittliche Stapelhöhe	Sehr relevant	0,764	3

Die am höchsten bewerteten Ergebnisparameter sind *Zykluszeit überschritten*, *Verbrauch Produktion* und *Durchschnittliche Stapelhöhe*, welche als sehr relevant eingestuft werden. Das Distanzmaß stellt einen Kehrwert der auf das Intervall von Null bis Eins transformierten Entfernungen zu dem Zentrum der jeweiligen Klasse dar. D. h. je höher der Wert, desto zuverlässiger ist die Zugehörigkeit zu der eingeordneten Klasse. In Abbildung 3 ist die Verteilung des Parameters *Zykluszeit überschritten* dargestellt. Dieser Parameter beschreibt, wie viele LKWs ihre Zykluszeit in dem Experimentlauf überschritten haben. Es ist deutlich erkennbar, dass die Normalverteilung von einer Spitze im Intervall [0, 12] durchbrochen wird. Das führt zu der Vermutung, dass es Wechselwirkungen zwischen den Faktoren gibt, die zu einer deutlich stabileren Einhaltung der Zykluszeit führen. Eine genauere Untersuchung im Laufe der Data-Farming-Studie in Bezug auf diesen Ergebnisparameter ist daher sinnvoll.

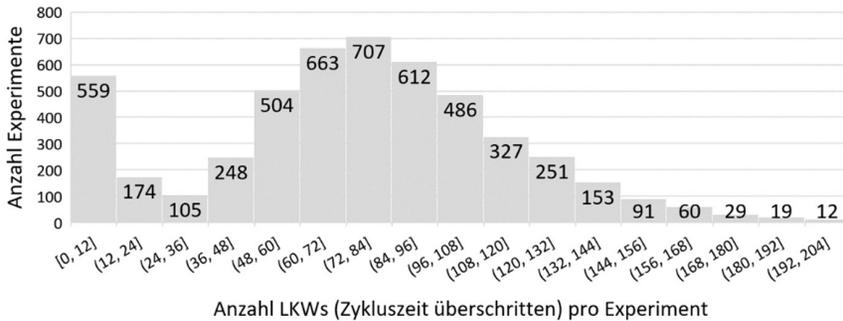


Abbildung 3: Häufigkeitsverteilung der LKWs, welche die Zykluszeit überschreiten.

Auch bei anderen als interessant eingeordneten Ergebnisparametern lassen sich Auffälligkeiten finden. Z. B. fällt bei dem Verbrauch der Produktion im Bereich des Medians ein deutlicher Einbruch in der Verteilung auf. Auch hier kann eine tiefere Analyse zu Wissensgewinn führen, da der Verbrauch der gesamten Produktion einen essenziellen Informationsgehalt über die Performanz der einzelnen Experimentläufe besitzt.

Um die prädizierten Ergebnisse zu validieren, wurde eine mehrdimensionale Klassifizierung des Modellverhaltens mittels eines Gaussian-Mixture-Modells und den genannten Ergebnisparametern durchgeführt. Bei fünf Clustern wurde auf Basis des Silhouettenkoeffizienten durch das Modell eine gute Strukturierung der Daten erreicht. Werden die klassifizierten Daten auf das Intervall [0, 12] des auffälligen Ergebnisparameters *Zykluszeit überschritten* gefiltert, wird deutlich, dass in diesem Abschnitt fast sämtliche Experimente dem Cluster 5 zugeordnet wurden. Gleichzeitig führen diese Experimente auch zu den höchsten Ergebniswerten des Verbrauchs der gesamten Produktion. Daraus schlussfolgernd können die

Experimente, die diesem Cluster angehören, als sehr performant eingestuft werden. Anschließend stellt sich die Frage: Welche Faktoren führen zu diesen performanten Experimenten? Abbildung 4 zeigt die Eingangsfaktoren *Produktionsvolumen*, *ProduktionAnteilA* und *ProduktionAnteilB* nach Cluster 5 gefiltert. Es wird deutlich, dass diese Faktoren eine große Bedeutung in Bezug auf die Clusterzuordnung haben. Insbesondere das Produktionsvolumen scheint einen großen Einfluss auszuüben, da in diesem Cluster ausschließlich Ausprägungen größer oder gleich 30 vorkommen. Entgegengesetzt sind die Faktoren *ProduktionAnteilA* und *ProduktionAnteilB* im niedrigen Bereich wesentlich stärker ausgeprägt. Daraus kann folgende erste Schlussfolgerung über die Systemperformanz gezogen werden: Ein hoher Verbrauch in der gesamten Produktion und ein zuverlässiges Einhalten der Zykluszeit durch die LKWs resultiert aus einem hohen Produktionsvolumen bei relativ niedrigen Produktanteilen von A und B. Dieser Zusammenhang eröffnet die Vermutung, dass ein Engpass in dem abgebildeten System vorliegt, da invers hohe Produktanteile zu einem niedrigen Gesamtverbrauch führen könnten.

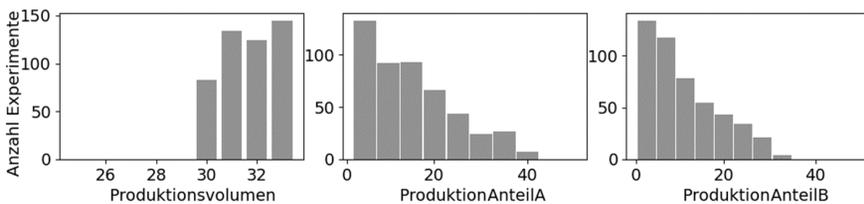


Abbildung 4: Eingangsfaktoren gefiltert nach Cluster 5.

5 Fazit und Ausblick

Dieser Beitrag hat ein Konzept zur automatisierten Identifizierung von statistisch auffälligen bzw. relevanten Ergebnisparametern im Kontext der Wissensentdeckung in Simulationsdaten vorgeschlagen und das entwickelte Modell an einem Anwendungsfall evaluiert. Dadurch konnte die prinzipielle Funktionsfähigkeit des entwickelten Konzepts sowie des trainierten Modells validiert werden. Kritisch zu hinterfragen ist der mögliche Bias, welcher durch die manuelle Klassifizierung der Trainingsdaten entstehen könnte sowie deren geringe verfügbare Menge. Durch eine breite Anwendung und einer online verfügbaren Metadatenbank, könnte dieser Kritikpunkt auf Grund einer kontinuierlichen Erweiterung der Trainingsdaten negiert werden. Ein weiterer notwendiger Schritt im Sinne der Grundidee des Beitrags ist die Unterstützung der anschließenden Prozessschritte der Wissensentdeckung in Simulationsdaten.

Literaturverzeichnis

- Brandstein, A.G.; Horne, G.E.: Data Farming: A Meta-technique for Research in the 21st Century. *Maneuver Warfare Science* (1998), S. 93–99.
- Brazdil, P.; Giraud-Carrier, C.; Soares, C.; Vilalta, R.: *Metalearning: Applications to Data Mining*. Berlin, Heidelberg: Springer 2009.

- Fayyad, U.; Piatetsky-Shapiro, G.; Smyth, P.: Knowledge Discovery and Data Mining: Towards a Unifying Framework. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, 1996, S. 82–88.
- Feldkamp, N.; Bergmann, S.; Schulze, T.; Straßburger, S.: Data Farming im Kontext von Produktion und Logistik. In: Wenzel, S.; Peter, T. (Hrsg.): Simulation in Produktion und Logistik. Kassel: kassel university press 2017, S. 169–178.
- Feldkamp, N.; Bergmann, S.; Strassburger, S.: Knowledge Discovery in Manufacturing Simulations. In: Taylor, S.J.E.; Mustafee, N.; Son, Y.-J. (Hrsg.): Proceedings of the ACM SIGSIM Conference on Principles of Advanced Discrete Simulation, London, United Kingdom, 10.06. - 12.06., 2015, S. 3–12.
- Horne, G.E.; Meyer, T.E.: Data Farming: Discovering Surprise. In: Ingalls, R.G. (Hrsg.): Proceedings of the Winter Simulation Conference, Washington, D.C., 05.12. - 08.12, 2004, S. 792–798.
- Keim, D.; Andrienko, G.; Fekete, J.-D.; Görg, C.; Kohlhammer, J.; Melançon, G.: Visual Analytics: Definition, Process, and Challenges. In: Hutchison, D.; Fekete, J.-D.; Kanade, T.; Kerren, A.; Kittler, J.; Kleinberg, J.M. et al. (Hrsg.): Information Visualization. Berlin, Heidelberg: Springer 2008, S. 154–175.
- Kleijnen, J.P.; Sanchez, S.M.; Lucas, T.W.; Cioppa, T.M.: State-of-the-Art Review: A User’s Guide to the Brave New World of Designing Simulation Experiments. *INFORMS Journal on Computing* 17 (2005) 3, S. 263–289.
- Law, A.M.: Simulation modeling and analysis. Dubuque: McGraw-Hill Education 2013.
- Lemke, C.; Budka, M.; Gabrys, B.: Metalearning: a survey of trends and technologies. *Artificial intelligence review* 44 (2015) 1, S. 117–130.
- März, L.; Krug, W.; Rose, O.; Weigert, G.: Simulation und Optimierung in Produktion und Logistik: Praxisorientierter Leitfaden mit Fallbeispielen. Heidelberg: Springer 2011.
- Ren, M.; Triantafillou, E.; Ravi, S.; Snell, J.; Swersky, K.; Tenenbaum, J.B.; Larochelle, H.; Zemel, R.S.: Meta-Learning for Semi-Supervised Few-Shot Classification. *International Conference on Learning Representations* (2018).
- Rivolli, A.; Garcia, L.P.; Soares, C.; Vanschoren, J.; Carvalho, André C. P. L. F.: Characterizing classification datasets: a study of meta-features for meta-learning (2018).
- Sanchez, S.M.: Simulation experiments: Better data, not just big data. In: Tolk, A. (Hrsg.): Proceedings of the Winter Simulation Conference, Savannah, GA, USA, 7.12. - 10.12, 2014, S. 805–816.
- Vanschoren, J.: Meta-Learning Architectures: Collecting, Organizing and Exploiting Meta-Knowledge. In: Kacprzyk, J.; Jankowski, N.; Duch, W.; Grąbczewski, K. (Hrsg.): Meta-Learning in Computational Intelligence. Berlin, Heidelberg: Springer Berlin Heidelberg 2011, S. 117–155.
- Verein Deutscher Ingenieure VDI 3633: Simulation von Logistik-, Materialfluß- und Produktionssystemen. Berlin: Beuth Verlag GmbH, 2000.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Entwicklung einer integrierten Lösung für das Data Farming und die Wissensentdeckung in Simulationsdaten

Development of an Integrated Solution for Data Farming and Knowledge Discovery in Simulation Data

Jonas Genath, Sören Bergmann, Niclas Feldkamp, TU Ilmenau, Ilmenau(Germany),
jonas.genath@tu-ilmenau.de, soeren.bergmann@tu-ilmenau.de,
niclas.feldkamp@tu-ilmenau.de

Sven Spieckermann, Stephan Stauber, Simplan AG, Hanau (Germany),
sven.spieckermann@simplan.de, stephan.stauber@simplan.de

Abstract: Simulation is an established methodology for planning and evaluating manufacturing and logistics systems. In contrast to classical simulation studies, the method of knowledge discovery in simulation data uses a simulation model as a data generator (data farming). Subsequently, hidden, previously unknown and potentially useful cause-effect relationships can be uncovered on the generated data using data mining and visual analytics methods. So far, however, there is a lack of integrated, easy-to-use software solutions for the application of the data farming in operational practice. This paper presents such an integrated solution, which allows for generating experiment designs, implements a method to distribute the necessary experiment runs, and provides the user with tools to analyze and visualize the result data.

1 Einleitung

Simulation ist ein etabliertes Werkzeug zur Planung und Steuerung komplexer Produktions- und Logistiksysteme und hat sich unter anderem hinsichtlich der Herausforderungen im Kontext der Industrie 4.0 als eine wichtige Schlüsselkomponente erwiesen (Krückhans und Meier 2013). Der Ablauf traditioneller Simulationsstudien ist üblicherweise darauf ausgelegt, durch manuelles Experimentieren einen vorher definierten Projektumfang abzudecken oder ein konkretes (Projekt-)Ziel zu erreichen. Darunter fällt z. B. die Optimierung eines Fertigungslayouts (Law 2003).

Mit steigender Rechenleistung und der allgemeinen Verfügbarkeit von Big-Data-Infrastrukturen und cloudbasierten Lösungen sowie erheblichen Fortschritten auf dem Gebiet des Data Mining ergibt sich eine weitere Anwendungsmöglichkeit für

Simulationsmodelle: Das Durchführen einer sehr großen Bandbreite von Experimenten, um versteckte, vorher unbekannte und möglicherweise nützliche Wirkzusammenhänge aufzudecken. Gerade in komplexen Systemen können möglicherweise Beziehungen, Probleme oder sogar Lösungen vorhanden sein, die über das definierte Projektziel einer klassischen Simulationsstudie hinausgehen, aber dennoch oder gerade deshalb zur Entscheidungsunterstützung beitragen können. Grundlage bildet hierbei die Methodik des Data Farming (Horne und Meyer 2010).

Ein Vorgehensmodell, speziell für die Analyse von Produktions- und Logistiksystemen, das die Methode des Data Farmings aufgreift und mit Methoden des Data Mining und der Visual Analytics ergänzt, wurde von Feldkamp et al. (2017) unter dem Namen Wissensentdeckung in Simulationsmodellen entwickelt. Erste Fallstudien haben erhebliche Potentiale aufgezeigt (Feldkamp et al. 2020a; Feldkamp 2020).

Ein breiter Transfer in die betriebliche Praxis scheitert bisher am Fehlen einer integrierten Softwarelösung, die auch Nicht-Simulations- bzw. Data-Farming-Experten befähigt, Wissensentdeckung in Simulationsmodellen zu betreiben.

In diesem Beitrag wird eine derartige integrierte Lösung vorgestellt, die als neues Modul (4farm) die bestehende Softwarelösung SimAssist (vgl. Sokoll und Clausing 2020) zunächst prototypisch erweitert. Die Entwicklung erfolgte im Rahmen des BMBF-Forschungsprojekts „Entwicklung einer integrierten Lösung für das Data Farming und die Wissensentdeckung in Simulationsdaten (DaWiS)“. Die zu betrachtenden Teilaspekte sind hierbei: Verfahren des intelligenten Experimentdesigns, Methoden zur (cloudbasierten) Verteilung der Experimente sowie die Auswahl und Adaption geeigneter Data-Mining- und Visual-Analytics-Methoden, so dass mit geringem Einarbeitungsaufwand effektiv Data Farming bzw. die Methode der Wissensentdeckung in Simulationsmodellen nach Feldkamp (2020) angewendet werden kann.

Im Rahmen dieses Beitrags werden die Methode und die implementierte Softwarelösung anhand eines Beispiels aus der Automobilbranche erläutert. Die eigentliche Simulation erfolgt exemplarisch in der verbreiteten Simulationssoftware Siemens Plant Simulation.

Die folgenden Kapitel dieses Beitrags gliedern sich wie folgt: Zunächst werden der Stand der Forschung und die notwendigen theoretischen Grundlagen des Data Farming und der Methode zur Wissensentdeckung in Simulationsdaten (Knowledge Discovery in Simulation Data – KDS) kurz dargestellt. Darauf folgend wird im Hauptteil des Beitrags die integrierte Methode schrittweise, den Teilaspekten folgend, vorgestellt und durch ein Workflow-Beispiel illustriert. An den entsprechenden Stellen wird hierbei insbesondere auf die technische Umsetzung eingegangen. Der Beitrag wird mit einem Fazit sowie einem Ausblick auf mögliche Erweiterungen der integrierten Lösung abgeschlossen.

2 Grundlagen und Stand der Forschung

Beim so genannten Data Farming wird ein vorab validiertes Simulationsmodell als Datengenerator genutzt, um mit Hilfe von intelligentem Experimentdesign und High Performance Computing ein möglichst großes oder gar vollständiges Spektrum an Modell- bzw. Systemverhalten (Wirkungsraum) abdecken zu können (Sanchez

2007; Horne und Meyer 2010). Die „Farming“-Metapher drückt hierbei aus, dass analog zu einem Farmer, der sein Land möglichst effizient kultiviert, um seinen Ernteertrag zu maximieren, beim Data Farming das Ziel ist, den Datenertrag des Simulationsmodells zu maximieren (Sanchez 2014).

Entscheidend ist hierbei unter anderem die Forschung zu verbesserten Verfahren für die Gestaltung von Simulationsexperimentplänen. Diese erlauben mögliche Wertekombinationen von Eingabeparametern umfassend abzubilden und garantieren trotzdem gleichzeitig eine vertretbare Datenmenge und Zahl von Experimentläufen (Kleijnen et al. 2005; Sanchez und Wan 2009). Gerade im Kontext der Simulation von Produktions- und Logistiksystemen ist die Auswahl einer der Designmethoden oder sogar die Auswahl einer geeigneten Kombination von verschiedenen Designmethoden von großer Bedeutung. Zur Durchführung der Experimente wird in der Data-Farming-Literatur oft auf entsprechendes High Performance Computing verwiesen (Horne und Meyer 2010).

Auf der erzeugten Datenbasis können dann mit Hilfe von verschiedenen Data-Mining- bzw. Visual-Analytics-Methoden interessante Zusammenhänge in den erzeugten Daten aufgedeckt werden (Keim et al. 2008). So können möglicherweise bisher unbekannt Beziehungen, Probleme oder sogar Lösungen identifiziert werden. Eine Auswahl möglicher Data-Mining-Methoden, z. B. das Clustering, und das Wechselspiel der Methoden untereinander stellt Feldkamp (2020) vor. Empfohlen wird hierbei, dass die eigentliche Analyse der generierten Simulationsergebnisdaten sowie die Beziehungen zwischen Eingangsparametern (Faktoren) und Ergebnisdaten (Kennzahlen) idealerweise mit interaktiver, visueller Analyse unterstützt wird. Visualisierung ist allgemein ein entscheidendes Werkzeug, wenn eine Interpretation von Daten gefordert ist. Eine konsequente Verzahnung, wie es allgemein in der Forschungsdisziplin Visual Analytics empfohlen wird, zwischen interaktiver Visualisierung, z. B. mittels interaktiv adaptierbare Animationen, Zeitreihendiagramm, Graphen, und der Datenanalyse mittels Data-Mining-Methoden ermöglicht dem Anwender, die menschliche Fähigkeit zur Schlussfolgerung bestmöglich einzubringen (Keim et al. 2008; Feldkamp et al. 2020b).

Zusammenfassend ist zum Stand von Wissenschaft und Technik einzuschätzen, dass die grundlegenden Einzelmethoden (Data Farming, Intelligentes Experimentdesign, Data Mining und Visual Analytics) im Stand der Forschung durchaus einen guten Stand erreicht haben. Auch erste erfolversprechende prototypische Anwendungen im Kontext der Simulation von Produktions- und Logistiksystemen sind publiziert worden. Es ist aber zu konstatieren, dass es bisher keine ganzheitliche Lösung zur Übertragung der Methoden als Ganzes oder zumindest in wesentlichen Teilen in ein von Nicht-Experten bedienbares Framework gibt, welches zudem auf den Bereich der Simulation in Produktion und Logistik fokussiert. Darüber hinaus fehlen Methoden zur (Teil-)Automatisierung der Prozesse und zur Unterstützung von Nicht-Experten im Allgemeinen.

3 Integrierte Lösung für das Data Farming und die Wissensentdeckung in Simulationsdaten

Wie in der Einleitung bereits erwähnt, ist das Ziel des BMBF-Forschungsprojekts DaWiS die Entwicklung einer Softwarelösung, ergänzt durch Best-Practice-Vorgehensweisen, welche es auch Nicht-Experten erlaubt, Wissen auf Basis von Data Farming mittels Data-Mining- und Visual-Analytics-Methoden zu erwerben.

Hierzu wird die erprobte modular aufgebaute Software SimAssist (vgl. Sokoll und Clausning 2020) der Simplan AG, welche bisher bereits umfangreiche Assistenzfunktionen zur Verwaltung, Analyse, Visualisierung und Dokumentation von Ergebnisdaten klassischer Simulationsprojekte bereitstellt, erweitert. Die Erweiterungen werden hierbei in einem neuen Modul (4farm) zusammengefasst. So werden jeweils für die bereits angesprochenen Teilaspekte, dem intelligenten Experimentdesign, der Verteilung von Experimenten sowie für das Data Mining und Visual Analytics, entsprechende Komponenten konzipiert und prototypisch implementiert. Die grundlegende Architektur des Moduls ist Abbildung 1 zu entnehmen.

Vorab ist zu bemerken, dass insgesamt bei der Konzeption und Entwicklung hohe Anforderungen an die Gestaltung der Nutzerinteraktionen gestellt werden, sodass alle nötigen Teilaspekte (vom Experimentdesign über die Verteilung der Experimentläufe bis hin zur Analyse und Visualisierung der Daten) ohne Wechsel der Software über eine integrierte Oberfläche verfügbar sind.

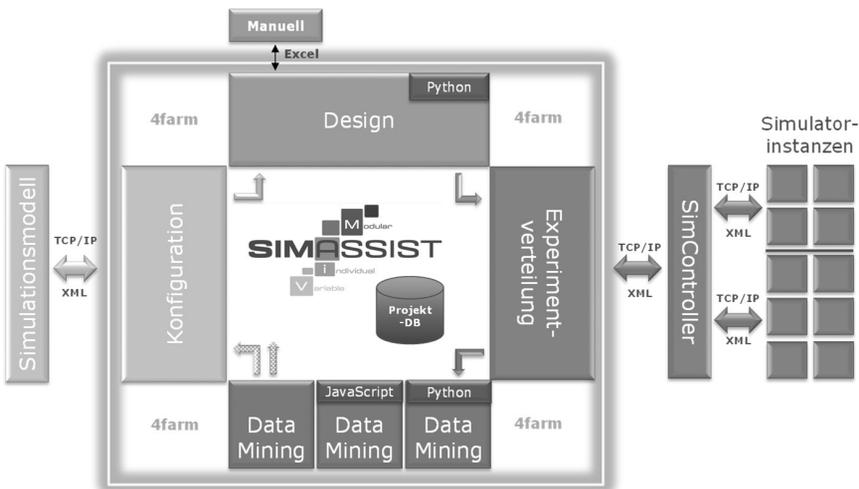


Abbildung 1: Grobarchitektur der Softwarelösung (SimAssist – 4farm-Modul) für das Data Farming und die Wissensentdeckung in Simulationsdaten

Des Weiteren werden vor dem Endanwender möglichst viele technische Details insbesondere des Experimentdesigns, der Data-Mining- und Visual-Analytics-Methoden verborgen oder die erforderlichen Entscheidungen werden basierend auf Best-Practices vorweggenommen. Ist dies nicht möglich, werden die nötigen Einstellungen auf eine intuitiv verständliche fachliche Anwendungsebene

transferiert. Dies geschieht, indem erforderliche Parameter bei der Nutzung der Methoden abgefragt werden. Dabei werden dem Nutzer jeweils Listen mit Auswahlmöglichkeiten angeboten. Die entsprechenden Hinweise zur Nutzung sowie den Vor- und Nachteilen der einzelnen Varianten sind in der Software in Form von Informationstexten oder Entscheidungsbäumen o. ä. hinterlegt. Der Anwender kann sich somit vollständig auf das Modell und die fachliche Aufgabe fokussieren.

Aufgrund der sich rasant weiterentwickelnden Forschung im Bereich des Data Mining und Visual Analytics aber auch der Vielzahl an Möglichkeiten bezüglich des Experimentdesigns ist eine weitere an die Software gestellte Anforderung die Eigenschaft, dass in Zukunft leicht neue Methoden inklusive Visualisierungen über einen standardisierten Mechanismus hinzugefügt werden können.

3.1 Workflow-Beispiel – Versorgung einer PKW-Produktionslinie mit Batterien

Im Rahmen dieses Beitrags werden ausgewählte Methoden sowie die implementierte Softwarelösung – insbesondere die Auszüge der Nutzeroberflächen – anhand eines Beispielszenarios vorgestellt. Unter anderem wurde dieses Beispiel im DaWiS-Projekt von der Anforderungsanalyse bis zur Demonstration der Methoden und der Software genutzt.

Das Workflow-Beispiel umfasst eine typische logistische Problemstellung, in welcher die Versorgung einer laufenden PKW-Produktion mit zwei verschiedenen Batterietypen sowie die Entsorgung der anfallenden leeren Ladungsträger betrachtet wird. Die Lieferung der Batterien in Ladungsträgern und die Abholung der leeren Ladungsträger erfolgt per LKW an einem Entladedock. Die eigentliche Handhabung der Batterien erfolgt durch Gabelstapler (Abbildung 2).

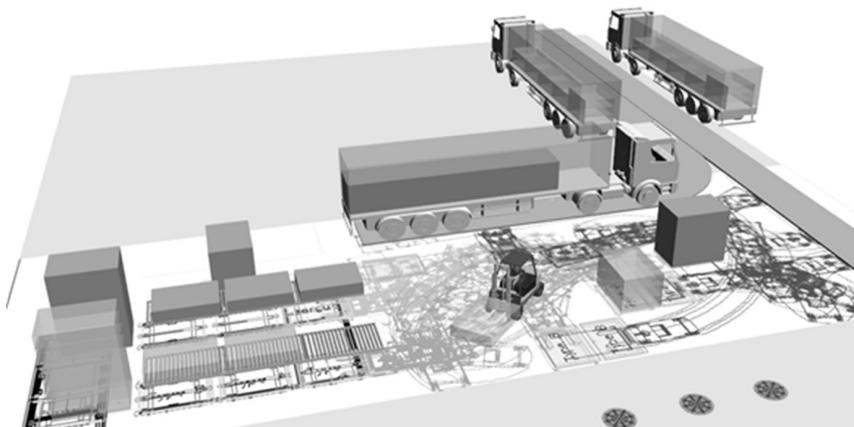


Abbildung 2: Bildschirmausschnitt des Modells einer Be-/Entladung von Batterien zur Versorgung einer laufenden Produktion.

Zu den variierbaren Eingangsparametern (Faktoren) des Modells gehören der Anteil der Batterietypen A (PAR_ProduktionAnteilA) und B (PAR_ProduktionAnteilB) am gesamten Produktionsprogramm der angeschlossenen Fahrzeugmontage (die

selbst nicht Teil des betrachteten Simulationsmodells ist). Weitere Faktoren des Simulationsmodells sind das Gesamtproduktionsvolumen (PAR_Produktionsvolumen), die Zyklen, in welchen LKWs neue Batterien anliefern (PAR_ZykluszeitA, PAR_ZykluszeitB) sowie die Größe der Puffer für Vollgut (PAR_Pufferplätze_Vollgut). Zudem sind vier verschiedene Szenarien (PAR_Szenarien) definiert, die sich je durch die genutzten Staplervarianten (5t- bzw. 8t-Stapler) und Stapleranzahlen (1–3 Stapler) unterscheiden. Zur Analyse der Ergebnisse werden 31 Ergebnisparameter gespeichert, u. a. die Staplerauslastungen oder die Stillstandzeiten der angeschlossenen Montage aufgrund fehlender Batterien.

Für die eigentliche Modellierung des Produktions- und Logistiksystems, mit der Software Siemens Plant Simulation, wurden Bausteine entwickelt, welche das Setzen der Faktoren (Stellgrößen) sowie das Auslesen der vorab definierbaren Ergebnisdaten (XML-Format) ermöglichen. Mit Hilfe dieser Bausteine können Modelle mit Metainformationen über die Faktoren und Ergebnisparameter sowie deren Datentypen und gegebenenfalls vorhandene Wertebereiche angereichert werden. Diese Informationen können durch die integrierte Lösung ausgewertet und dem Nutzer präsentiert und als Grundlage z. B. für das konkrete Experimentdesign genutzt werden.

3.2 Experimentdesign

Wie bereits beschrieben, ist das Generieren geeigneter Experimentpläne die erste große Herausforderung im Prozess des Data Farmings und somit auch bei der darauf aufbauenden Methode zur Wissensentdeckung in Simulationsdaten. Aktuell wurden fünf verschiedenen Experimentdesignmethoden ausgewählt und implementiert: das vollfaktorielle Design, das 2^k Design, das Central Composite Design und das Latin Hypercube Sampling (LHS) sowie ein Design, in welchem ein LHS mit einem weiteren Faktor bzw. Design gekreuzt werden kann.

Zu beachten ist, dass Experimentdesignmethoden neben den Parametern, welche die Faktoren beschreiben, also den Namen der Faktoren, dem Datentyp und Wertebereich des Faktors mitunter auch methodenspezifische Parameter benötigen. So muss etwa beim LHS die Zahl der Experimente angegeben werden. Alle Designmethoden sind als Python-Skripte realisiert. Die Skripte nutzen hierbei eine einheitliche Bibliothek für den XML-basierten Datenimport und -export und können zu Beginn entsprechende Metainformationen als Kommentare enthalten. Die Informationen über die Faktoren können manuell gesetzt werden oder – wie im vorherigen Abschnitt angedeutet – aus dem Simulationsmodell ausgelesen werden. Die designmethodenspezifischen Parameter werden in SimAssist beim Nutzer abgefragt. Welche Parameter abgefragt werden und wie sich die Oberfläche im 4farm-Modul gestaltet, wird in den Metainformationen des jeweiligen Skriptes definiert. Ein Beispiel ist in Abbildung 3 zu sehen, in der eine gekreuztes LHS parametrisiert wird.

Das Hinzufügen weiterer Experimentdesignmethoden ist jederzeit auch ohne Neustart der Software möglich. Hierzu müssen die Methoden nur als mit den angesprochenen Metainformationen annotierte Python-Skripte durch Kopieren in ein festgelegtes Verzeichnis bekannt gemacht werden (sogenanntes Hot-Deployment). Die entsprechende Auswahlmöglichkeit und auch die Oberflächen werden ad hoc generiert und können sofort genutzt werden.

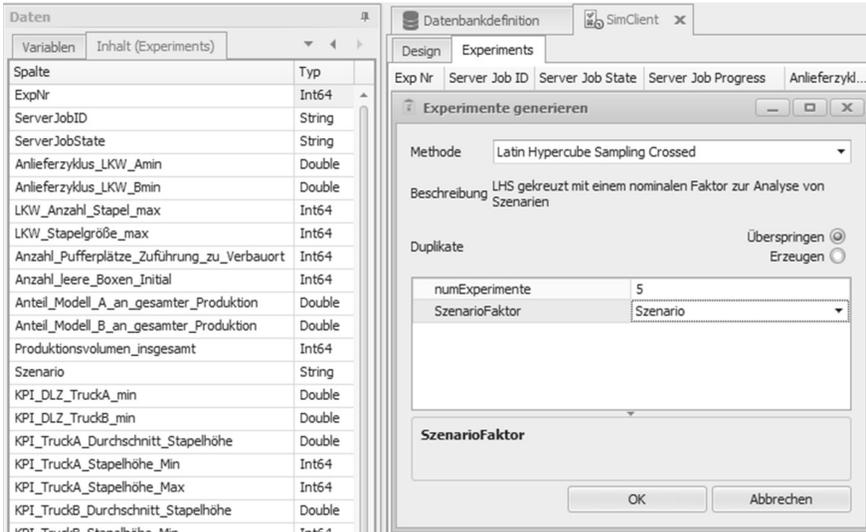


Abbildung 3: Ausschnitt der Oberfläche der Experimentdesignkomponente des 4farm Moduls

Bei der Auswahl der Experimentdesignmethoden wird der Nutzer zu einem durch Hilfetexte zu den einzelnen Methoden unterstützt. Zum anderen ist eine Hilfestellung in Form eines Entscheidungsbaums (Abbildung 4) vorhanden, in welchen durch die Beantwortung einfacher Fragen eine Designmethode vorgeschlagen wird. Im hier angesprochenen Beispielszenario wurde auf Grund der verschiedenen Szenarien ein gekreuztes LHS-Design mit 15.000 Experimenten (5000 LHS * 3 Szenarien) angewendet (grüner/gestrichelter Pfad in Abbildung 4).

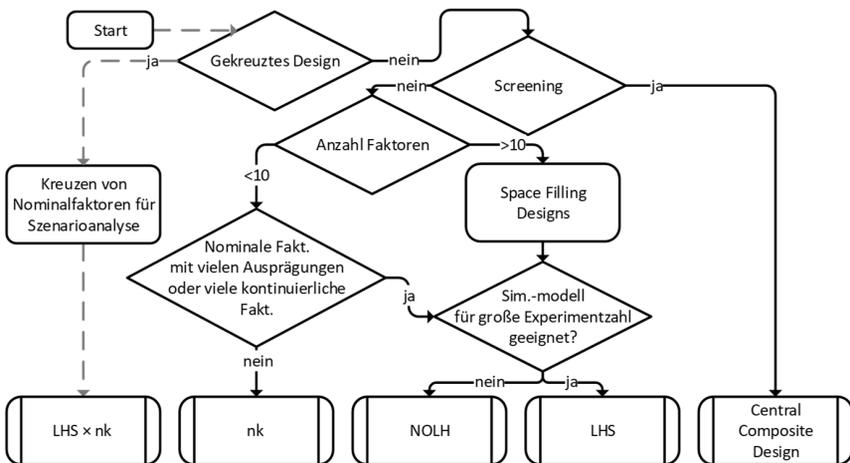


Abbildung 4: Flowchart zur Auswahl geeigneter Experimentdesignmethoden

3.3 Experimentverteilung

Die Durchführung der Experimente im Data Farming ist auf Grund der recht hohen Anzahl an Simulationsläufen auf einer einzelnen Recheninstanz (einzelner Rechner bzw. einzelner Prozessorkern auf einem einzelnen Rechner) meist nicht sinnvoll. Wünschenswert ist vielmehr das Verteilen einzelner Experimente auf verschiedene Recheninstanzen. Die im DaWiS Projekt genutzte Technologie basiert hierbei auf einer aus einem Vorprojekt der Simplan AG stammenden Softwarekomponente, dem so genannten SimController. Dieser wurde so angepasst, dass es nun möglich ist, Experimente, d. h. Modelle und die konkreten Faktorenwerte in Form von XML-Daten über eine zentrale Instanz auf sich selbst registrierende Clientinstanzen mittels TCP/IP zu verteilen. Die einzelnen Instanzen melden laufend ihren Status. Nach Durchlauf eines Experimentes werden die definierten Kennwerte (und genutzten Faktorwerte) an die zentrale Instanz zurückgemeldet und in Form einer SQLite Datenbank in SimAssist gespeichert. Diese Daten können im Anschluss sehr unkompliziert weiter genutzt und analysiert werden. Die Oberfläche der Experimentverteilung, in welcher u. a. der aktuelle Bearbeitungsfortschritt und die Ergebnisdaten zu sehen sind, ist Abbildung 5 zu entnehmen.

Design		Experiments				
Exp Nr	Server J...	Server Job ...	Server Job Progress	KPI_DLZ_Truck A...	KPI_DLZ_Truck B...	KPI_Verbrauch_Pr...
51	1723	Finished	100 %	3721,18622324097	7862,89464324...	54264
50	1724	Started	71 %			
49	1725	Started	72 %			
48	1726	Finished	100 %	2931,70201137141	8109,91975358...	54264
47	1727	Started	0 %			

Abbildung 5: Ausschnitt der Oberfläche der Experimentverteilungskomponente

3.4 Analyse der Ergebnisdaten - Data Mining und Visual Analytics

Analog zu den Experimentdesignmethoden ist die Zahl möglicher Data Mining Methoden und Visualisierungsmethoden sehr groß. Im Vorgehensmodell zur Wissensentdeckung in Simulationsdaten analysiert Feldkamp (2020) verschiedene Gruppen von Methoden und gewichtet diese bzgl. ihres Nutzens für die Wissensentdeckung. Aufbauend auf dieser Forschung wurden Verfahren der deskriptiven Statistik, der Korrelationsanalyse, zum Clustering (k-means und gaussian mixture) sowie zur Regressionsanalyse und zur Bildung von Klassifizierungsbäumen als für die praktische Anwendung am wertvollsten eingestuft und im Prototyp implementiert. Hinzu kommen jeweils geeignete Visualisierungen wie Heatmaps für die Korrelationsanalyse oder Parallelkoordinaten- und Streudiagramme für die Auswertung geclusterter Daten. Die technische Umsetzung erfolgt hierbei analog der Umsetzung der Experimentdesignmethoden, d. h. jede der Methoden ist als annotiertes Python-Skript realisiert und der Datenaustausch mit dem Skript erfolgt wiederum per XML-Format. Auch hier ist es somit einfach möglich, weitere Data-Mining-Methoden und Visualisierungen zu implementieren, welche dem Anwender sofort über generische Dialoge aus dem 4farm-Modul zur Wissensentdeckung zur Verfügung stehen. Beispielhaft ist ein Dialog, hier zum Starten der Clusteranalyse auf den Daten des Workflow-Beispiels, in Abbildung 6 zu sehen.

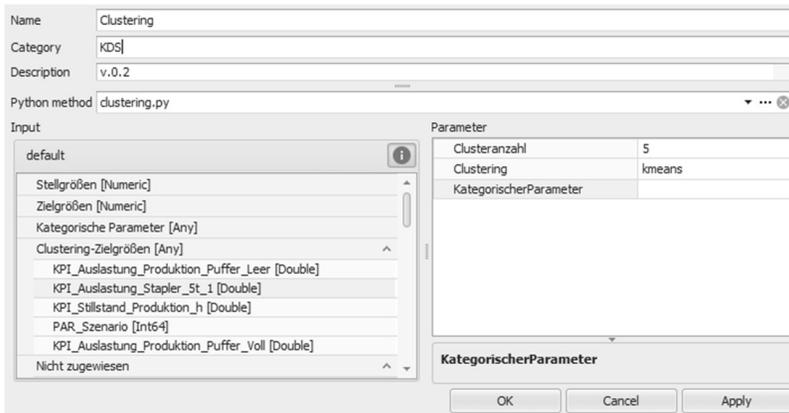


Abbildung 6: Ausschnitt der Oberfläche der Data-Mining-Komponente des 4farm-Moduls

Die Datenanalysen des Workflow-Beispiels sind aktuell noch nicht vollständig abgeschlossen. Erste Erkenntnisse zeichnen sich aber ab. Es kristallisiert sich z. B. heraus, dass bei Annahme des aktuellen Bedarfs an Batterien jedes Szenario zu einer sicheren Versorgung der Produktion führt, aber bereits bei moderat steigenden Anteilen an batterieelektrischen Fahrzeugen, insbesondere wenn ein Batterietyp deutlich häufiger nachgefragt wird, Szenarien mit zwei Staplern, wobei mindestens einer ein 8t Stapler ist, besser funktionieren.

4 Fazit und Ausblick

Der Beitrag stellt eine sich in Entwicklung befindende integrierte Softwarelösung für die Wissensentdeckung in Simulationsdaten vor. Hierzu wurden der Bedarf für und die Anforderungen an eine solche Lösung hergeleitet sowie die wesentlichen Teilaspekte der Methode und ihre nutzerfreundliche prototypische Umsetzung näher betrachtet. Weitere Entwicklungsschritte sind u. a. die Implementierung weiterer Experimentdesign- und Data-Mining-Methoden sowie zusätzlicher Visualisierungen. Zudem sind weitere Tests in Anwendungsfällen aus der Praxis nötig, insbesondere um die implementierten Schnittstellen und Datei-Formate zu validieren. Des Weiteren sind Forschungen zur weiteren (Teil-) Automatisierung der Data-Mining-Methoden, z. B. mittels Meta-Learning zu Bestimmung geeigneter Hyperparameter, oder auch die Nutzung von Methoden zur Robustheitsanalyse denkbar.

Danksagung

Das DaWiS Forschungsprojekt wird gefördert vom Bundesministerium für Bildung und Forschung. Die Förderung startete im Januar 2020 und wird 2 Jahre dauern.



Bundesministerium
für Bildung
und Forschung

Literatur

- Feldkamp, N.: Wissensentdeckung im Kontext der Produktionssimulation. Ilmenau: Universitätsverlag Ilmenau 2020.
- Feldkamp, N.; Bergmann, S.; Strassburger, S.: Knowledge Discovery in Simulation Data. *ACM Transactions on Modeling and Computer Simulation* 30 (2020a) 4, S. 1–25.
- Feldkamp, N.; Bergmann, S.; Strassburger, S.: Visualization and Interaction for Knowledge Discovery in Simulation Data. In: Bui, T. (Hrsg.): *Proceedings of the 53rd Hawaii International Conference on System Sciences, Maui, 7-10.01, 2020, 2020b*, S. 1340–1349.
- Feldkamp, N.; Bergmann, S.; Straßburger, S.; Schulze, T.: Data Farming im Kontext von Produktion und Logistik. In: Wenzel, S.; Peter, T. (Hrsg.): *Simulation in Produktion und Logistik 2017, Kassel, 20-22.09, 2017*, S. 169–178.
- Horne, G.E.; Meyer, T.: Data farming and defense applications. In: Armstrong, R.; McNamara, J.; Pinelli, T.E. (Hrsg.): *MODSIM World Conference and Expo, Hampton, VA, USA, 13-15.10, 2010*, S. 74–82.
- Keim, D.A.; Mansmann, F.; Schneidewind, J.; Thomas, J.; Ziegler, H.: Visual Analytics: Scope and Challenges. In: Simoff, S.J.; Böhlen, M.H.; Mazeika, A. (Hrsg.): *Visual Data Mining. Berlin, Heidelberg: Springer 2008*, S. 76–90.
- Kleijnen, J.P.; Sanchez, S.M.; Lucas, T.W.; Cioppa, T.M.: State-of-the-Art Review: A User’s Guide to the Brave New World of Designing Simulation Experiments. *INFORMS Journal on Computing* 17 (2005) 3, S. 263–289.
- Krückhans, B.; Meier, H.: Industrie 4.0 – Handlungsfelder der Digitalen Fabrik zur Optimierung der Ressourceneffizienz in der Produktion. In: Dangelmaier, W.; Laroque, C.; Klaas, A. (Hrsg.): *Proceeding der 15. ASIM Fachtagung Simulation in Produktion und Logistik 2013, Paderborn, 09.-11.10, 2013*, S. 31–40.
- Law, A.M.: How to conduct a successful simulation study. In: Chick, S.; Sanchez, P.J.; Ferrin, D.; Morrice, D.J. (Hrsg.): *Proceedings of the 2003 Winter Simulation Conference, New Orleans, LA, USA, 7.12-10.12, 2003*, S. 66–70.
- Sanchez, S.M.: Work Smarter, Not Harder: Guidelines for Designing Simulation Experiments. In: Henderson, S.G.; Biller, B.; Hsieh, M.-H.; Shortle, J.; Tew, J.D.; Barton, R.R. (Hrsg.): *Proceedings of the 2007 Winter Simulation Conference, Washington, DC, USA, 9-12.12, 2007*, S. 84–94.
- Sanchez, S.M.: Simulation Experiments: Better Data, Not Just Big Data. In: Tolk, A.; Diallo, S.D.; Ryzhov, I.O.; Yilmaz, L.; Buckley, S.; Miller, J.A. (Hrsg.): *Proceedings of the 2014 Winter Simulation Conference, Savannah GA, 7.12.-10.12., 2014*, S. 805–816.
- Sanchez, S.M.; Wan, H.: Better than a petaflop: The power of efficient experimental design. In: Rossetti, M.D.; Hill, R.R.; Johansson, B.; Dunkin A.; Ingalls, R.G. (Hrsg.): *Proceedings of the 2009 Winter Simulation Conference (WSC 2009), Austin, TX, 13.-16. Dezember, 2009*, S. 60–74.
- Sokoll, K.; Clausing, M.: Methoden und Werkzeuge der Simulationsassistentz. In: Mayer, G.; Pöge, C.; Spieckermann, S.; Wenzel, S. (Hrsg.): *Ablaufsimulation in der Automobilindustrie. Berlin: Springer Vieweg 2020*, S. 349–364.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Simulationsgestützter Ansatz zur Verbesserung der Leistungsprognose maschineller Tunnelvortriebe mithilfe baubegleitender Prozessdaten

***Simulation-based approach to enhance the performance prediction of
mechanized tunneling using actual data during construction***

Annika Jodehl, Yara Salloum, Markus König, Markus Thewes, Ruhr-Universität
Bochum, Bochum (Germany), annika.jodehl@rub.de, yara.salloum@rub.de,
koenig@inf.bi.rub.de, markus.thewes@rub.de

Abstract: Performance prognoses in mechanised tunnelling are complex because production and support processes are closely interdependent while depending on typically highly variable ground conditions. Process simulation is a suitable tool to analyse these complex interrelationships. So far, developed simulation models were mainly aiming at planning phase. Extending the simulations towards integrating real-time data over the execution of a project could support continuous improvement of the models, decision-making in the event of unexpected disturbances and a robust incident management. In mechanised tunnelling, this is particularly interesting because in planning many input variables are based on assumptions and often subject to uncertainties and fuzziness and therefore deviate during execution. This paper presents an approach to update input data of simulation models using actual data during construction. By continuously adapting the model and analysing the impact on the project duration, the aim is to improve performance prediction and support decision making during construction.

1 Einleitung

Die Leistungsprognose maschineller Tunnelvortriebe ist aufgrund vieler Prozessinteraktionen und der sich stetig ändernden Randbedingungen während des Vortriebs sowie der Variabilität des Baugrunds sehr komplex. Die Verfügbarkeit eines Vortriebs hängt dabei in einem hohen Maße von der Leistungsfähigkeit der Support- und Logistikprozesse sowie der Robustheit einzelner Abläufe gegenüber Leistungsschwankungen und ungeplanten Stillständen ab. Durch den Einsatz von Prozesssimulationen ist es möglich die Interaktionen der Prozesse zu untersuchen und leistungsmindernde Elemente zu identifizieren.

Im maschinellen Tunnelbau existieren bereits einige Simulationsansätze zur Analyse der Leistungsfähigkeit von Tunnelvortrieben. In den meisten Ansätzen werden jedoch die Produktions- und die Logistikprozesse getrennt voneinander betrachtet (Al-Bataineh et al., 2013; Ebrahimi et al., 2011). Eine konkrete Analyse der Interaktion der Logistikprozesse mit den Produktionsprozessen findet selten statt. Für die Modellierung solcher Prozessabläufe und -interaktionen werden im Bauwesen in der Regel ereignisdiskrete Simulationen eingesetzt (Akhavian und Behzadan, 2013). Die Eingangsdaten der Simulationen basieren oft auf Verteilungsfunktionen, um Schwankungen oder Unsicherheiten der Eingangsdaten (z.B. Prozessdauern, Ausfallwahrscheinlichkeiten) adäquat abzubilden. Die Anwendung dieser Simulationsansätze ist jedoch häufig auf die frühen Planungsphasen beschränkt und findet kaum Betrachtung in der Ausführung (Akhavian und Behzadan, 2013; Wu et al., 2019).

Eine Erweiterung der Simulationen um Komponenten zur Integration von Echtzeitdaten könnte eine kontinuierliche Verbesserung der Modelle, eine Entscheidungsfindung bei unerwarteten Stillständen während der Ausführung sowie ein robustes Störfallmanagement unterstützen. Im maschinellen Tunnelbau ist dies besonders interessant, da in der Planung viele Eingangsdaten auf Annahmen basieren und häufig mit Unsicherheiten und Unschärfen behaftet sind. Daher treten häufig Abweichungen während der Ausführungsphase auf. Zusätzlich werden die Annahmen der Eingangsdaten durch die gegebenen Randbedingungen und Baugrundverhältnisse des individuellen Projekts erschwert, welche prozessübergreifend im Vergleich zu anderen Bereichen der Baubranche besonders stark voneinander abhängen.

Dieser Beitrag stellt einen Ansatz zur Aktualisierung von stochastisch verteilten Eingangsdaten eines Simulationsmodells unter Verwendung von baubegleitenden Ist-Daten für den maschinellen Tunnelbau vor. Durch eine kontinuierliche Anpassung des Simulationsmodells und stetige Analyse der Auswirkungen auf die Projektdauer soll eine Verbesserung der Leistungsprognose und eine Unterstützung der Entscheidungsfindung während des Bauablaufs erzielt werden.

2 Grundlagen

In diesem Abschnitt wird die Funktionsweise einer Tunnelbohrmaschine am Beispiel einer Schildmaschine mit flüssigkeitsgestützter Ortsbrust kurz beschrieben. Darauf folgt als Hintergrund für das vorgestellte Konzept zur Prognoseanpassung ein kurzer Überblick über Entwicklungen in der Echtzeitsimulation.

2.1 Maschinelles Schildvortrieb

Für Tunnelvortriebe in Lockergestein, in instabilen Baugrundverhältnissen oder unter dem Grundwasserspiegel werden Schildmaschinen eingesetzt. Schildmaschinen sind von einer zylindrischen Stahlkonstruktion, dem Schildmantel, umgeben, die den ausgebrochenen Hohlraum bis zum Einbau der endgültigen Tunnelsicherung stützt (Abb. 1). Bei Vortrieben mit nicht standfester Ortsbrust oder unterhalb des Grundwasserspiegels muss außerdem eine aktive Stützung der Ortsbrust erfolgen. Schildmaschinen mit einer flüssigkeitsgestützten Ortsbrust bringen diese Stützung durch eine unter Druck stehende Suspension in der Abbaukammer auf.

Der maschinelle Schildvortrieb besteht aus zwei alternierenden Kernprozessen, dem Vortrieb und dem Ringbau. Während des Vortriebs wird der Boden durch ein mit

Werkzeugen bestücktes Schneidrad abgebaut. Anschließend folgt die Sicherung des erzeugten Hohlraums mit Stahlbetonsegmenten, den sogenannten Tübbing. Die Kernprozesse werden von zahlreichen Supportprozessen kontinuierlich unterstützt. Ein Ausfall einzelner Prozesse kann zu einem Ausfall des gesamten Vortriebs führen und durch Fortpflanzungseffekte weitere Störungen hervorrufen.

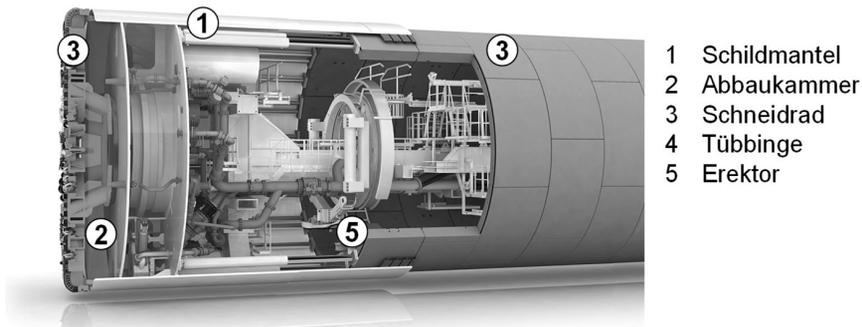


Abbildung 1: Aufbau einer Schildmaschine mit flüssigkeitsgestützter Ortsbrust (Herrenknecht AG, 2021)

Während des Vortriebs werden im maschinellen Tunnelbau durch eine Vielzahl von Sensoren große Datenmengen aufgezeichnet, um die Produktionsprozesse mithilfe eines Prozesscontrolling-Systems zu überwachen, zu bewerten und zu steuern. Sämtliche für den Vortrieb relevanten Maschinendaten wie die Penetration, die Vorschublänge oder der Stützdruck werden so erfasst. Weitere Datenquellen sind Schichtprotokolle, Daten aus dem Geomonitoring oder aus dem Werkzeugmanagement (Hegemann et al., 2020). Mithilfe von Datenmodellierungstechniken können diese Daten verwendet werden, um die Eingangsdaten für Logistiksimulationen zu verbessern, Muster und Zusammenhänge in den Daten für die Aktualisierung von Prognosen zu identifizieren.

2.2 Echtzeitsimulationen

In der Forschung werden in den letzten Jahren zunehmend Methoden zur Echtzeitsimulation untersucht. Auch im Bauwesen existieren erste Ansätze zur baubegleitenden Anwendung von Prozesssimulationen. Viele Ansätze beschäftigen sich speziell mit der Aktualisierung von Eingangsdaten, die mit Unsicherheiten behaftet sind. Zahlreiche Ansätze basieren dabei auf der Anwendung des Bayes'schen Theorems (Bradley et al., 2015; Wu et al., 2019). Das Bayes'sche Theorem erlaubt die Generierung einer neuen Verteilungsfunktion für die Prognose von Daten durch die Kombination neuer Beobachtungen mit einem a-priori Wissen. Das a-priori Wissen, bspw. die Informationen aus der Vorplanung, wird dabei in der Regel mit Wahrscheinlichkeitsverteilungen abgebildet, wodurch eine systematische Berücksichtigung von Unsicherheiten ermöglicht wird. Die Qualität der Ergebnisse ist jedoch in hohem Maß von der Wahl der a-priori-Verteilung abhängig. Bisher entwickelte Ansätze sind daher häufig nicht allgemein übertragbar.

Im Tunnelbau untersuchten Haas und Einstein (2002) die Anwendung des Bayes'schen Theorems für das Update des Decision Aid for Tunneling (DAT), einem Modell zur Prognose von Projektzeiten und Kosten unter Berücksichtigung von Unsicherheiten. Sie aktualisierten beispielhaft geotechnische Parameter und demonstrierten die Auswirkung der Aktualisierung an einem theoretischen Fallbeispiel. Chung T.H. et al. (2004) erforschten den Einsatz des Bayes'schen Theorems zur Aktualisierung der Penetrationsrate für die Simulation eines Tunnelvortriebs. Wu et al. (2019) untersuchten den Einsatz Bayes'scher Inferenzen gekoppelt mit Markov-Ketten, um den Einsatzbereich der Aktualisierungen durch das Bayes'sche Theorem zu erweitern. Durch die Koppelung der Markov-Ketten gelingt es, einen numerischen Ansatz für die Aktualisierung von Verteilungsfunktionen zu finden, bei welchen keine analytische Lösung möglich ist.

Andere Ansätze, wie die von Song und Eldin (2012), die die Methode der vorausschauenden Planung (Look-ahead scheduling) erforschten, sind meist auf einzelne Anwendungsfälle begrenzt und daher nicht für alle Anwendungen geeignet (Wu et al., 2019).

In diesem Beitrag wird ein ganzheitlicher Ansatz speziell für den Tunnelbau vorgestellt, welcher die Aktualisierung der Eingangsdaten unter Berücksichtigung von tunnelbauspezifischen Besonderheiten der Eingangsparameter ermöglicht.

3 Simulationsbasierte Prognoseanpassung

In vorangegangener Forschung wurde ein Ansatz zum flexiblen Aufbau von ereignisdiskreten Simulationsmodellen zur Analyse von Produktionsvorgängen und Störungen im maschinellen Tunnelbau in der Planung entwickelt und in der Simulationssoftware AnyLogic implementiert (Scheffer et al., 2016; Rahm et al., 2015; Conrads et al., 2018). Dieser Simulationsansatz wird nun für eine baubegleitende Anwendung erweitert. Es wird ein Konzept zur kontinuierlichen Aktualisierung der Leistungsprognose eines Vortriebs während der Ausführungsphase vorgestellt und die Aktualisierung stochastisch verteilter Eingangsdaten thematisiert.

3.1 Integration des aktuellen Baufortschritts

Für eine baubegleitende, simulationsbasierte Prognoseanpassung ist eine Schnittstelle zwischen den aktuellen Ist-Daten der Baustelle und der Prozesssimulation erforderlich. Das entwickelte Konzept sieht dafür vor, die Ist-Daten der Baustelle, die i. d. R. in Echtzeit im Prozesscontrolling-System gespeichert werden, durch eine API-Anfrage eines online verfügbaren Servers kontinuierlich abzufragen und in einer separaten Datenbank zu speichern (Abb. 2). Eine zwischengeschaltete Datenverarbeitung greift auf diese Datenbank zu und ermöglicht eine Bearbeitung und Auswertung der projektspezifischen Ist-Daten. Sie bietet weiterhin die Möglichkeit, die stochastisch verteilten Eingangsdaten der Simulation durch die projektspezifischen Ist-Daten zu aktualisieren. Der aktuelle Baufortschritt wird über eine Verknüpfung der Parameter des Simulationsmodells mit der Datenbank automatisch und echtzeitnah integriert. Die Schnittstelle wurde von Salloum et al. (2020) verifiziert.

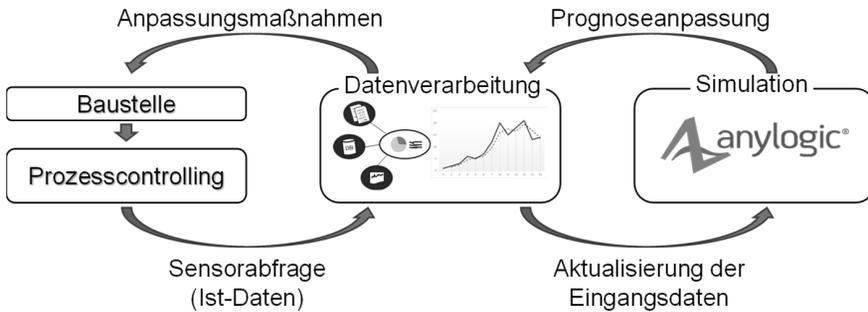


Abbildung 2: Schematischer Ablauf der vortriebsbegleitenden, simulationsgestützten Prognoseanpassung

Des Weiteren ist vorgesehen, dass die aktualisierte Prognose an die Datenverarbeitung zurückgegeben wird. Durch deren Analyse können neue Erkenntnisse gewonnen und ein vortriebsbegleitendes Störfallmanagement unterstützt werden.

3.2 Anpassung der Eingangsdaten

Die Integration des aktuellen Baufortschritts in eine Simulation ist der erste Schritt zur Aktualisierung einer Leistungsprognose. Die Auswertung der projektspezifischen Ist-Daten des bisherigen Bauablaufs erlaubt außerdem eine Verbesserung stochastisch verteilter, in der Planung angenommener Eingangsdaten der Simulation.

Bei der Datenauswertung müssen die vorliegenden Ist-Daten gewisse Anforderungen erfüllen. Wichtige Anforderungen, die das VDI Zentrum Ressourceneffizienz formuliert, sind Aktualität, Nachvollziehbarkeit, Konsistenz, Eindeutigkeit und Vergleichbarkeit der Daten. Eine stetige Überprüfung der Daten auf diese Anforderungen ist unerlässlich.

Für die Anpassung der Eingangsdaten in dieser Arbeit werden die während des Bauablaufs erfassten Ist-Daten unter Berücksichtigung tunnelbau-spezifischer Kenntnisse statistisch ausgewertet (Abb. 3). Zunächst müssen hierfür die zu aktualisierenden Eingangsparameter gewählt werden. Eine vorherige Experteneinschätzung oder Sensitivitätsanalyse ist empfehlenswert für die Bestimmung der Parameter, welche den größten Einfluss auf das zu untersuchende System haben und demnach aktualisiert werden sollten. Eine Aktualisierung aller Parameter wird aufgrund der hohen Anzahl an Eingangsdaten nicht empfohlen (Salloum et al., 2020).

Die Bestimmung der Abweichung der Ist-Daten zu den zu aktualisierenden Eingangsdaten bildet den nachfolgenden Schritt. Tritt eine Abweichung auf, wird für jeden Parameter separat eine geeignete Anpassungsstrategie gewählt.

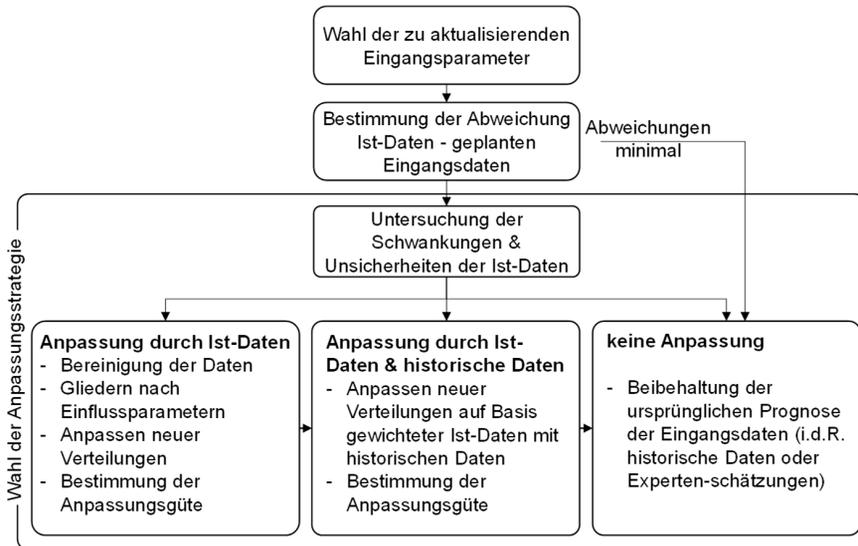


Abbildung 3: Baubegleitende Anpassung der Eingangsdaten anhand von aktuellen Ist-Daten

Im Tunnelbau sind viele Eingangsdaten abhängig von projektspezifischen Randbedingungen, die sich während des Vortriebs ändern können (bspw. Baugrundverhältnisse). Diese Einflussfaktoren müssen für die Anpassung der Eingangsdaten bestimmt und die Parameter den Einflussfaktoren entsprechend gegliedert angepasst werden. So ist es sinnvoll, die Vortriebsgeschwindigkeit für jede Bodenschicht einzeln zu aktualisieren und nicht pauschal für das gesamte Projekt eine Anpassung vorzunehmen. Eine vorherige Bereinigung der Daten um Werte, die auf einen Sensorausfall (Messfehler) schließen lassen, ist außerdem sinnvoll. Fehlenden Werte („missing values“) und Ausreißer in den Datensätzen gehören zu den nachträglich aus den Daten identifizierbaren Messfehlern. Um die Bereinigung von „echten“ Ausreißer, also Extremwerten zu vermeiden, schlagen die Autoren vor plausible Grenzwerte für die jeweiligen Parameter zu definieren und die Daten anhand dieser zu bereinigen.

Die Wahl der Anpassungsstrategie ist maßgeblich von der Verlässlichkeit der Ist-Daten, Schwankungen und Extremwerten sowie der Größe der erfassten Datenmenge abhängig und ist für jeden Parameter einzeln durchzuführen. Als Mindestgröße der Datenbasis sind in der Literatur Anhaltswerte von prozentual 9 % der gesamten Projektdaten bzw. absolut 50 Daten zu finden (Chung T.H. et al. 2004). Wichtig ist, dass die Datenbasis erwartete Schwankungen und Extremwerte der jeweiligen Prozessdaten enthalten. Die Berücksichtigung dieser sind für eine robuste Prognose und Planung sowie eine Abbildung von Leistungsspitzen essenziell. Ist für einen Parameter eine große Datenbasis vorhanden, die eine verlässliche Prognose des weiteren Projektablaufs erlaubt, können für die Aktualisierung der Eingangsdaten neue Verteilungsfunktionen rein auf Basis der Ist-Daten generiert werden. Die Anpassung neuer Verteilungsfunktionen an die Ist-Daten ist dabei kontinuierlich von der Überprüfung der Anpassungsgüte durch Anpassungstests (Goodness-of-fit tests)

zu begleiten. Bei einer unzureichenden Anpassungsgüte sollte eine andere Anpassungsstrategie gewählt werden.

Ist die Datenbasis für eine ausreichend gute Anpassung zu gering, weisen die Ist-Daten zu große Schwankungen auf oder besitzen diese keine verlässliche Aussagekraft für den weiteren Projektablauf, müssen die Ist-Daten für eine robuste Aktualisierung um historische Daten oder Expertenschätzungen ergänzt werden.

Für manche Parameter wird empfohlen, trotz Abweichungen die ursprünglich angenommenen Verteilungen beizubehalten. Dies trifft auf Parameter zu, die selten auftreten oder dessen Einflussparameter großen Schwankungen unterliegen, sodass keine verlässliche Datenbasis für eine Anpassung besteht.

4 Anwendungsbeispiel

Zur Demonstration und Evaluation des vorgestellten Konzepts wird der Vortrieb eines 2000 m langen Tunnels mit einem Durchmesser von 9 m, welcher mit einer TBM mit flüssigkeitsgestützter Ortsbrust aufgeföhren wird, simuliert. Der Vortrieb erfolgt in quartären Kies-Sand-Böden und durchfährt drei noch nicht ausgehobene Haltestellen. Eine Übersicht des Tunnels mit den zu durchörternden Bereichen ist Abbildung 4 zu entnehmen.

Im Rahmen des Anwendungsbeispiels werden die Daten eines abgeschlossenen Referenzprojekts verwendet. Die Eingangsdaten der Simulation basieren auf Daten aus der Vorplanung, auf historischen Daten und ingenieurtechnischen Annahmen. Für die Nachbildung der baubegleitenden Simulation wurden die Daten des Referenzprojekts in dem Rhythmus, in welchem sie in Echtzeit erfasst wurden, mithilfe von API-Anfragen aus einem Prozesscontrolling-System abgefragt und in Abhängigkeit des bekannten Baufortschritts in definierten Zeitabständen als quasi Echtzeiteingangswerte verwendet. So wurden die Randbedingungen einer baubegleitenden Simulation mit Echtzeit-Datenanbindung nachgebildet und das Beispiel konnte für die Evaluierung des entwickelten Konzepts herangezogen werden.

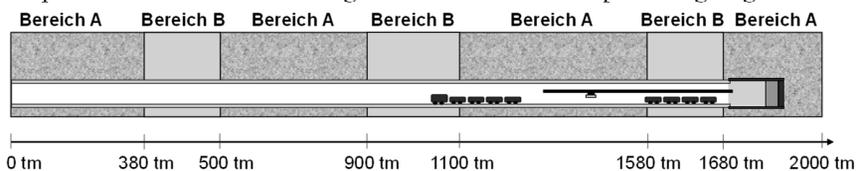


Abbildung 4: Übersicht des Anwendungsbeispiels mit Vortrieb zwischen den Haltestellen (Bereich A) und Durchörterung der Haltestellen (Bereich B)

Für die Anpassung der Eingangsdaten werden die Prozessdauern der beiden Kernprozesse Vortrieb und Ringbau in drei Schritten aktualisiert und der aktuelle Baufortschritt in das Simulationsmodell integriert.

Das Anwendungsbeispiel wurde unter Verwendung eines Multimethodenansatzes in der Simulationssoftware AnyLogic (The AnyLogic Company) implementiert. Eine detaillierte Beschreibung des verwendeten Modells ist bei Salloum et al. (2020) zu finden. Die Ergebnisse der Simulationsstudie werden zur Berücksichtigung der Unsicherheiten der Eingangsdaten durch Monte-Carlo-Simulationen erzeugt.

4.1 Aktualisierung des Parameters Ringbaudauer

Die Ringbauzeit ist im Wesentlichen abhängig von dem Erfahrungs- und Einarbeitungsgrad des Personals und von projektspezifischen Randbedingungen wie bspw. der Anzahl und Größe der Tübbinge je Ring oder der Präzision und Geschwindigkeit des Erektors (Maidl und Wingmann, 2009). Die projektspezifischen Faktoren sind innerhalb eines Projekts konstant. Der Einarbeitungsgrad verändert sich jedoch mit Projektfortschritt und muss bei der Aktualisierung berücksichtigt werden. Abhängig vom Einarbeitungsgrad müssen die aktuellen Ist-Daten gefiltert und entsprechend angepasst werden.

Maidl und Wingmann (2009) nennen Richtwerte für Ringbauzeiten in Abhängigkeit der Tunnellänge basierend auf Referenzprojekten. Angelehnt an diese Werte werden in diesem Anwendungsbeispiel vereinfacht drei Bereiche definiert, die den Einarbeitungseffekt widerspiegeln. Der Einarbeitungsbereich 1 wird für den Vortrieb von 0 bis 200 Tunnelmetern (tm) definiert, der zweite Einarbeitungsbereich wird bis zu einer Tunnellänge von 500 tm definiert. Ab einer Tunnellänge von 500 tm wird angenommen, dass die Einarbeitung abgeschlossen ist und ein Regelbetrieb stattfindet.

Für die einzelnen Aktualisierungsschritte wurden die zu dem jeweiligen Zeitpunkt vorliegenden Ist-Daten in diese drei Bereiche unterteilt und Ausreißer, die auf einen Sensorausfall schließen lassen, bereinigt. Zur Erkennung von Ausreißern wurden vorab Grenzwerte für plausible Prozesszeiten definiert. Die Grenzwerte berücksichtigen dabei mögliche Extrema. Anschließend wurden neue Verteilungsfunktionen für die jeweiligen Bereiche an die aktuellen Ist-Daten angepasst (Tab. 1) und die Anpassungsgüte durch Anpassungstests (Chi²-Test, Kolmogorov-Smirnov-Test, Anderson-Darling-Test) überprüft. Die Leistungsprognose wurde dann durch die Integration des aktuellen Baufortschritts sowie der aktualisierten Verteilungen in der Simulation aktualisiert.

Tabelle 1: Anpassung des Parameters Ringbaudauer [min] im Einarbeitungsbereich 1 (bis 200 tm), Einarbeitungsbereich 2 (bis 500 tm) und Regelbetrieb (ab 500 tm)

Aktualisierungsschritt		Verteilung	Parameter
Planung	Bereich 1:	Weibull (E)	loc = 20.78; $\alpha = 3.399$; $\beta = 53.87$
	Bereich 2:	Weibull	loc = 0.689; $\alpha = 3.406$; $\beta = 53.975$
	Regelbetrieb:	Weibull	loc = 16.776; $\alpha = 13.035$; $\beta = 2.152$
1. Update (500 tm)	Bereich 1:	Log-Logistic(E)	loc = 19.547; $\alpha = 24.459$; $\beta = 2.486$
	Bereich 2:	Log-Logistic(E)	loc = 16.776; $\alpha = 13.035$; $\beta = 2.152$
	Regelbetrieb:	keine Anpassung	-
2. Update (1000 tm)	Bereich 1:	s. Update 1	-
	Bereich 2:	s. Update 1	-
	Regelbetrieb:	Lognormal	$\mu = 2.532$; $\sigma = 0.830$; min = 15.108
3. Update (1500 tm)	Bereich 1:	s. Update 1	-
	Bereich 2:	s. Update 1	-
	Regelbetrieb:	Johnson SB	LEP = 14.883; UEP = 505.005; $\alpha = 4.552$; $\beta = 1.236$

4.2 Aktualisierung des Parameters Vortriebsgeschwindigkeit

Die Vortriebsgeschwindigkeit wird maßgeblich von den geologischen Randbedingungen und den angetroffenen Baugrundverhältnissen beeinflusst. Die Anpassung der Vortriebsgeschwindigkeit erfolgt demnach je Bodenschicht.

In dem Anwendungsbeispiel werden auf der freien Tunnelstrecke zwischen den Haltestellen vergleichbare Bodenverhältnisse angenommen. Bei der Durchörterung der Haltestellen ist jedoch eine erhebliche Reduzierung der Vortriebsgeschwindigkeit vorgeschrieben. Daraus ergeben sich zwei Bereiche, für die jeweils zwei unterschiedliche Verteilungen der Vortriebsgeschwindigkeit angenommen werden.

Äquivalent zu dem Vorgehen der Anpassung der Ringbaudauer werden für die Aktualisierungsschritte die aktuellen Ist-Daten der Vortriebsgeschwindigkeit in die Bereiche A (zwischen den Haltestellen) und B (Durchörterung der Haltestellen) unterteilt und die Verteilungsfunktionen für diese Bereiche je nach bekanntem Baufortschritt aktualisiert. Die aktualisierten Verteilungen für die Anpassungsschritte sind in Tabelle 2 dargestellt.

Tabelle 2: Anpassung der Eingangsdaten Vortriebsgeschwindigkeit [mm/min] im Bereich A zwischen den Haltestellen und Bereich B Durchörterung der Haltestellen

Aktualisierungsschritte	Verteilung	Parameter
Planung	Bereich A: Beta	$\alpha_1 = 15.690$; $\alpha_2 = 10.426$; LEP = 20.639; UEP = 69.518
	Bereich B: Beta	$\alpha_1 = 15.639$; $\alpha_2 = 10.396$; LEP = 0.688; UEP = 49.493
1. Update (500 tm)	Bereich A: JohnsonSB	$\alpha_1 = -1.152$; $\alpha_2 = 1.170$; LEP = 0.075; UEP = 56.835
	Bereich B: Log-Logistics(E)	loc = 12.172; $\alpha = 11.759$; $\beta = 4.135$
2. Update (1000 tm)	Bereich A: Weibull	loc = 0; $\alpha = 44.841$; $\beta = 6.186$
	Bereich B: Weibull (E)	loc = 4.458; $\alpha = 17.220$; $\beta = 2.490$
3. Update (1500 tm)	Bereich A: Weibull (E)	Loc = 13.360; $\alpha = 30.433$; $\beta = 4.521$
	Bereich B: s. 2. Update	-

4.3 Ergebnisse

Für die Auswertung der einzelnen Aktualisierungsschritte werden die Summen der Vortriebsdauer, der Ringbaudauer und die prognostizierte Projektzeit gegenübergestellt (Tab. 3).

Tabelle 3: Ergebnisse der Simulationsstudie

Aktualisierung	Vortriebsdauer [d]	Ringbaudauer [d]	Projektzeit [d]
Prognose Planung	33,02	29,14	121,04
1. Update (500 tm)	46,05	27,38	155,86
2. Update (1000 tm)	44,11	25,76	270,10
3. Update (1500 tm)	43,23	24,82	291,37
Ist-Daten	42,77	23,63	305,00

Der Vergleich der Aktualisierung der prognostizierten Projektzeit zeigt eine signifikante Verbesserung bereits ab dem zweiten Aktualisierungsschritt. Das Ausmaß der Verbesserung ist nicht nur auf die Aktualisierung der Parameter Vortriebgeschwindigkeit und Ringbaudauer zurückzuführen, sondern auch auf die Integration des aktuellen Baufortschritts in die Simulation. Mit fortschreitender Projektzeit reduzieren sich zunehmend die Unsicherheiten der weiteren Prognose, unvorhergesehene Stillstände finden in der Prognose außerdem nachträglich Berücksichtigung.

Der reine Effekt der Anpassung der Verteilungsfunktionen für die Ringbaudauer und die Vortriebgeschwindigkeiten zeigt sich bei der Prognose der jeweiligen Prozessdauern. Dieses Ergebnis unterstreicht, dass nicht nur eine Integration des aktuellen Baufortschritts, sondern auch die gezielte, baubegleitende Anpassung der Eingangsdaten einer Simulation essenziell für eine verbesserte Prognose ist.

5 Zusammenfassung und Ausblick

Die Aktualisierung eines Simulationsmodells mit aktuellen Ist-Daten während des Bauablaufs ermöglicht eine kontinuierliche Verbesserung der Leistungsprognose. In diesem Paper wurde ein Konzept vorgestellt, welches die baubegleitende Aktualisierung einer Simulation mit aktuellen Ist-Daten unter Berücksichtigung tunnelspezifischer Kenntnisse ermöglicht. Ein erstes Anwendungsbeispiel zeigt die Verbesserung der Leistungsprognose durch die Integration des aktuellen Baufortschritts und die Aktualisierung der stochastischen Eingangsdaten für die Kernprozessdauern Vortrieb und Ringbau.

Für die Integration von Ist-Daten in Simulationen ist eine Datenbasis, die vortriebsbegleitend in Echtzeit verfügbar ist, essenziell. Die Heterogenität der Datenquellen im Tunnelbau (Maschinendaten, Schichtprotokolle etc.) stellt hier immer noch eine Herausforderung dar. Durch die fortschreitende Digitalisierung der Baustellen werden jedoch zunehmend digitale, offene Lösungen für die Organisation und Speicherung heterogener Daten entwickelt. Hegemann et al. (2020) stellen bspw. einen integralen Ansatz zur Verknüpfung unterschiedlicher Datenquellen und Bereitstellung der Daten vor. Für weitere Forschung sollte die Kopplung eines solchen Ansatzes an die baubegleitende Simulation untersucht werden.

Die Aktualisierung weiterer Parameter und die Untersuchung der Auswirkungen dieser auf das Gesamtsystem stellen die nächsten Forschungsziele dar. Der Einsatz der baubegleitenden Simulation zur Unterstützung eines vortriebsbegleitenden Störfall-managements wird außerdem angestrebt.

Danksagung

Die Autoren danken der Deutschen Forschungsgesellschaft (DFG) für die finanzielle Förderung dieser Arbeit im Rahmen des Teilprojekts C3 des Sonderforschungsbereichs SFB 837 „Interaktionsmodelle für den maschinellen Tunnelbau“.

Literatur

- Akhavian, R.; Behzadan, A.H.: Knowledge-Based Simulation Modeling of Construction Fleet Operations Using Multimodal-Process Data Mining. *Journal of Construction Engineering and Management* 139 (2013) 11.
- Al-Bataineh, M.; AbouRizk, S.; Parkis, H.: Using Simulation to Plan Tunnel Construction. *Journal of Construction Engineering and Management* 139 (2013) 5, S. 564–571.
- Bradley, R.; Bergmann, J.J.; Noble, J.S.; McGarvey, R.G.: Evaluating a bayesian approach to demand forecasting with simulation. *Proceedings of the 2015 Winter Simulation Conference* (2015).
- Chung T.H.; Mohamed Y.; AbouRizk S.M. (Hrsg.): *Simulation Input Updating using Bayesian Techniques* 2004.
- Conrads, A.; Scheffer, M.; König, M.; Thewes, M.: Robustness evaluation of cutting tool maintenance planning for soft ground tunneling projects. *Underground Space* 3 (2018) 1, S. 72–85.
- Ebrahimi, Y.; AbouRizk, S.M.; Fernando, S.; Mohamed, Y.: *Symphony Supply Chain Simulator*. *SIMULATION* 87 (2011) 8, S. 657–667.
- Haas, H.; Einstein, H.H.: Updating the Decision Aids. *Journal of Construction Engineering and Management* (2002) 128, S. 40–48.
- Hegemann, F.; Stascheit, J.; Maidl, U.: Digitalisierung und IoT im maschinellen Tunnelbau – Ein integraler Ansatz. *Geomechanics and Tunnelling* 13 (2020) 4, S. 406–420.
- Herrenknecht AG, 2021: Herrenknecht AG. Online verfügbar unter <https://www.herrenknecht.com/de/produkte/productdetail/mixschild/>, zuletzt geprüft am 05.05.2021.
- Maidl, U.; Wingmann, J.: Leistungsprognose von Erddruckschilden im Lockergestein. *Geomechanics and Tunnelling* 2 (2009) 2, S. 189–197.
- Rahm, T.; Scheffer, M.; Thewes, M.; König, M.; Duhme, R.: Evaluation of Disturbances in Mechanized Tunneling Using Process Simulation. *Computer-Aided Civil and Infrastructure Engineering* 31 (2015) 3, S. 176–192.
- Salloum, Y.; Jodehl, A.; Thewes, M.; König, M.: From Forecasting to real-time process controlling with continuous simulation model updates. *Proceedings of the 37th International CIB W78 Conference* (2020).
- Scheffer, M.; Rahm, T.; König, M.; Thewes, M.: Simulation-Based Analysis of Integrated Production and Jobsite Logistics in Mechanized Tunneling. *Journal of Computing in Civil Engineering* 30 (2016) 5, C4016002.

Song, L.; Eldin, N.N.: Adaptive real-time tracking and simulation of heavy construction operations for look-ahead scheduling. *Automation in Construction* 27 (2012), S. 32–39.

The AnyLogic Company: <http://www.anylogic.com/>: 8.5.2 University.

VDI Zentrum Ressourceneffizienz: Grundlagen zur Echtzeitdatenerfassung und -verarbeitung, VDI Zentrum Ressourceneffizienz. Online verfügbar unter <https://www.ressource-deutschland.de/themen/echtzeitdaten/grundlagen/>, zuletzt geprüft am 02.07.2021.

Wu, L.; Ji, W.; AbouRizk, S.M.: Bayesian Inference with Markov Chain Monte Carlo-Based Numerical Approach for Input Model Updating. *Journal of Computing in Civil Engineering* 34 (2019).



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Simulation-based demand forecast generation to analyze forecast accuracy and its influence on logistical performance

Simulationsbasierte Forecastdaten-Generierung zur Analyse der Forecastqualität und deren Einfluss auf das logistische Potential.

Thomas Felberbauer, St. Poelten UAS, St. Poelten (Austria),
thomas.felberbauer@fhstp.ac.at

Wolfgang Seiringer, Klaus Altendorfer, UAS Upper Austria, Steyr (Austria),
wolfgang.seiringer@fh-steyr.at, klaus.altendorfer@fh-steyr.at

Abstract: Managers face the problem of different demand models and forecast information for their products. Uncertainty and trends in forecasts influence the logistical performance of the companies and the question which planning method or planning parameter level fits best for the respective product is still open for practitioners. This paper gives an overview about practical inspired demand models to mimic different behaviours of customer provided forecasts, and presents a simulation based forecast generation module to analyse and improve forecast accuracy. The applications of this model presented in this paper show that measuring the forecast accuracy, understanding the forecasting behaviour, and applying respective correction methods is an important issue to be focussed in research.

1 Introduction

In manufacturing companies, demand forecast data is used for the underlying production planning. The demand information can either be self-generated with common forecasting techniques or provided by the customers in a rolling horizon. In a make-to-stock production system, manufacturing companies use classical forecast methods to predict the demand, or the customers provide forecast data in a rolling horizon with periodic order amount updates. In a make-to-order setting manufacturing companies use the customer provided order amounts with a specific due date to plan their production. Norouzi and Uzsoy (2014) highlight that the two mentioned make-to-stock forecast systems are not completely contrary. The authors refer to Chen and Lee (2009) who showed that many commonly used time-series forecast models can be interpreted as special cases of the forecast evolution model. In all mentioned systems the provided information is somehow stochastic, and companies are looking for ways how to deal with the underlying uncertainty. Literature on inventory and

production planning shows that understanding, analysing and improving the forecast accuracy has a large improvement potential for production system performance (Enns, 2002; Sanders and Graman, 2009; Metters, 1997; Chen and Lee, 2009). For streamlined production systems and inventory models, analytical models can help to set planning parameters to be more efficient, which means on-time deliveries in the agreed order amount with a minimum of stock levels (Norouzi and Uzsoy, 2014). Nevertheless, for real production systems facing multiple demand uncertainties simulation is an appropriate method to investigate forecast accuracy and its influence on logistical performance. The proposed paper presents results of the research project InnoFIT. The overall project aim is to improve the forecast quality of customer provided forecasts by analysing and visualizing forecast data and developing appropriate mathematical models to correct systematic forecast errors. Due to the project cooperation with four automotive suppliers, the presented demand models and the developed forecast correction models are inspired by real world problems and partly evaluated with real company data.

Contribution one of this paper is the introduction and explanation of a discrete event simulation model module to generate forecast data with systematic and unsystematic forecast effects. The developed simulation model module enables to mimic the evolution of customer provided forecasts over time. In the forecast evolution model, forecasts are not static, but forecast updates appear in a rolling horizon until the due date occur. Such effects can be observed in real data often disturbed with other uncertainty effects. However, to study different information sharing types and their interaction and influence on production planning, framed data sets need to be generated where simulation can be used. Contribution two of the paper shows the output of the mentioned research project where the simulation model module is used in different publications and its contribution to scientific progress is discussed.

This paper is organized as follows. In Section 2 the demand models, the forecast accuracy measures, and one forecast correction method are introduced which are implemented in the simulation model module. Section 3 presents the simulation model application and selected results. Finally, in Section 4 we summarize the results of the presented study and mention some interesting topics for future research.

2 Demand Models and Simulation Model Implementation

To enable a thorough discussion and analysis of customer provided forecast behaviours, a simulation model module that mimics the forecast evolution demand process has been developed and is introduced in this section. The software AnyLogic© is used to build the simulation model. On the one hand, this model can be applied for static forecast data generation to evaluate the forecast accuracy and the improvement potential of forecast correction methods. On the other hand, this model can be applied as simulation module for production planning simulation to evaluate the forecast accuracy and investigate its influence on production order accuracy and key performance indicators such as service level, tardiness, and inventory. The following subsections introduce the forecast evolution demand process which is extended by several practically observed forecasting behaviours. Furthermore, a few forecast accuracy measures, which have been applied in the different studies from the InnoFIT project are introduced. Additionally, one forecast correction method has been

developed and is implemented in the simulation model and described in the following section.

2.1 Demand model development

To simulate real forecasting behaviour the following demand model is developed.

2.1.1 Basic forecast evolution model

The idea of this basic forecast evolution demand model is rather simple. Customers provide a constant long-term forecast \tilde{x}_i for due date i in the future and start to update this forecast with a due date horizon H . This means the first update occurs with $j = H$ where j are the periods before delivery, and then each period an update occurs. For a single item with normal distributed update terms this can be written as:

$$\begin{aligned} x_{i,j} &= x_{i,j+1} + \varepsilon_{i,j}(\tilde{x}_i, 0, \alpha) \\ \varepsilon_{i,j}(\tilde{x}_i, 0, \alpha) &= N(0, \alpha \alpha_j \tilde{x}_i) \end{aligned} \quad (1)$$

Whereby $x_{i,j}$ denotes the forecast for due date i stated j periods in advance and $\varepsilon_{i,j}$ is the normal distributed forecast update for due date i generated j periods in advance. The parameter α_j is applied to model the stochastic error related to periods before delivery and α is applied to enable an easy change in the level of uncertainty. In the following presented studies the α_j vector is usually kept constant and only α is varied. Note that the original martingale model of forecast evolution as applied in Norouzi and Uzsoy (2014) includes multivariate normal distributed updates with possible relations between the items. This is not included in our simulation model since the focus is more on extending this model by practically observed behaviours. If a pure customer required lead time setting is to be modelled, then only one update occurs with j being the customer required lead time. The basic forecast evolution model is used in the simulation studies Felberbauer and Altendorfer (2014), Zeiml et al. (2019), and Zeiml et al. (2020).

2.1.2 Biased forecasts

The basic forecast evolution model assumes that each update includes an information gain, i.e. the new updated value has on average less distance to the final order amount. Since the mean of the forecast updates $\varepsilon_{i,j}$ is zero, the updates can be positive or negative. However practical observations as well as supply chain literature suppose that there might be a customer which provide biased forecast information for example due to rationing (Lee and Whang, 2000). This leads to the following demand model:

$$\begin{aligned} x_{i,j} &= x_{i,j+1} + \varepsilon_{i,j}(\tilde{x}_i, \beta, \alpha) \\ \varepsilon_{i,j}(\tilde{x}_i, \beta, \alpha) &= N(\beta b_j \tilde{x}_i, \alpha \tilde{x}_i) \end{aligned} \quad (2)$$

The relation between forecast bias and periods before delivery is included with b_j while the level of forecast error is parametrized with β . Note that in the following presented studies the b_j vector is usually kept constant and only β is varied. In this

modelling frame the forecast bias can either be positive or negative and it is also possible to model long-term forecast errors, i.e. the long-term forecast might be too high or too low. A simulation model with biased forecasts is investigated in the papers Zeiml et al. (2019) and Zeiml et al. (2020).

2.1.3 Temporary Outliers

A further customer behaviour observed is that planning instabilities in the customer's planning process lead to changes in the demand forecasts that can be high in their amount, but occur only temporary. Such a behaviour is called temporary outliers in our demand model. Note that such temporary outliers imply that again the information quality does not improve with each update. The following modelling is applied:

$$\begin{aligned}
 x_{i,j} &= x_{i,j+1} + \varepsilon_{i,j}(\tilde{x}_i, \beta, \alpha) + P_{i,j}(\gamma c_j) \lambda_{i,j}(\tilde{x}_i, \delta, e) \\
 &- P_{i,j+v}(\gamma c_j) \lambda_{i,j+v}(\tilde{x}_i, \delta, e) \\
 \lambda_{i,j}(\tilde{x}_i, \delta, e) &= N(\delta \tilde{x}_i, \delta e \tilde{x}_i) \\
 P_{i,j}(\gamma c_j) &= 1 \text{ with Probability } \gamma c_j \text{ and } P_{i,j}(\gamma c_j) = 0 \text{ otherwise}
 \end{aligned} \tag{3}$$

In this formulation, an outlier occurs j periods before delivery with probability γc_j , whereby again c_j includes the link to the periods before delivery and γ defines the level of occurrence probability. The amount of an outlier is normally distributed with mean $\delta \tilde{x}_i$ and standard deviation $\delta e \tilde{x}_i$, i.e. the level of outliers is defined with δ while its stochastic is defined with e . The parameter v indicates for how much periods an outlier occurs (Seiringer et al., 2021).

2.1.4 Demand shifting between periods

Looking at real data sets from the company partners of the research project uncovered a further customer behaviour that is included in our demand model. In this setting, the demands are shifted between the single due dates, e.g., demand for due date i is reduced by a certain amount and demand of due date $i+1$ is increased by the same amount. Such a behaviour can be based on the assumption of cumulated demand being known by the customers, but the respective due dates may change. The following model includes this behaviour.

$$\begin{aligned}
 x_{i,j} &= x_{i,j+1} + \varepsilon_{i,j}(\tilde{x}_i, \beta, \alpha) + P_{i,j}(\gamma c_j) \lambda_{i,j}(\tilde{x}_i, \delta, e) \\
 &- P_{i,j+v}(\gamma c_j) \lambda_{i,j+v}(\tilde{x}_i, \delta, e) \\
 &+ \sum_{m \in \{-r, \dots, -1, 1, \dots, r\}} Y_{i+m, -m}(\psi f_j) l_{j+m} \mu_i - \sum_{k \in \{-r, \dots, -1, 1, \dots, r\}} Y_{i,k}(\psi f_j) l_j \mu_i \\
 Y_{i,k}(\psi f_j) &= 1 \text{ with Probability } \psi f_j \text{ and } Y_{i,k}(\psi f_j) = 0 \text{ otherwise}
 \end{aligned} \tag{4}$$

According to this formulation, a demand shift occurs with probability ψf_j again with ψ stating the general level of amount shifts and f_j stating the relation to periods before delivery. Furthermore, a demand shift can only occur in a range of r periods and the l_j identifies the amount of demand that is shifted. This practical extension finalizes the

demand model that is generated in the simulation model to mimic real customer forecasting behaviour.

2.2 Forecast Accuracy Measures

To evaluate the forecast accuracy, several forecast error measures could be applied. In the study of Zeiml et al. (2019) several forecast error measures are introduced whereby the normalized root mean squared error related to j periods before delivery $RMSE_j$ is found to be a good applicable measure for identifying the general forecast accuracy. To evaluate systematic behaviours, the mean-percentage-error MPE_j for j periods before delivery is suggested there. The following equation shows these measures.

$$MPE_j = \frac{\sum_{i=1}^n (x_{ij} - x_{i,0})}{\left(\frac{1}{n} \sum_{i=1}^n x_{i,0}\right)} RMSE_j = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ij} - x_{i,0})^2}}{\frac{1}{n} \sum_{i=1}^n x_{i,0}} \quad (5)$$

In most of the papers presented above these measures have been applied and in some papers a forecast correction method is developed (see next subsection). There also a correction effectiveness is discussed calculated as $E_j = \frac{(RMSE_j - CRMSE_j)}{RMSE_j}$. Note that $CRMSE_j$ is the normalized root-mean-squared-error when the respective correction is applied.

2.3 Forecast bias correction

Since a forecast bias is a systematic behaviour that a customer shows and this bias can be measured very well with the MPE_j , in Zeiml et al. (2020) a respective forecast correction method is developed that is also included in the simulation model. The following equation shows how forecasts are corrected applying this method:

$$\hat{x}_{i,j} = \begin{cases} \frac{x_{i,j}}{1 + MPE_j} : \sigma_{MPE_j} \leq \sigma_t \\ x_{i,j} : \text{otherwise} \end{cases} \quad (6)$$

According to this formulation, the forecast is corrected to $\hat{x}_{i,j}$ by applying the MPE_j whenever the standard deviation of the respective MPE_j is below a certain threshold. This implies that such a correction should only be applied if the bias is stable.

3 Model applications and selected results

In this section the most important results of simulation studies and the respective publications are presented that apply the simulation model module to discuss forecast accuracy effects and the performance of the presented correction model. Note that the contribution of this paper lies in providing a comprehensive view on the potential of

applying such a simulation model of forecasting behaviour to tackle different aspects of the respective process.

3.1 Demand model features and its influence on RMSE

Our first results presented in Figure 1 show different demand model features, which are presented in Section 2.1.1-2.1.4, i.e. basic forecast evolution model, biased forecasts, temporary outliers, and demand shifting between periods, and its influence on normalized root mean squared error.

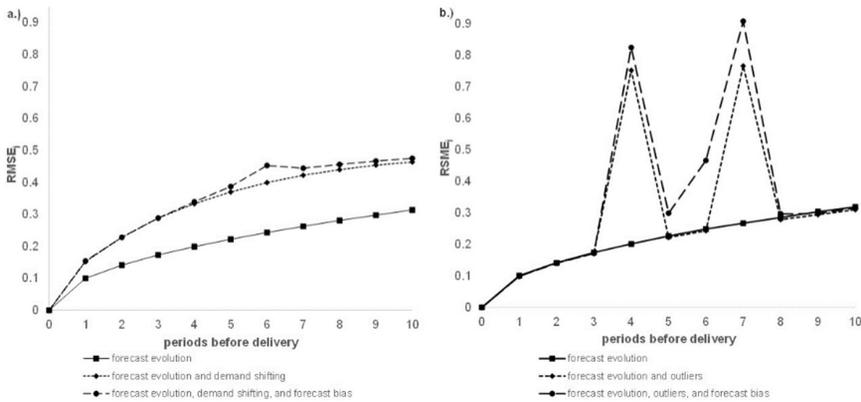


Figure 1: Demand model features and its influence on RMSE

Figure 1 shows the RMSE with respect to periods before delivery for the demand model features. The forecast accuracy is good if RMSE value is low. Figure 1a shows the indicator for basic forecast evolution demand feature and when forecast bias and demand shifting disturbances are added. Figure 1b shows RMSE for the demand models with outliers and forecast bias again compare to the basic forecast evolution scenario. Results confirm the intuitive finding that the more disturbances features are present the lower forecast accuracy is and the higher the respective RMSE value. Comparing basic forecast evolution with forecast evolution and demand shifting shows that for all periods before delivery the RMSE value is higher with demand shifting. Additionally, Figure 1a shows for the RMSE values of the scenario with forecast evolution, demand shifting and forecast bias a peak in the respective periods before delivery where the forecast bias occurs $j \in \{4,5,6,7,8\}$. Figure 1b shows higher RMSE values in periods where outliers happen $j \in \{4,7\}$ compared to the basic forecast evolution model. When in addition to basic forecast evolution and outliers also the demand model mimics forecast bias results of RMSE is again higher than in the scenario with basic forecast evolution model and outliers.

3.2 Customer-required-lead-time versus forecast evolution behaviour

Felberbauer and Altendorfer (2014) apply a first version of the simulation model module to compare the forecast evolution behaviour, where customers provide a forecast quantity for a specific due date for a long horizon in advance and update their

forecast quantities periodically, with a customer-required-lead-time-behaviour, where customers demand stochastic order amounts with a stochastic customer-required-lead-time in advance. In this study the authors assume that the customer required lead times are usually shorter than the forecast evolution horizon, but order quantities do not further change. To compare the influence of both order behaviours for a very streamlined hierarchical production system, a simulation study is performed where among other KPIs the production system overall costs (that are the sum of inventory, tardiness and capacity costs) are discussed.

Overall Costs												
	forecast quality measure											
	0.13	0.16	0.18	0.21	0.24	0.27	0.29	0.32	0.35	0.37	0.39	∅
FEV	12,856	12,899	12,966	13,008	13,082	13,159	13,269	13,382	13,492	13,627	13,762	13,227
CRL	12,475	12,575	12,806	13,289	13,881	14,540	15,202	16,036	16,496	16,803	17,929	14,730

Figure 2: Overall costs for forecast evolution (FEV) and customer required lead time (CRL) behaviour with respect to forecast quality (Felberbauer and Altendorfer, 2014)

Figure 2 shows that for both, forecast evolution (FEV) and customer required lead time order behaviour, the overall costs increase with respect to decreasing forecast quality. In this study the higher the forecast quality measure value the lower the forecast accuracy. Nevertheless, in this study the authors can verify that the hierarchical production planning approach can mitigate the forecast uncertainty better for the forecast evolution behaviour. This means that the average overall costs for all tested forecast quality measures are 11.4% higher for the customer required lead-time behaviour compared to the forecast evolution behaviour. This leads to the insight that for the hierarchical production planning approach more frequent order updates are easier to handle than a very limited number of updates.

3.3 Performance of moving average forecasting if forecast bias occurs

In the paper of Zeiml et al. (2019), the simulation model module is applied to compare forecasts that follow the forecast evolution without bias to forecasts that are biased. Firstly, different forecast error measures are introduced and it is shown that MPE_j is a good measure to identify systematic forecast errors, i.e. to show if the customer's forecasts are too high or too low with respect to j periods before delivery. $RMSE_j$ is identified to be a good measure for evaluating the overall level of forecast uncertainty j periods before delivery, however, it also includes the forecast bias induced errors. As stated above, biased forecasts do not imply an information increase with each update and the level of randomness in the stochastic forecast updates significantly influences the value of their respective information. Therefore, it is conjectured that upon a certain level of forecast bias and a certain level of uncertainty, self-generated forecasts, in this setting generated by the moving average method based on the historical final orders, will perform better than the customer provided forecasts. The simulation model module for forecast generation is in Zeiml et al. (2019) applied to generate customer provided forecasts and it is extended by the simple moving average method. The study compares the $RMSE_j$ for customer provided forecasts with

different α and β values with the $RMSE_j$ values when the simple moving average method is applied. Figure 3 shows the threshold above which the simple moving average forecasts lead to lower $RMSE_j$ values for biased forecasts (highest forecast bias is 6 periods before delivery) whereby the variable α on the y-axis represents the unsystematic forecast error level and the three lines correspond to different systematic forecast error values β . The scaling factor of the unsystematic forecast error level in this study $a_j = 1$. The respective results show that for high forecast bias values, it is even for very low unsystematic effects better to apply the simple moving average forecast method. From a managerial point of view this is an interesting result as it shows that the forecast bias significantly diminishes the value of demand forecast information. This result motivates the study presented in the next subsection where the performance of a forecast bias correction is evaluated.

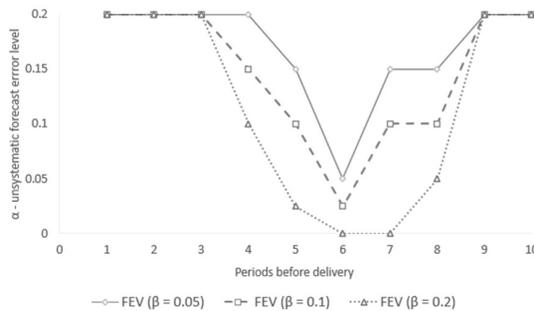


Figure 3: Simple moving average forecasting compared to forecast evolution method (Zeiml et al., 2019)

For the pure unsystematic forecast error scenarios, i.e. $\beta = 0$, the simulation results show that customer provided forecasts always lead to a lower $RMSE_j$ in comparison to the simple moving average method. This is an interesting managerial insight, since it implies that whenever there is no systematic forecast bias using the customer provided and continuously updated forecasts is the best choice.

3.4 Forecast bias correction model

Based on the findings of Zeiml (2019) a forecast correction model (see Section 2.3) is integrated in the simulation forecast module to mitigate systematic forecast errors. The correction model is tested in scenarios with and without unsystematic and systematic errors. Zeiml et al. (2020) presents a decision model to decide in a rolling horizon, whether in the current time period the presented forecast correction model should be applied or not. Therefore, the constant forecast data generation module is extended to generate seasonal demand behaviours. Additionally, we introduce the possibility to parameterize a dynamic planning behaviour, where the systematic bias β can also change over time.

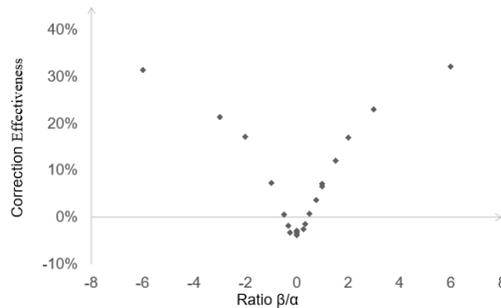


Figure 4: Correction effectiveness with respect to the ratio of the systematic and unsystematic forecast error β/α (Zeiml et al., 2020)

To answer the question about the performance of the forecast correction model Figure 4 give some insights. The forecast correction effectiveness measure (Section 2.2) describes the improvement (positive value) or the deterioration (negative value) of the forecast accuracy comparing the situations with and without the correction model. Results show that the ratio β/α influences the correction effectiveness. The higher $|\beta/\alpha|$, the better the correction mechanism works. Nevertheless, applying the correction model in situations where the unsystematic error α dominates, leads to an increase of RMSE and on the same time reduces forecast accuracy. Additionally, the study of Zeiml et al. (2020) shows that the developed decision model, whether to apply the correction or not, leads to significant forecast quality improvements, compared to its continuous application.

4 Conclusion

This paper presents the status of the simulation model module for forecast data generation to mimic different customer provided forecast behaviours. Basic forecast evolution behaviour, biased forecasts, forecast with outliers and forecasts with demand shifting are implemented. Additionally, forecast error measures are modelled to discuss the influence of the demand model features and its parameterization. The third feature of the simulation model are correction models to improve forecasts based on the information of historical forecast evolutions. The most interesting results of already published simulation studies, which use the presented simulation model module for forecast data generation, are presented. The highlights of the mentioned papers are discussed concerning the differences of forecast accuracy and their performance of the presented correction model with respect to the customer provided forecast behaviours. The contribution of this paper lies in providing a comprehensive view on the potential of applying the simulation model module for forecast data generation to inspire other researchers to contribute in the area of customer provided forecast. After the InnoFIT project it is planned to share the simulation model module for forecast data generation on phaidra, which is a repository for the permanent secure storage of digital assets. After that, researchers will have free access to the simulation model module to further extend advanced forecasting techniques and forecast correction models for customer provided forecasts. Additionally, the simulation model module will be integrated in a generic and scalable simulation model, to test the developed forecast correction models for more complex production systems.

Acknowledgement

The work described in this paper is supported by the Austrian Research Promotion Agency (FFG), “Produktion der Zukunft” via the InnoFIT research project (No. 867471).

References

- Chen, L.; Lee, H.L.: Information Sharing and Order Variability Control Under a Generalized Demand Model. *Management Science* 5 (2009) 55.
- Enns, S.T.: MRP performance effects due to forecast bias and demand uncertainty. *European Journal of Operational Research* 138 (2002) 1, pp. 87–102.
- Felberbauer, T.; Altendorfer, K.: Comparing the performance of two different customer order behaviors within the hierarchical production planning. In: Tolk, A.; Diallo, S.Y.; Ryzhov, I.O.; Yilmaz, L.; Buckley, S.; Miller, J.A. (Hrsg.): *Proceedings of the Winter Simulation Conference 2014*, 2014, pp. 2227–2238.
- Lee, H.L.; Whang, S.: Information sharing in a supply chain. *International Journal of Manufacturing Technology and Management* 1 (2000) 1, pp. 79.
- Metters, R.: Quantifying the bullwhip effect in supply chains. *Journal of Operations Management* 15 (1997) 2, pp. 89–100.
- Norouzi, A.; Uzsoy, R.: Modeling the evolution of dependency between demands, with application to inventory planning. *IIE Transactions* 46 (2014) 1, pp. 55–66.
- Sanders, N.; Graman, G.: Quantifying costs of forecast errors: A case study of the warehouse environment☆. *Omega* 37 (2009) 1, pp. 116–125.
- Seiringer, W.; Altendorfer, K.; Felberbauer, T.: Simulation of stochastic rolling horizon forecast behavior with applied outlier correction to increase forecast accuracy (submitted). In: S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo and M. Loper (Hrsg.): *Proceedings of the 2021 Winter Simulation Conference, Virtual Conference, 2021*,
- Zeiml, S.; Altendorfer, K.; Felberbauer, T.; Nurgazina, J.: Simulation based forecast data generation and evaluation of forecast error measures. In: N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son (Hrsg.): *Proceedings of the 2019 Winter Simulation Conference, National Harbor, MD, USA, 2019*, pp. 2119–2130.
- Zeiml, S.; Seiler, U.; Felberbauer, T.; Altendorfer, K.: Simulation Evaluation of Automated Forecast Error Correction Based on Mean Percentage Error. In: K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing (Hrsg.): *Proceedings of the 2020 Winter Simulation Conference, Online, 2020*,



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Combining Engineering Data, Sensor Data and Artificial Intelligence for Automated Edge Network Infrastructures

Verknüpfung von Engineering-Daten, Sensordaten und künstlicher Intelligenz für automatisierte Edge-Netzwerkinfrastrukturen

André Hürkamp, Franziska Aschersleben, Marvin Czarski, Sebastian Gellrich, Christoph Herrmann, Klaus Dröder, Technische Universität Braunschweig, Institute of Machine Tools and Production Technology, Braunschweig (Germany),
a.huerkamp@tu-braunschweig.de, f.aschersleben@tu-braunschweig.de,
m.czarski@tu-braunschweig.de, s.gellrich@tu-braunschweig.de,
c.herrmann@tu-braunschweig.de, k.droeder@tu-braunschweig.de

Abstract: The paper describes a concept of monitoring and controlling flexible production processes by means of engineering data, sensor data and artificial intelligence for an automatic control by edge-computing. In particular, surrogate models from numerical simulations obtained at the engineering phase are derived using Proper Orthogonal Decomposition and machine learning. Among the six investigated methods, Random Forest yields the best surrogate model in terms of accuracy and real-time feasibility on the edge server for the use case of thermoplastic composites.

1 Introduction

For a real-time capable monitoring and control of flexible production systems, status information on the process and the components at different process steps are needed. Ideally, that information can be obtained without time and cost intensive component testing and without occupying the IT infrastructure with too much raw data.

Decentralized data fusion of real-time capable and valid quality gates at the edge of the factory network can be a solution to overcome these burdens. Process knowledge from available engineering data (i.e., simulation results) together with sensor data and methods of artificial intelligence can be combined to form the basis of virtual quality gates with reduced data streams. Furthermore, the use of quality gates at the individual process and provision of consolidated KPIs at the process chain level (e.g. error probability, source of error, etc.) provides a powerful real-time capable monitoring and control of a flexible production system.

The “Incremental Manufacturing Lab” (IML) of TU Braunschweig is such a flexible manufacturing cell that combines different tool-independent manufacturing technologies. It consists of a welding robot for 3D-printing of steel with high deposition rates, a robot for extrusion of thermoplastics that can also be used for surface machining and handling as well as a milling robot for post-processing. In order to be able to implement a lean solution for improving the part quality and process efficiency, a data processing infrastructure is required that can be universally applied to the process chain, is sufficiently powerful and enables low latencies.

The objective of this contribution is the development of real-time applications on edge devices that are able to learn from simulation data after the engineering phase. In this way, systematic deviations from the simulation model can be corrected online and under real-time conditions with the stream of machine, process, and sensor data. Such an approach is also described as the concept of hybrid twins (Chinesta et al., 2018; Sancarlos et al., 2020).

2 State of the Art

Incremental manufacturing is a concept for a flexible production system that combines additive manufacturing, subtractive manufacturing processes and other production technologies like forming or injection moulding. Low-cost semi-finished parts serve as the basis and functionalities can be added incrementally e.g. by injection moulding, individualized by screw-based extrusion and post-processed by milling. The main advantage over conventional, serially linked manufacturing strategies is the flexibility of the value chain, so that the high customer requirements for individualization of products can be met since the design space is enlarged in contrast to traditional manufacturing technologies and complex geometries can be realized. In addition, the production process is scalable due to the partially tool-free manufacturing, so that a wide variety of materials can be used and product variants can be produced (Reichler et al., 2019; Dröder et al., 2019). However, due to the high flexibility, the process planning, supervision and quality assurance are challenging. Therefore, real-time data is needed for in-process quality monitoring especially for parametrization e.g. for different materials or in changing environments.

In today's product and production engineering, simulation tools are a matter of course to support development processes. Computer simulations are able to analyse the material behaviour of products and processes using numerical methods based on physical principles (Panchal et al., 2013; Allison et al., 2006). However, those simulations are computationally expensive and time consuming (Zimmerling et al., 2020). Due to that, surrogate modelling can be applied to reach computationally feasible models (Han and Zhang, 2012). Applications can be found in the field of biomechanics (Liang et al., 2018), composite manufacturing (Pfrommer et al., 2018; Zambal et al., 2018; Hürkamp et al., 2020a) and in the context of virtual quality gates (Hürkamp et al., 2020b).

The advent of low-cost information and communication technology (ICT) has enabled various industries to employ data gathering and data processing for their production. With the goal to automatically make decisions and predict outcomes regarding the quality and progression of a product (Thiede et al., 2019; Filz et al., 2020). A challenge is the increasing complexity, the high dimensionality of the available data and the mostly chaotic structures. A viable solution for solving these problems is the

application of artificial intelligence (AI) methods, especially with supervised AI models (Wuest et al., 2016). A manufacturer usually employs centralized or cloud computing to train AI models. Since there are limitations imposed by the network infrastructure regarding the data transfer it may not be possible to deploy the AI models, because they require a constant stream of live-data and so it may be necessary to perform on-site signal pre-processing (Wang et al., 2015).

A technology to perform the processing on site is edge computing. The objective of edge computing is to process data as close as possible to the source. With edge computing the computational load can be distributed between edge devices in one network to reduce the computation time. It can extend one manufactures capabilities in the domains: computation, networking and storage with the benefit of reducing latency and bandwidth (Wu et al., 2017).

3 Methods

The proposed method is based on the edge computing concept for flexible manufacturing, which is displayed in Figure 1 on the right side. The foundation is the flexible process chain in the bottom.

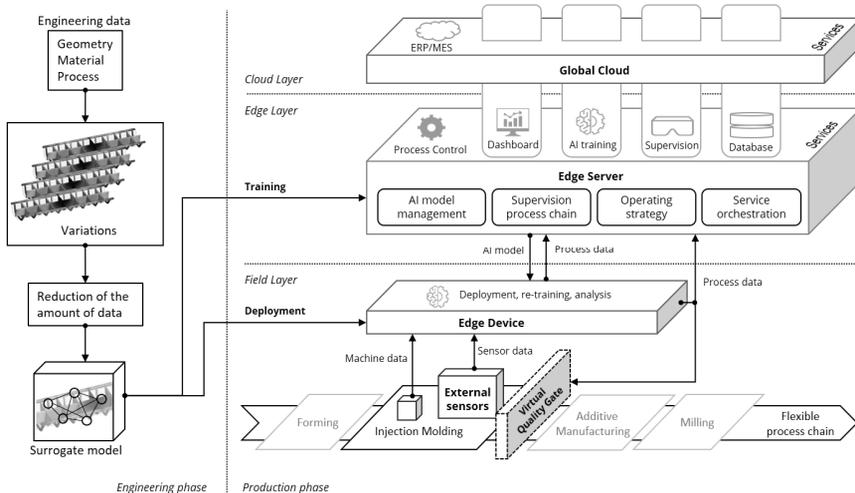


Figure 1: Edge computing concept for flexible manufacturing cell

An exemplary chain is shown with focus on the here utilized injection moulding. Each process step is equipped with external sensors and internal measurement capabilities. These sensor data are then transferred to an edge device that runs a pre-processing to decrease and aggregate the large amount of data to be further processed. AI methods are deployed for a fast and intelligent analysis. Especially simulation data from the engineering phase (Figure 1 on the left side) are used to train a surrogate model that is able to predict process results on an edge device in real time.

The edge device is mounted closely to the process, so that a very low latency can be achieved. Therefore, it is still assigned to the field layer. In the second layer, the

process-superordinate calculations take place on the edge server. The forwarded data from the edge devices are processed to derive parameter adjustments for the process. In addition, the process data is visualized here and the entire process chain is monitored by evaluating virtual quality gates, which are provided after each process step, based on the processed data. Possibilities for compensating for errors in the subsequent process are included here. The edge server is therefore the main computing unit, where also management and training of AI models is performed due to a large computational capacity. In addition, the process data required for training is stored in a database here. All these modules are executed as a service to maximize the flexibility by using Kubernetes. This means that services can be moved to the cloud layer in order to use central computing units, which leads to better resource utilization and but also an increase in latency. The associated orchestration evaluates the requirements for the service, such as the maximum tolerated latency and the computing overhead in order to distribute the services reasonably. This data is subsequently transformed into a real-time capable surrogate model by machine learning that can be applied to edge devices. Embedded into networking interfaces this model can function as the digital representation of the production system (digital twin).

In order to improve this digital twin continuously in the further course of production, suitable sensors provide data with which additional production data is collected. For this purpose, algorithms are being developed to provide robust and reliable automation of flexible manufacturing processes. They will thus provide solutions for automated and decentralized edge infrastructure processes to achieve better process automation and in-line quality control. The goal is to provide methods and infrastructure that can leverage services with high performance and scalability requirements.

3.1 Numerical Modelling and Data Sampling

For the development of the digital twin, numerical parameter studies are performed during the engineering phase. These methods are able to represent the physical interactions of process and material properties and to determine three-dimensional process-dependent product properties. As a use case, the thermoplastic composite structure shown in Figure 2 a) is used. It consists of a continuously fibre reinforced thermoplastic (organo sheet) as a planar part insert and injected polymer ribs on top of it. The structure is designed for the manufacturing of testing specimens in order to experimentally determine the interface bond strength since this is limiting factor in such thermoplastic composites.

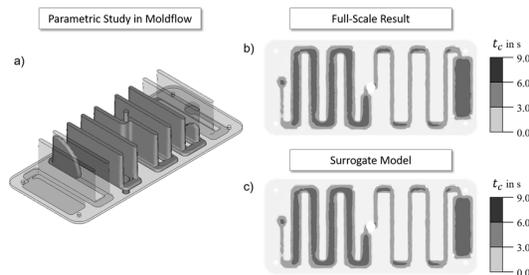


Figure 2: a) Investigated thermoplastic composite structure and exemplary result of b) the full-scale computation and c) the corresponding surrogate

During the engineering phase, potential process and material variations are captured in the resulting data set of numerical studies that cover the entire process parameter space. Within a Latin Hypercube Sampling (LHS), the process parameters part insert temperature $20 \leq T_{insert} \leq 240^\circ\text{C}$, the mould temperature $30 \leq T_{mould} \leq 80^\circ\text{C}$ and the flow rate $10 \leq v_{inj} \leq 100 \text{ cm}^3/\text{s}$ are varied. For the use case, 100 parameter combinations are used as input for simulations of the injection moulding process. As target value, the contact time t_c is computed for each individual simulation. It describes the time the interface between part insert and injected polymer is heated above melting temperature of the used polymer. This variable has been identified to correlate directly with the final bond strength and hence the part quality (Hürkamp et al., 2021). The interface between organo sheet and injected polymer consists of 8196 nodes. Hence, each solution contains 8196 values describing the spatial distribution of the contact time t_c at the interface (cf., Figure 2 b) and c)).

3.2 Combining Proper Orthogonal Decomposition and Data-driven Surrogate Modelling

The LHS produces a large amount of data that needs to be processed. For this purpose, a Proper Orthogonal Decomposition (POD) is used, which on the one hand reduces the amount of data and on the other hand provides a basis for model reduction and surrogate modelling. POD has been proven to be a powerful tool in combination with machine learning for the design of surrogates with physical constraints (Swischuk et al., 2019). The POD of an arbitrary solution vector U can be written as

$$U(x, p) \approx \sum_{i=1}^k \phi_{i(x)} \cdot a_i(p), \quad (1)$$

where $a_i = \phi_i^T u$ denotes the POD coefficients. The vector $\phi_i(x)$ is referred to as spatial mode and contains the geometric information. In contrary, the coefficients $a_i(p)$ depend only on the input p , which in the present case represents the set of injection moulding parameters. A reduced representation of the solution is obtained when $k < n$ (with n denoting the number of precomputed solutions) modes are used in the series expansion (1). The goal of the subsequent machine learning procedure is the prediction of the POD coefficients a_i in dependence of the process parameters p .

Consequently, in contrast to previous studies on surrogate modelling of FEA data (Hürkamp et al., 2020b), within this study the reduced FEA data set by POD is taken as input for machine learning. This leads to three major differences in the data set characteristics. First, the target variable is no longer a physical property (e.g., temperature or contact time), but multiple POD coefficients per target variable (here: first 50 POD coefficients a_i). Due to this, a multi-target regression approach is pursued. Second, the feature set only consists of the simulated (and measurable) process parameters, e.g. flow rate, mould temperature and part insert temperature as the geometrical information is embedded within the spatial modes $\phi_i(x)$. Third, due to the removed geometrical data, only one data point per process parameter variation is needed, which significantly decreases the data volume. In total, the training data set (if a share of 20 % is hold out for testing) is sized 80×3 with 50 target variables, in contrast to $655,680 \times 6$ and a single target variable for FEA. Within pre-processing, the input features are standardized by MinMax scaling. As for FEA in the previous study, a benchmark across six data-driven approaches is performed, which are AdaBoost, Decision Tree, Gradient Boosting (GradBoost), Polynomial Regression,

RandomForest and Extreme Gradient Boosting (XGBoost). The training is implemented by the scikit-learn library (Pedregosa et al., 2011) of python using a 5-fold cross-validation. As shown in Table 1 a sound hyperparameter optimization is performed. The POD and FEA columns list the hyperparameters that result in best scores for R^2 for both approaches. The comparison shows that the hyperparameters differ between the two approaches, which is reasonable, because of the significantly changed properties of the underlying data sets of the same thermoplastic composite structure.

Table 1: Overview of applied hyperparameter optimization of six data-driven methods for the POD and FEA surrogate modelling approaches

Method	Hyperparameters	Range	# Steps	POD	FEA
AdaBoost	learning rate	0.1-1	10	0.9	0.7
	# estimators	10-100	7	90	30
Decision Tree	max depth	10-None	7	200	None
	min samples leaf	1-10	10	3	6
GradBoost	learning rate	0.1-1	10	0.1	0.7
	# estimators	10-500	7	300	500
Polynomial Reg.	degree	1-7	7	3	4
RandomForest	max depth	10-70	4	30	50
	min samples leaf	1-10	5	1	1
	# estimators	10-500	9	300	100
XGBoost	learning rate	0.1-1	10	0.1	0.7
	# estimators	10-500	7	500	500

3.3 Preparing the Surrogate to run on the edge

A challenge when developing software solutions to run on the edge is that the development is usually different from the deployment environment. In this case, the development of the simulation and the surrogate model was on personal computers (workstation) and the result needs to be deployed on a separate edge-device. One technology to transfer software from one environment to the other without compatibility issues is Docker containers. Containers are packaged software environments made up of different layers. The bottom layer is the Host-OS usually a Linux distribution with additional software, libraries and configurations on top. The main benefit of containers is that they can be seamlessly shared between different host systems independent from the local environment. For this study the base image with the Host-OS is `python3.6:slim-buster` from `docker-hub` (public, online container repository). The image contains Linux Debian 10.9 with Python 3.6 already installed. In the next step, the surrogate application comprised of the Python code and trained machine learning models, is added to the container. Additional libraries (e.g., XGBoost) are installed. In order to access the surrogate model inside the container an application programmable interface (API) is employed. The API takes simulation parameters and the machine learning method as an input. With the input the API

executes the surrogate model inside the container and provides the output in a structured format (JSON). Request and response are transferred using the http protocol, which is very widespread and makes the API integrable in nearly any other software.

4 Results

Machine Learning Results

Regarding the prediction of each single POD coefficient, it can be seen that significantly weak predictions can occur by Polynomial Regression. Further, there are no clear performance differences or trends across the 50 POD coefficients for all six models. Therefore, in Figure 3, a comparison of the R^2 scores for the two approaches POD (x-axis) and FEA (y-axis) across the six data-driven methods is depicted. The depiction shows that in contrast to the FEA approach all six models are feasible for the POD surrogate modelling task as DecisionTree performs weakest for POD with $R^2 = 0.893$ and a mean of all models of $R^2 = 0.936$ (mean FEA: 0.646). For both approaches RandomForest ($R^2_{\text{POD}} = 0.969$; $R^2_{\text{FEA}} = 0.925$) outperforms all other models, although GradBoost ($R^2_{\text{POD}} = 0.968$) and XGBoost ($R^2_{\text{POD}} = 0.966$) are comparable, especially for POD-based surrogate modelling.

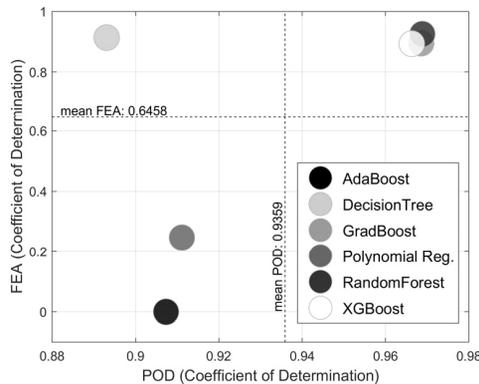


Figure 3: Comparison of R^2 between the approaches FEA and POD surrogate modelling (right) for all six modelling approaches.

Approximation Quality

From the estimated POD coefficients, the solution given in Eq. (1) can be computed. Since it represents an approximation of the full-scale solution, the corresponding error versus the number of modes k used in the reduced representation is displayed in Figure 4.

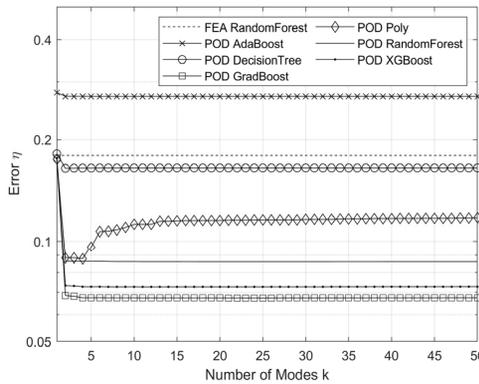


Figure 4: Comparison of Model error with respect to full-scale simulations.

As reference, the error achieved by Random Forest used for FEA is displayed. Note, that the plotted errors are the average of 100 results. It can be seen, that POD performs in general better (except AdaBoost) than the direct FEA approach. DecisionTree shows the lowest coefficient of determination. However, the error is even slightly smaller than for the best FEA approach. In the present study, GradBoost and XGBoost achieve the smallest error. Even though Random Forest performed best in approximating the POD coefficients (cf., Figure 3), these two methods lead to a smaller error compared to the full-scale simulation. In this context, it should be mentioned that with increasing number of modes, the influence of the individual modes decreases. Hence, the first modes are most important for the POD approximation. When the approximation of the first mode is already insufficient, a comparable large model error can be expected. From the graph, we observe that each method converges after five modes (except Polynomial Regression which in fact shows an error increase when using more than five modes), which implies that the number of target variables can be significantly decreased and only the first modes are important to predict.

Testing the container on an Edge Device and the Edge Server

In order to verify the edge suitability of the container with the surrogate model it is executed on the Edge Server (Workstation with 3.9 GHz Xeon W2245, 128 GB RAM) and the Edge Device (1.9 GHz i7-3517U, 4 GB RAM). Both are running on a Windows 10 OS and have a x84-64 CPU architecture. The Server can handle the FEA and the POD approach, but the device is not able to compute the FEA surrogate model because it is running out of RAM. Figure 5 depicts the various calculation times depending on the machine learning method for POD surrogates. Every method was tested with 100 different parameter configurations. The bar shows the mean calculation time and the error bar shows the standard deviation, which is significantly larger for the Edge Device. As expected, the Edge Server vastly outpaces the Edge Device. Most methods take below 1 second to compute in both environments with the exception of GradBoost an XGBoost.

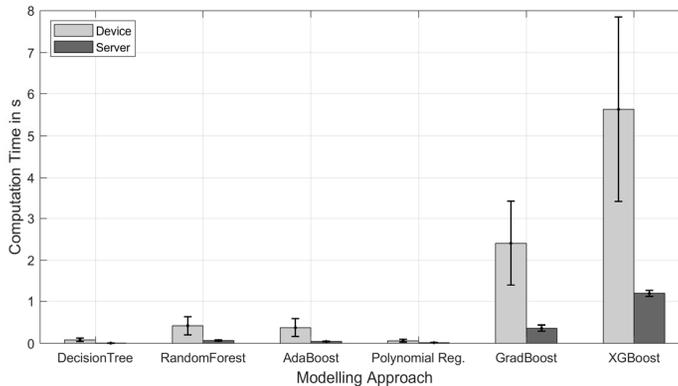


Figure 5: Calculation times for the POD surrogate models on the edge server and on the edge device.

5 Concluding Remarks

The combination of sensor data and real time-feasible surrogate models obtained from engineering data in flexible production systems shows great potential. Since continuous data streams need to be evaluated, low-data demanding and fast algorithms are necessary to define virtual quality gates. Utilizing POD from offline obtained simulation results extracts the main features of the underlying physics and forms the basis for a reduced representation. The advantage for this kind of surrogate modelling lies in the reduced number of data that need to be processed during training and evaluation which makes it suitable for edge computing. Furthermore, it has been shown in the given use case of a thermoplastic composites that the surrogate error even decreases when using a POD approach instead of a pure data driven modelling of the FEA data set. The smallest error in this cases is obtained with GradBoost. However comparing the overall metrics and the calculation time on the edge server, RandomForest is most suitable. In future works, the investigated process chain will be extended and the edge computing concept for flexible manufacturing cells will be implemented and tested within the IML.

References

- Allison, J.; Backman, D.; Christodoulou, L.: Integrated computational materials engineering: A new paradigm for the global materials profession. *JOM* 58 (2006) 11, pp. 25–27.
- Chinesta, F.; Cueto, E.; Abisset-Chavanne, E.; Duval, J.L.; Khaldi, F.E.: Virtual, Digital and Hybrid Twins: A New Paradigm in Data-Based Engineering and Engineered Data. *Archives of Computational Methods in Engineering* (2018).
- Dröder, K.; Reichler, A.-K.; Mahlfeld, G.; Droß, M.; Gerbers, R.: Scalable Process Chain for Flexible Production of Metal-Plastic Lightweight Structures. *Procedia CIRP* 85 (2019), pp. 195–200.

- Filz, M.-A.; Gellrich, S.; Turetskyy, A.; Wessel, J.; Herrmann, C.; Thiede, S.: Virtual Quality Gates in Manufacturing Systems: Framework, Implementation and Potential 2020.
- Han, Z.-H.; Zhang, K.-S.: Surrogate-Based Optimization. In: Roeva, O. (Hrsg.): Real-World Applications of Genetic Algorithms: InTech 2012,
- Hürkamp, A.; Dér, A.; Gellrich, S.; Ossowski, T.; Lorenz, R.; Behrens, B.-A.; Herrmann, C.; Dröder, K.; Thiede, S.: Integrated computational product and production engineering for multi-material lightweight structures. *The International Journal of Advanced Manufacturing Technology* 110 (2020a) 9-10, pp. 2551–2571.
- Hürkamp, A.; Gellrich, S.; Dér, A.; Herrmann, C.; Dröder, K.; Thiede, S.: Machine Learning and Simulation Based Surrogate Modeling for Improved Process Chain Operation. *The International Journal of Advanced Manufacturing Technology* (2021).
- Hürkamp, A.; Gellrich, S.; Ossowski, T.; Beuscher, J.; Thiede, S.; Herrmann, C.; Dröder, K.: Combining Simulation and Machine Learning as Digital Twin for the Manufacturing of Overmolded Thermoplastic Composites. *Journal of Manufacturing and Materials Processing* 4 (2020b) 3, pp. 92.
- Liang, L.; Liu, M.; Martin, C.; Sun, W.: A deep learning approach to estimate stress distribution: a fast and accurate surrogate of finite-element analysis. *Journal of the Royal Society, Interface* 15 (2018) 138.
- Panchal, J.H.; Kalidindi, S.R.; McDowell, D.L.: Key computational modeling issues in Integrated Computational Materials Engineering. *Computer-Aided Design* 45 (2013) 1, pp. 4–25.
- Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; others: Scikit-learn: Machine learning in Python. *the Journal of machine Learning research* 12 (2011), pp. 2825–2830.
- Pfrommer, J.; Zimmerling, C.; Liu, J.; Kärger, L.; Henning, F.; Beyerer, J.: Optimisation of manufacturing process parameters using deep neural networks as surrogate models. *Procedia CIRP* 72 (2018), pp. 426–431.
- Reichler, A.-K.; Gerbers, R.; Falkenberg, P.; Türk, E.; Dietrich, F.; Vietor, T.; Dröder, K.: Incremental Manufacturing: Model-based part design and process planning for Hybrid Manufacturing of multi-material parts. *Procedia CIRP* 79 (2019), pp. 107–112.
- Sancarlos, A.; Cameron, M.; Abel, A.; Cueto, E.; Duval, J.-L.; Chinesta, F.: From ROM of Electrochemistry to AI-Based Battery Digital and Hybrid Twin. *Archives of Computational Methods in Engineering* (2020).
- Swischuk, R.; Mainini, L.; Peherstorfer, B.; Willcox, K.: Projection-based model reduction: Formulations for physics-based machine learning. *Computers & Fluids* 179 (2019), pp. 704–717.
- Thiede, S.; Turetskyy, A.; Kwade, A.; Kara, S.; Herrmann, C.: Data mining in battery production chains towards multi-criterial quality prediction. *CIRP Annals* 68 (2019) 1, pp. 463–466.
- Wang, P.; Gao, R.X.; Fan, Z.: Cloud Computing for Cloud Manufacturing: Benefits and Limitations. *Journal of Manufacturing Science and Engineering* 137 (2015) 4.
- Wu, D.; Liu, S.; Zhang, L.; Terpenney, J.; Gao, R.X.; Kurfess, T.; Guzzo, J.A.: A fog computing-based framework for process monitoring and prognosis in cyber-manufacturing. *Journal of Manufacturing Systems* 43 (2017), pp. 25–34.

- Wuest, T.; Weimer, D.; Irgens, C.; Thoben, K.-D.: Machine learning in manufacturing: advantages, challenges, and applications. *Production & Manufacturing Research* 4 (2016) 1, pp. 23–45.
- Zambal, S.; Eitzinger, C.; Clarke, M.; Klintworth, J.; Mechin, P.: A digital twin for composite parts manufacturing: Effects of defects analysis based on manufacturing data. In: 2018 IEEE 16th International Conference on Industrial Informatics (INDIN), 2018, pp. 803–808.
- Zimmerling, C.; Poppe, C.; Kärger, L.: Estimating Optimum Process Parameters in Textile Draping of Variable Part Geometries - A Reinforcement Learning Approach. *Procedia Manufacturing* 47 (2020), pp. 847–854.





*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Entwicklung einer dynamischen Routenfindung in hybriden Mensch-Roboter-Kommissioniersystemen

Development of dynamic routing in hybrid human-robot order picking systems

Dirk Kauke, Angelika Rett, Johannes Fottner, Technische Universität München,
Lehrstuhl für Fördertechnik Materialfluss Logistik, Garching (Germany),
dirk.kauke@tum.de, angelika.rett@tum.de, j.fottner@tum.de

Abstract: Over the past years the e-commerce sector has been growing continuously while the shortage of personnel has worsened. Therefore, companies increasingly use mobile picking robots in their order picking systems. These robots are capable of picking articles independently, it is however not yet possible to rely solely on them. When robots and humans work alongside in hybrid order picking systems, they inevitably interact with each other, which can create blockages in the picking aisles and reduces the performance of the whole system. This paper investigates a dynamic approach to routing of the robots within the warehouse to actively avoid interactions with other participants. It is shown that independent of the size of a warehouse this strategy has the potential to increase the performance of a hybrid order picking system.

1 Einleitung

Das anhaltende Wachstum des Onlinehandels sowie ein steigender Fachkräftemangel in der Transport- und Logistikindustrie (Kohl; Pfrezschner, 2019) hat dazu geführt, dass bereits seit einigen Jahren die Einsatzmöglichkeiten von Robotern in der Intralogistik untersucht werden. Besonders durch den Einsatz von autonom kommissionierenden Robotern in Person-zu-Ware Kommissioniersystemen, sogenannten Mobile-Picking-Robots (MPRs), erhofft man sich Erleichterung in dieser Situation. Aktuell stellt jedoch die Greiffunktion der Roboter noch eine große Herausforderung dar. Die Roboter können aufgrund der unterschiedlichen Maße der gelagerten Produkte nicht alle Artikel des Sortiments sicher greifen. Somit ist es zumindest bisher im Allgemeinen noch nicht möglich, komplett auf menschliche Arbeitskräfte zu verzichten. Es bietet sich allerdings an, ein hybrides Kommissioniersystem zu nutzen, in welchem die MPRs parallel zu den Menschen selbständig kommissionieren, um so die Vorteile, die der Einsatz von Robotern bietet, möglichst weit auszunutzen.



Durch die Zusammenarbeit kommt es aber unvermeidbar zu Interaktionen zwischen Menschen und Robotern, was zu Verzögerungen führt. Durch erzeugte Staus oder Blockaden in den Lagergassen wird die Leistung des gesamten Systems gesenkt. Daher soll eine dynamische Routenführung die Interaktionen so weit wie möglich vermeiden.

2 Stand der Forschung

Der Einsatz mobiler Roboter innerhalb der Kommissionierung wird bereits seit vielen Jahren in der Wissenschaft untersucht. Bereits 2008 stellten Wurmann und D'Andrea ihr Konzept zum Einsatz von mobilen Robotern in der Kommissionierung vor (Wurman et al., 2008; Andrea und Wurman, 2008). Das Konzept, welches heute als *Robotic Mobile Fullfilment Center (RMFC)* beschrieben wird, sieht eine Flotte an Fahrerlosen Transportfahrzeugen (FTF) vor, die mobile Regale zu Pickplätzen transportieren, an denen der Mensch die geforderten Produkte entnimmt. Die Besonderheit hier ist die räumliche Trennung der Arbeitsbereiche von Menschen und Robotern.

Einen ähnlichen Einsatz findet man bei den *Pick-Support-AGVs*. In diesem Fall assistieren die FTFs dem Menschen bei der Kommissionierung, indem sie beispielsweise Leer- oder Vollgut transportieren und der Mensch lediglich das Greifen der Artikel übernimmt. Im Gegensatz zu den RMFCs teilen sich Mensch und Roboter einen Arbeitsbereich. Erste Untersuchungen finden sich bei Azadeh et al. (2020), Loffler et al. (2019) und Žulj et al. (2020). In keiner der Arbeiten werden jedoch Interaktionen zwischen Menschen und Robotern und daraus resultierende Leistungsverluste berücksichtigt.

Die dritte Gruppe an mobilen Robotern in der Kommissionierung, die *Mobile-Picking-Robots*, übernehmen neben dem Transport ebenfalls eigene Pickvorgänge. Die MPRs können somit autonom vom Menschen im System agieren. Zwar unterliegt insbesondere das Greifen von heterogenen Artikeln umfassender Forschung, jedoch haben erst vereinzelte Arbeiten eine ganzheitliche Ausgestaltung von solch hybriden Kommissioniersystemen untersucht (Kauke und Fottner, 2020; Kauke et al., 2021).

Vernachlässigt wird bis dato vor allem die dynamische Anpassung von Routen, der in der außerbetrieblichen Logistik in den letzten Jahren hingegen viel Aufmerksamkeit gewidmet wurde. Ein Großteil der Veröffentlichungen betrachtet als Ursache für die Dynamik den Eingang von Kundenaufträgen, nachdem Touren gestartet wurden, was dazu führt, dass Routen nachträglich angepasst werden müssen, um diese Kundenaufträge zu erfüllen (Kucharska, 2019; Lorini et al., 2011; Moretti Branchini et al., 2009). Es lassen sich aber auch Arbeiten finden, die Routen basierend auf der benötigten Gesamtfahrzeit optimieren und dabei die aktuelle Verkehrslage analysieren (Fleischmann et al., 2004; Güner et al., 2012). Die vermutlich bekannteste Umsetzung in diesem Bereich ist Google Maps.

Unter dem Stichwort *Conflict-Free AGV Routing* lassen sich auch im Bereich der Intralogistik zahlreiche Veröffentlichungen finden, die sich mit der Routenplanung von Automated Guided Vehicles (AGVs) beschäftigen. Jedoch werden die zukünftigen Aufenthaltsorte aller AGVs lediglich basierend auf der ihnen zugewiesenen Tour bestimmt, wobei ungeplante Verzögerungen nicht berücksichtigt werden (Möhring et al., 2005; Solichudin et al., 2020).

Weitaus weniger Arbeiten untersuchen eine dynamische Anpassung der Route basierend auf der aktuellen Auslastung des Systems (Maza und Castagna, 2005; Nguyen Duc et al., 2020; Secchi et al., 2015).

3 Zielsetzung

Sowohl im Bereich der dynamischen Routenfindung als auch im Bereich der hybriden Kommissioniersysteme ist deutlicher Forschungsbedarf zu erkennen. Dieser Beitrag setzt daher gezielt an der Schnittstelle der beiden Bereiche an und untersucht, ob der Einsatz einer dynamischen Routingstrategie die Gesamtleistung eines hybriden Systems positiv beeinflussen kann. Es stellt sich außerdem die Frage, ob sich Unterschiede hinsichtlich des Leistungszuwachses in Abhängigkeit von der Agentenanzahl feststellen lassen.

4 Simulationsmodell

Um die komplexen Interaktionen zwischen Menschen und Robotern sowie die zahlreichen dynamischen Einflüsse auf die Systemleistung eines hybriden Systems untersuchen zu können, wird im Zuge dieser Arbeit ein Simulationsmodell genutzt. So ist es möglich, die entwickelten Strategien in unterschiedlichen Szenarien und unter zufälliger Variation definierter Parameter zu untersuchen, was analytisch nicht sinnvoll umsetzbar wäre.

4.1 Grundkonzept

Im Zuge früherer Arbeiten (Kauke und Fottner, 2020; Kauke et al., 2021) wurde bereits ein Planungstool in Form eines Simulationsmodells erstellt, welches mit der Simulationssoftware Tecnomatix Plant Simulation umgesetzt ist. Dieses Modell ermöglicht es, unterschiedlichste Parameter eines dargestellten Lagers zu variieren, wie beispielsweise das Layout oder die Anzahl der eingesetzten Agenten. Das Tool soll nun um die benötigten Funktionen erweitert werden, sodass der Einfluss von dynamischer Routenführung auf die Leistung des Systems untersucht werden kann.

4.1.1 Layout

Das Layout (Abb. 1) stellt die gewählte Konfiguration des Lagers sowie die Bewegungen der Agenten grafisch dar. Hellgraue Kacheln symbolisieren befahrbare Knoten, dunkelgraue Kacheln im oberen Teil der Abbildung die Regale. Die einzelnen dunkelgrauen Kacheln am unteren Rand des Layouts repräsentieren Depots, welche Ausgangs- und Zielort einer Kommissioniertour sind. Befahrbare Kacheln können sowohl vertikal als auch horizontal durchquert werden, das diagonale Wechseln auf ein angrenzendes Feld ist jedoch nicht möglich. Auch kann ein Feld immer nur von einem Agenten gleichzeitig befahren werden.

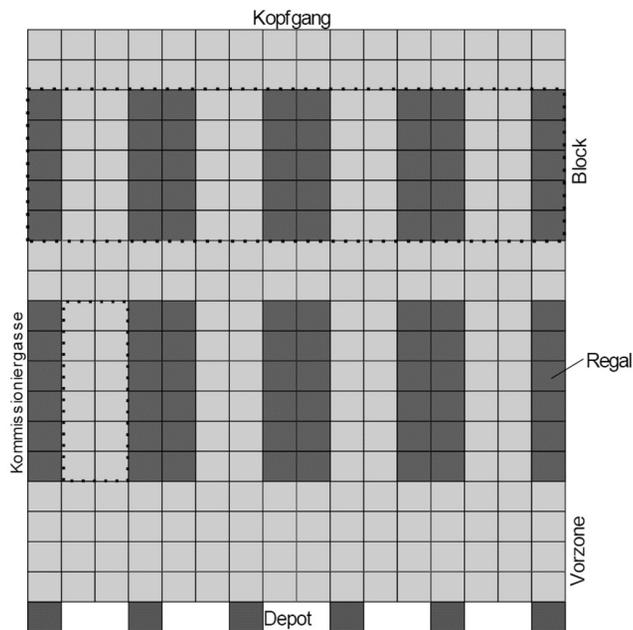


Abbildung 1: Beispielhaftes Lagerlayout

Frei variierbare Parameter sollen es einerseits ermöglichen, ein reales System zu Planungszwecken möglichst genau nachzubilden und andererseits viele verschiedene Layoutvarianten simulativ untersuchen zu können. Zu diesen Parametern zählen:

- Anzahl der Blöcke
- Anzahl der Gassen je Block
- Länge der Gassen
- Anzahl der Depots

4.1.2 Interaktionen

MPRs verfügen üblicherweise über Sicherheitstechnik, die Zusammenstöße mit Hindernissen verhindern soll. Nähert sich der Roboter einem Hindernis, sodass sich dieses im vor ihm liegenden Warnbereich befindet, verringert er seine Geschwindigkeit, bewegt sich aber dennoch weiter. Detektieren die Sensoren jedoch ein Hindernis in der Sicherheitszone des Roboters, so hält dieser sofort an und setzt seinen Weg erst fort, wenn das Hindernis verschwunden ist. Um Zusammenstöße zu verhindern, verhalten sich Menschen intuitiv ähnlich wie die Roboter bei Hindernissen in der Sicherheitszone. Sie beschränken sich dabei allerdings auf die Geschehnisse direkt vor ihnen, weshalb die Sicherheitszone des Menschen im Modell weniger groß ausgeprägt ist. Es wird außerdem davon ausgegangen, dass der Mensch sehr viel flexibler ist als der Roboter und daher besser auf Hindernisse reagieren kann. Aus diesem Grund benötigt der Mensch keinen Warnbereich. In Abbildung 2 sind die Sicherheitszone und der Warnbereich des Roboters sowie die Sicherheitszone des Menschen dargestellt.

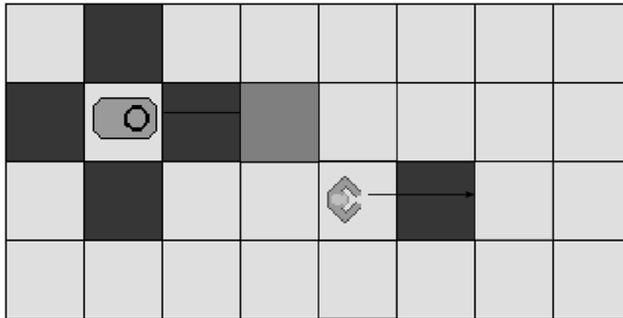


Abbildung 2: Sicherheitszone (dunkelgrau) und Warnbereich (hellgrau) von Roboter und Mensch

Für den Fall, dass zwei Agenten aufeinandertreffen, verfügen diese über Strategien zum Ausweichen. Einer der Agenten bleibt dabei jeweils stehen, während der andere seine Route anpasst. Die Ausweichmanöver werden vorrangig vom Menschen durchgeführt, da dieser flexibler und damit vor allem schneller als der Roboter ist. Jedoch verfügen auch die Roboter über die nötige Steuerung, um Hindernissen ausweichen zu können. Ausweichvorgänge führen allerdings zwangsläufig zu Verzögerungen, die minimiert werden können, wenn durch geschickte Routenplanung die Wahrscheinlichkeit des Aufeinandertreffens von zwei Agenten verringert wird.

4.1.3 Statisches Routing

Bevor Agenten das Depot verlassen, wird ihnen eine optimierte Kommissioniertour zugewiesen. Diese Tour enthält eine feste Anzahl an zu kommissionierenden Aufträgen, die zu einem Batch zusammengefasst sind, sowie die exakte Route zwischen den einzelnen Stopps. Die Reihenfolge der anzufahrenden Stopps und die Route ist bereits hinsichtlich einer minimalen Gesamtdauer der Tour optimiert. Die verwendeten Optimierungsverfahren betrachten jedoch lediglich die absoluten Entfernungen zwischen zwei Punkten und die Geschwindigkeit der Agenten, nicht jedoch die aktuelle Auslastung des Systems.

4.2 Dynamisches Routing

Ziel des dynamischen Routings ist es, während die Agenten im Lager unterwegs sind, flexibel auf aktuelle Hindernisse oder Engstellen in den Gassen reagieren zu können und die Routen der betroffenen Agenten so anzupassen, dass die Anzahl an Interaktionen minimiert wird. Es wird angenommen, dass der Mensch von Natur aus flexibel genug ist, um Hindernisse souverän und schnell zu umgehen, weshalb die vorgestellte Strategie nur bei den Robotern zum Tragen kommt.

Grundlage für die dynamische Route ist die zu Beginn einer Tour statisch ermittelte Route, welche noch nicht die Auslastung des Gesamtsystems berücksichtigt. Immer wenn ein Agent einen der ihm vorgegebenen Stopps erreicht und die geforderten Artikel gepickt hat, soll durch das dynamische Routing die geplante Route zum nächsten Stopp überprüft und gegebenenfalls entsprechend der aktuellen Auslastung optimiert werden. Es sollen zwei Optimierungsvarianten untersucht werden, welche im Folgenden erläutert werden.

4.2.1 *Optimierungsvariante 1: Anpassung der Route*

Einerseits soll die Möglichkeit untersucht werden, die Route zwischen den einzelnen Stopps anzupassen. Hierzu wird an jedem Stopp, nachdem der Agent die erforderlichen Artikel an dieser Position gepickt hat, eine Momentaufnahme der aktuellen Auslastung des Systems betrachtet und mit der vorgegebenen Route zum nächsten Stopp verglichen. Führt das nächste Teilstück über aktuell belegte Felder oder durch die Sicherheitszone anderer Agenten, so wird eine alternative Route zum nächsten Stopp ermittelt. Für die Berechnung wird der A*-Algorithmus verwendet, der die beste Route entsprechend der benötigten Wegzeit ermittelt. Um die momentane Auslastung berücksichtigen zu können, wird der Zeitbedarf für das Befahren eines belegten Feldes oder des Sicherheitsbereiches der anderen Agenten stark erhöht, wodurch diese Felder in der neuen Route meist vermieden werden. Der Roboter folgt der neuen Teilroute zum nächsten Stopp, pickt dort die geforderten Artikel und überprüft dann erneut die aktuelle Auslastung, um gegebenenfalls die nächste Teilroute zu optimieren.

4.2.2 *Optimierungsvariante 2: Anpassung der Reihenfolge der Stopps*

Aufbauend auf der Anpassung der Route soll des Weiteren untersucht werden, ob es mit der veränderten Route lohnenswert ist, die Reihenfolge der noch nicht angefahrenen Stopps zu verändern. Hierfür wird überprüft, ob die veränderte Teilroute bereits an späteren Stopps vorbeiführt. Sollte dies der Fall sein, so werden die betroffenen Stopps vorgezogen und vor dem eigentlich nächsten Stopp abgearbeitet. Dies soll verhindern, dass der Roboter wieder zurückfahren muss und Wegstrecke doppelt zurücklegt, da besonders Wendevorgänge viel Zeit in Anspruch nehmen. Nachdem die Reihenfolge der Stopps angepasst wurde, folgt der Agent der neu ermittelten Teilroute zu dem nun vorgezogenen Stopp, pickt dort die entsprechenden Positionen und prüft dann wieder die Auslastung auf dem Routenabschnitt zum nächsten geplanten Stopp.

5 Ergebnisbetrachtung

Als primärer Indikator für die Leistungsfähigkeit des Gesamtsystems wird der Durchsatz an Aufträgen innerhalb von zwei Schichten mit je 8 Stunden betrachtet. Des Weiteren gibt die durchschnittliche Wartezeit während einer Tour Auskunft darüber, ob die Funktionalität der angewandten Steuerung gewährleistet ist.

Es werden zwei Lagerkonfigurationen betrachtet, die ein kleines und ein großes Lager darstellen sollen. Die Konfigurationsparameter exakten Konfigurationen sind in Tabelle 1 zu finden. Sowohl die Gesamtanzahl an Agenten als auch die Verteilung von Menschen und Robotern werden systematisch variiert, um ihren Einfluss auf die Wirksamkeit des dynamischen Routings zu untersuchen.

Tabelle 1: *Untersuchte Layoutkonfigurationen*

	Anzahl Blöcke	Gassen je Block	Länge der Gassen	Anzahl Depots	Maximale Agentenanzahl eines Typs
Layout 1 – klein	1	15	12 m	8	12
Layout 2 – groß	1	11	48 m	8	18

5.1 Allgemeine Erkenntnisse

Es lassen sich basierend auf den Ergebnissen einige allgemeine Aussagen über hybride Kommissioniersysteme treffen, welche teilweise bereits in früheren Arbeiten getroffen wurden (Kauke und Fottner, 2020; Kauke et al., 2021), die der Vollständigkeit halber aber dennoch angeführt werden.

- Alle beschriebenen Effekte sind in beiden untersuchten Layouts zu beobachten.
- Mit steigender Agentenanzahl steigt tendenziell auch der Output. Es wird jedoch ein Leistungsmaximum erreicht, ab welchem die Performance aufgrund der steigenden Interaktionen wieder sinkt. In dem kleinen Layout liegt dieses Maximum bei 15 Agenten, das große Layout weist die maximale Leistungsfähigkeit mit 21 Agenten auf.
- Bei gleichbleibender Agentenzahl und steigendem Anteil der Roboter steigt die durchschnittliche Wartezeit pro Tour der Roboter. Es kommt zu mehr Roboter-Roboter Interaktionen, welche aufgrund der geringeren Geschwindigkeit der MPRs mehr Zeit in Anspruch nehmen als Ausweichvorgänge, an denen ein Mensch beteiligt ist. Der Anstieg der Wartezeit führt zu einem sinkenden Output.

5.2 Anpassung der Route

Um zunächst die grundsätzliche Frage der Funktionalität der dynamischen Routenanpassung zu klären, wird die Wartezeit je Tour betrachtet. Es ist klar zu erkennen, dass die Wartezeit des Roboters stark abnimmt, was bedeutet, dass die Anzahl der Interaktionen mit Beteiligung von Robotern deutlich verringert wurde. Auch die Wartezeit der Menschen sinkt leicht, was ebenfalls mit der Vermeidung von Interaktionen zu begründen ist. Der beschriebene Effekt ist in beiden Layoutvarianten zu erkennen, exemplarisch sind in Abbildung 3 die Ergebnisse aus dem kleinen Layout dargestellt.

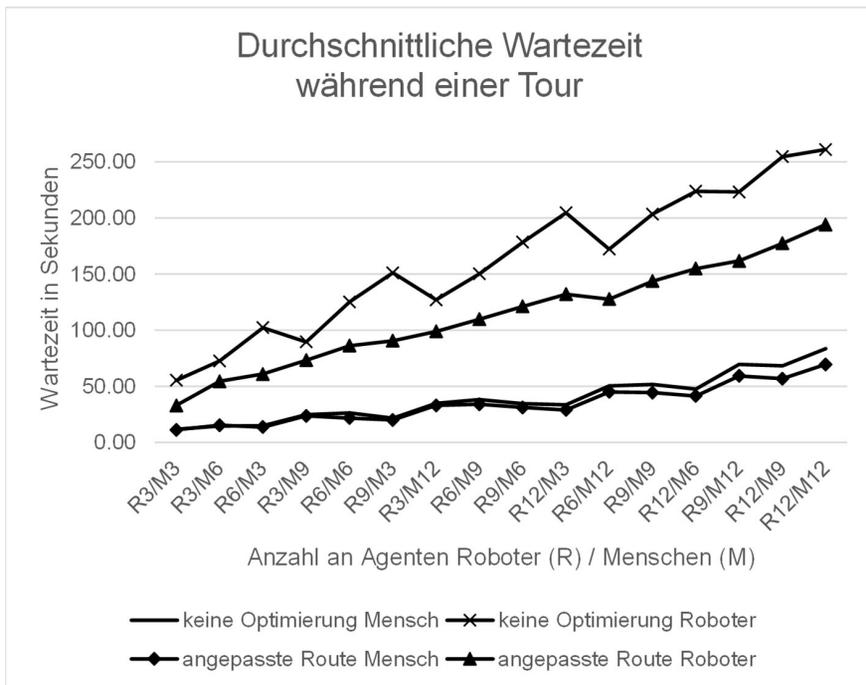


Abbildung 3: Entwicklung der Wartezeit von Mensch und Roboter in Layout 1

Um Zeitverluste durch Hindernisse zu vermeiden, wird beim dynamischen Routing ein längerer Weg in Kauf genommen. Trotz der erhöhten Fahrzeit sollte der Output dennoch nicht sinken, damit eine dynamische Steuerung vorteilhaft ist. Abbildung 4 zeigt den Einfluss, den die Routenanpassung auf den Durchsatz im untersuchten Layout 2 hat. Es wird deutlich, dass der Output durch die angewandte Strategie grundsätzlich gesteigert werden kann. Befinden sich jedoch nur sehr wenige Agenten im Lager, so wirkt sich die längere Route negativ auf die Gesamtleistung aus. Dies ist der Tatsache geschuldet, dass die Neuplanung der Route auch bei nur einem einzelnen Hindernis auf dem nächsten Wegabschnitt durchgeführt wird, was jedoch nicht immer die längere Wegstrecke rechtfertigt.

Mit wachsender Agentenanzahl steigt das Potential des dynamischen Routings und es ist ein positiver Effekt zu erkennen. Jedoch wird auch hier, wie beim Gesamtdurchsatz, ein Maximum erreicht, ab welchem die Verbesserung wieder sinkt. Befinden sich zu viele Agenten im Lager, so kommt es sowohl auf der ursprünglichen als auch auf der neu gewählten Route vermehrt zu Interaktionen, wodurch der positive Effekt des Umwegs weniger stark ausfällt.

Bei genauer Betrachtung des Diagramms 4 ist zu erkennen, dass der Verlauf bei konstanter Agentenanzahl jeweils ein lokales Maximum aufweist. Steigt der Anteil der Roboter, so wird auch der Effekt des dynamischen Routings zunächst größer, da die Strategie grundsätzlich öfter angewendet wird und so die Vorteile besser ausgenutzt werden können. Steigt der Roboteranteil jedoch weiter, so verringert sich

der Effekt wieder, da die Anzahl an zeitintensiven Roboter-Roboter Interaktionen zunimmt und so die Vorteile überschattet werden.

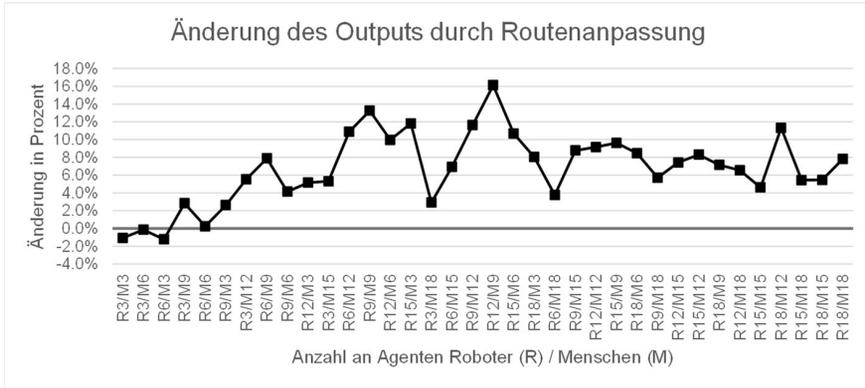


Abbildung 4: Prozentuale Veränderung des Outputs durch die Routenanpassung in Layout 2

5.3 Anpassung der Reihenfolge der Stops

Nachdem bereits gezeigt werden konnte, dass eine dynamische Anpassung der Route positive Effekte hervorruft, soll der Output durch eine Anpassung der Reihenfolge der Stops entsprechend der neuen Route noch weiter gesteigert werden. Die Ergebnisse der Untersuchung sind in Abbildung 5 dargestellt.

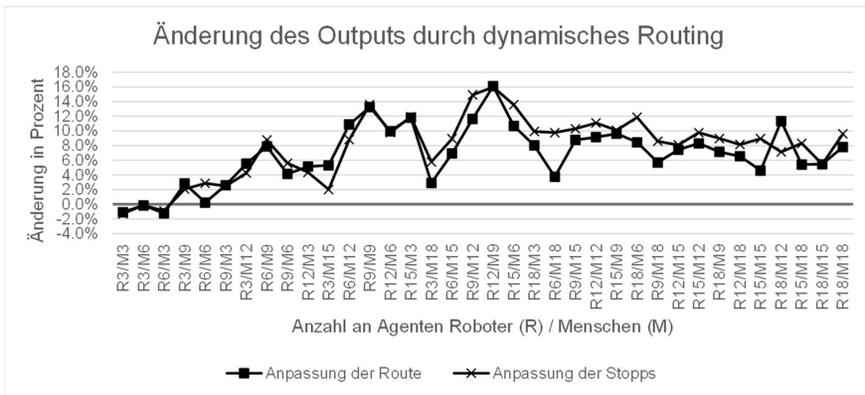


Abbildung 5: Prozentuale Veränderung des Outputs durch die beiden untersuchten Optionen des dynamischen Routings in Layout 2

Es wird schnell ersichtlich, dass der bereits erzielte positive Effekt durch die Anwendung beider vorgestellter Strategien im Allgemeinen noch stärker ausgeprägt ist. Es werden unnötige Wendevorgänge vermieden und Teilstrecken müssen nicht doppelt zurückgelegt werden. So kann die Fahrtzeit der Roboter noch weiter

verringert und der Output zusätzlich gesteigert werden. In seltenen Fällen wird der Output jedoch durch die zweite Optimierungsstufe im Vergleich zur ersten Stufe wieder verschlechtert, was darauf zurückzuführen ist, dass die Route zwar lokal verbessert, aber global verschlechtert wurde. Um dies zu vermeiden, sollte der Rest der Tour nach einer Anpassung der Reihenfolge der Stopps nochmal wie zu Beginn des statischen Routings mittels eines Optimierungsverfahrens global optimiert werden. Aus Gründen der Laufzeit und Rechenzeit wurde im Zuge dieser Untersuchungen allerdings darauf verzichtet.

6 Zusammenfassung und weiterer Forschungsbedarf

Wie der Stand der Forschung bestätigt, besteht im Bereich der hybriden Kommissioniersysteme besonders bei der dynamischen Routenführung großer Forschungsbedarf. Ziel dieser Untersuchung war es daher, den Einfluss von dynamischem Routing auf die Leistungsfähigkeit eines solchen Kommissioniersystems zu untersuchen. Da die Leistung in hybriden Systemen stark davon abhängt, wie harmonisch die Agenten zusammenarbeiten, soll die Zahl an zeitintensiven Interaktionen zwischen den Agenten verringert werden. Dazu wird die Route der MPRs hinsichtlich der aktuellen Auslastung dynamisch angepasst, um das Zusammentreffen mit anderen Agenten aktiv zu vermeiden und auf Hindernisse flexibel reagieren zu können.

Es wurde deutlich, dass eine dynamische Routenführung unabhängig von der Größe des Systems großes Potential bietet. Am Punkt des Leistungsmaximums des Systems ist der Effekt besonders stark ausgeprägt und der Output kann um bis zu 16% gesteigert werden. Ist die Anzahl an Agenten für die jeweilige Lagergröße jedoch sehr gering, so kommt es sogar zu leicht negativen Effekten. An dieser Stelle ist weiterer Forschungsbedarf zu erkennen. Wenn genaue Informationen über den Zeitbedarf der verschiedenen Arten von Interaktionen vorliegen, kann abgewogen werden, ob das aktuelle Hindernis den längeren Umweg tatsächlich rechtfertigt, oder ob die voraussichtliche Interaktion in Kauf genommen werden sollte.

Besonders im Hinblick auf die Anpassung der Reihenfolge der Stopps ist die Batchgröße ein Parameter, der in zukünftigen Untersuchungen betrachtet werden könnte. Je mehr Aufträge ein Agent innerhalb einer Tour zu erledigen hat, desto größer sollte auch die Wahrscheinlichkeit sein, dass einer der nächsten Stopps auf der neu gewählten Route liegt und das Potential der zweiten Optimierungsstufe ausgeschöpft werden kann. Erste Experimente im Zuge dieser Untersuchung haben hierzu noch keine eindeutigen Ergebnisse liefern können.

Die mögliche Leistungssteigerung des Gesamtsystems ist jedoch vor dem Hintergrund der Verlangsamung der Systemlandschaft auch kritisch zu betrachten, da die umfangreiche Steuerung einen hohen Rechenaufwand erfordert. Um reale Systeme in Echtzeit steuern zu können, muss ein solches System zunächst hinsichtlich der Laufzeit optimiert werden, bevor es sinnvoll eingesetzt werden kann.

Literatur

- Andrea, R.D.; Wurman, P.: Future Challenges of Coordinating Hundreds of Autonomous Vehicles in Distribution Facilities (2008), S. 80–83.
- Azadeh, K.; Roy, D.; Koster, M.R. de: Dynamic Human-Robot Collaborative Picking Strategies. *SSRN Electronic Journal* (2020).
- Fleischmann, B.; Gnutzmann, S.; Sandvoß, E.: Dynamic Vehicle Routing Based on Online Traffic Information. *Transportation Science* 38 (2004) 4, S. 420–433.
- Güner, A.R.; Murat, A.; Chinnam, R.B.: Dynamic routing under recurrent and non-recurrent congestion using real-time ITS information. *Computers & Operations Research* 39 (2012) 2, S. 358–373.
- Kauke, D.; Fottner, J.: Planungstool zur Entwicklung geeigneter Betriebsstrategien von mobilen Kommissionierrobotern. In: Kreuzfeldt, J. (Hrsg.): Tagungsband zum 16. Fachkolloquium der Wissenschaftlichen Gesellschaft für Technische Logistik e.V. (WGTL). Hamburg: Technische Universität Hamburg, Institut für Technische Logistik 2020,
- Kauke, D.; Sailer, F.; Fottner, J.: Mobile Picking Robots. In: Knapcikova, L.; Peraković, D.; Behúnová, A.; Perisa, M. (Hrsg.): 5th EAI International Conference on Management of Manufacturing Systems: A first study of the effects of human-robot-interactions in conventional order picking systems: Springer International Publishing 2021,
- Kohl, A.-K.; Pffretzschner, F., 2019: Logistikmonitor 2018: Der Wirtschaftszweig in Zahlen, BVL, Statista. Online verfügbar unter https://www.bvl.de/files/1951/1988/2128/Logistikmonitor_2018_-_Der_Wirtschaftszweig_in_Zahlen.pdf, zuletzt geprüft am 07.05.2021.
- Kucharska, E.: Dynamic Vehicle Routing Problem—Predictive and Unexpected Customer Availability. *Symmetry* 11 (2019) 4, S. 546.
- Löffler, M.; Boysen, N.; Schneider, M.: Picker routing in AGV-assisted order picking systems (2019), S. 44.
- Lorini, S.; Potvin, J.-Y.; Zufferey, N.: Online vehicle routing and scheduling with dynamic travel times. *Computers & Operations Research* 38 (2011) 7, S. 1086–1090.
- Maza, S.; Castagna, P.: Sequence Based Hierarchical Conflict-Free Routing Strategy of Bi-Directional Automated Guided Vehicles. *IFAC Proceedings Volumes* 38 (2005) 1, S. 168–173.
- Möhring, R.H.; Köhler, E.; Gawrilow, E.; Stenzel, B.: Conflict-free Real-time AGV Routing. In: Fleuren, H.; Hertog, D. den; Kort, P. (Hrsg.): *Operations Research Proceedings 2004*. Berlin, Heidelberg: Springer Berlin Heidelberg 2005, S. 18–24.
- Moretti Branchini, R.; Amaral Armentano, V.; Løkketangen, A.: Adaptive granular local search heuristic for a dynamic vehicle routing problem. *Computers & Operations Research* 36 (2009) 11, S. 2955–2968.
- Nguyen Duc, D.; Tran Huu, T.; Nananukul, N.: A Dynamic Route-Planning System Based on Industry 4.0 Technology. *Algorithms* 13 (2020) 12, S. 308.

- Secchi, C.; Olmi, R.; Rocchi, F.; Fantuzzi, C.: A dynamic routing strategy for the traffic control of AGVs in automatic warehouses. In: Institute of Electrical and Electronics Engineers (IEEE) (Hrsg.): 2015 IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, USA, 26.05.2015 - 30.05.2015, 2015, S. 3292–3297.
- Solichudin; Triwiyatno, A.; Riyadi, M.A.: Conflict-free dynamic route multi-AGV using dijkstra floyd-warshall hybrid algorithm with time windows. International Journal of Electrical and Computer Engineering (IJECE) 10 (2020) 4, S. 3596.
- Wurman, P.; D'Andrea, R.; Mountz, M.: Coordinating Hundreds of Cooperative, Autonomous Vehicles in Warehouses Volume 29 Number 1 (2008), S. 9–19.
- Žulj, I.; Goeke, D.; Scheiner, M.; Salewski, H.: Order batching and batch sequencing in an AGV-assisted picker-to-parts system (2020), S. 1–32.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Improving Blood Supply Chain Crisis Management by Simulation-based Optimization

Verbesserung des Krisenmanagements in Blutlieferketten durch den Einsatz von Simulationsbasierter Optimierung

Dennis Horstkemper, WWU Münster, Münster (Germany), horstkemper@ercis.org

Melanie Reuter-Oppermann, TU Darmstadt, Darmstadt (Germany),
oppermann@is.tu-darmstadt.de

Michael Middelhoff, Adam Widera, Bernd Hellingrath, WWU Münster, Münster
(Germany), middelhoff@ercis.org, widera@ercis.org, hellingrath@ercis.org

Abstract: Blood supply chains are a specialized type of supply chain dealing with several legally independent actors that cooperate to facilitate the demand of a resource with limited availability and shelf life. In this setting, adequate decision support tools are needed to ensure good design choices are made, as any interruption in the blood supply chain endangers human lives. To handle the complexity of such a supply chain, we propose a simulation-based optimization framework that enables the depiction of strategic and operational decision-making throughout the overall BSC. We illustrate the toolkits capabilities by applying it on a use case with the aim of improving the utilization of mobile blood donation facilities during a crisis.

1 Introduction

Blood products are a very critical resource in healthcare systems, affecting several procedures, such as scheduled or unplanned surgeries. Providing timely, sufficient supply of blood products to match demand is a challenging task, reliant on volunteer blood donations. Often, it is already challenging to fulfil everyday demands. During a crisis, existing structures and resources might not be available anymore, while high demand and supply fluctuations require short-term contingency plans. The COVID-19 pandemic has not only led to a shortage of blood products worldwide, but the loss of crucial staff due to sickness or public health restrictions significantly affected the blood supply chains (BSC) (Stanworth et al., 2020). Although decision support systems or crisis management plans can support BSC management, several characteristics of BSCs might explain the rather low uptake of current practices and research from the areas of supply chain management and logistics. Firstly, the multi-stakeholder setting of BSCs hinders the application of integrated supply chain design and planning approaches. Secondly, the degree of uncertainty in demand and supply

is significantly higher compared to other supply chains. Thirdly, it is not only the diverse shelf life of blood products making the application domain highly dynamic in terms of supply and demand: the need and stock of blood products is highly interrelated with disruptions in other areas, such as power outages, mass-casualty incidents, or large-scale disasters. Finally, a high degree of complexity and interdependency can be identified when considering the different variations of blood demands vs. supply as well as geographical and cultural contexts.

To handle the described challenges, we suggest the development and application of combined simulation-optimization approaches. The main rationale lies in the ability to vary and analyze assumptions that must be made for an optimization model, as for example input values, relevant scenarios, and functional areas, within the approach. Furthermore, stochastic influences can be tested within the simulation while keeping the optimization model deterministic, reducing the mathematical complexity. Finally, this approach combines rigorous methods with a continuous involvement of domain experts. Simulation-based optimization approaches have been applied to BSCs and production planning in disaster management before, yet existing approaches either assume a centralized decision problem, only focus on smaller excerpts of the BSC, or neglect the implications of disasters on the demand and supply of blood products (Osorio et al., 2015; Ramezani and Behboodi, 2017; Ghasemi et al., 2020). Based on related works in the respective fields, a systematic simulation-optimization-human loop is designed, enabling decision makers to identify both weak points and improvement potentials. An exemplary application of the toolkit for design decisions in the South-African BSCs is described afterwards. The toolkit uses deterministic optimization to configure the network structure for the stochastic simulation, which in turn analyses constraints and changes to the target function in an iterative process while considering further operational decision-making activities. The aim is to find an improved supply chain design that performs well under stochastic influences. Thus, we propose a toolset enabling the depiction and evaluation of the dynamics between different actors and their decisions throughout a complete BSC. The remaining paper is structured as follows: The relevant foundations presented within Section 2. In Section 3, we describe the proposed simulation environment as well as the tools to consider decision support within the simulation and present an exemplary use case. The paper closes with a discussion and outlook in Section 4.

2 Foundations

2.1 BSC management

Blood products are not or only partly substitutable resources and people in critical conditions may depend on their availability (Heidari-Fathian and Pasandideh, 2017). In order to increase availability, an efficient BSC is required, comprising the minimization of costs and the avoidance of wastage while meeting the demand. The BSC process can be divided into four stages or echelons, being: (1) collection, (2) production (including testing and processing, i.e. separation into different blood products), (3) storage and inventory, as well as (4) distribution, each containing main decisions on a strategic, tactical and operational level (Osorio et al., 2015). Some of those decisions are interdependent, so an isolated approach could lead to a myopic perspective which makes a real-life application impractical (Osorio et al., 2015).

This work focuses the South African blood transfusion system. There are two blood establishments operating in South Africa: the Western Cape Blood Service (WCBS), responsible for the Western Cape province, and South Africa National Blood Service (SANBS) providing blood transfusion services in the remaining parts of South Africa (Wise et al., 2020). As nonprofit organizations, WCBS and SANBS must operate cost-effectively. Both blood establishments collect blood from 100% voluntary, non-numerated donors and operate facilities on all relevant echelons of the BSC. SANBS operates 85 blood banks serving more than 600 hospitals. WCBS runs 7 blood banks providing blood products to approximately 190 hospitals.

2.2 Simulation-based optimization for blood logistics and mobile blood donation

While simulation-based systems have commonly been applied to evaluate and design BSCs, they usually only focus on single echelons and do not enable an integrated view of the whole BSC (e.g., Alfonso et al. (2012) and Lowalekar and Ravichandran (2010) focus on the collection stage). Only a few works that consider the whole BSC within their simulation models exist (e.g., Baesler et al. (2014)). Combining simulation and optimization approaches is an even rarer methodological approach. For example, Lang (2009) combine heuristics with a simulation to depict the impact of transshipments and substitutions in the BSC, while Osorio et al. (2017) describe a combination of linear programming and discrete-event simulation to improve production planning in BSCs. However, both approaches focus on specific problems and are not suited to evaluate a broad selection of practical tasks in BSC operations.

Several publications have addressed the concept of locating mobile facilities or routing collection vehicles in order to increase the number of blood donations. Sachdev et al. (2016) for example describe the concept of blood donation mobiles in India and analyze their use between 2012 and 2014. The authors list several advantages of using mobiles including the flexibility and the existing air-conditioning that is especially helpful during hot periods. Şahinyazan et al. (2015) study a mobile blood collection system in Turkey. Blood mobiles are routed through the area and stay at least one day at each chosen location, up to three days. Shuttles take the blood donations to the processing facilities. Gunpinar and Centeno (2016) present an IP model to simultaneously identify the daily number of blood mobiles to operate while minimizing the distance travelled. The authors allow the blood mobiles to relocate during the day and assume a service time of three to seven hours at each location. Two routing models for blood mobiles are analyzed by Rabbani et al. (2017), Heidari and Pasandideh (2018) as well as Karadağ et al. (2021) integrate mobile collection units into an Mixed-Integer Program (MIP) that optimizes the whole BSC. A two-stage stochastic selective-covering-inventory-routing model was proposed by Bashiri and Ghasemi (2018) with stochastic blood demand as well as blood donations.

3 Simulation

3.1 Simulation framework

The South African BSC structure is modelled using the HumLog Suite, which was developed in previous works by the authors (Widera et al., 2017a; Widera et al., 2017b). HumLog Suite offers an environment for agent-based simulations of

humanitarian logistics operations. It is based on the commercial simulation software AnyLogic and thereby extendable for specific use cases. In addition to AnyLogic, HumLog Suite offers a predefined set of common humanitarian actors and their basic decision logic, like beneficiaries representing needs, transport units for goods and people or medical and storage sites. These actors provide basic parameterization and state charts that can be specialized for the requirements of the modelled context. Furthermore, it offers a web-based dashboard to configure simulation scenarios and to view performance metrics derived through the simulations in a practitioner readable manner. In the past, HumLog Suite was applied in other contexts like flooding events and urban evacuation scenarios (Detzer et al., 2016).

In our case, the South African BSC is represented by a set of actors within the agent-based simulation. Stationery facilities are the collection points, at which blood is donated by the population, the testing centers, at which the infectious disease testing is performed, the blood processing center, which transform the collected blood into the various blood products, the blood banks, which store the blood or blood products for a period of time, and the hospitals, which administer the blood to patients and hence create the demand in the system. Each of the facilities is defined by its location in latitude and longitude, its processing capacities and resource requirements. The blood banks provide an additional storage capacity and hospitals create the demand for blood products. Individual facilities can also be situated at the same location, which eliminates transports between them and affects their cost structure. This is for example the case, if a hospital runs an own blood bank. The logistics operations between these facilities are taken care of by a logistics service provider, who offers services via drones and trucks. These transport units must fulfil specific requirements for the transportation of blood products, like for example a guaranteed steady cooling. They are defined by their overall availability at a certain facility, their transport capacities, costs factors and further parameters on their mode of transportation, like for example speed. Finally, the potential supply of blood to be donated is modelled in the population. The supply chain can be distinguished between a supply-driven “push” and a demand-driven “pull” segment. Starting from collection points via test and processing centers to the storage in the blank banks, the supply chain is supply-driven and provides as much blood and blood products as possible according to the configured donation and processing plan. The amount of blood donations of the various blood types at each collection point follows a normal distribution with a configured mean, standard deviation and size for each of the local donor populations. The hospitals have a demand-driven order policy with the blood banks. The demand for blood products is derived from a safety stock at the hospitals and the local population in size and blood type division.

To properly account for the influence of transports in humanitarian operations under potential crisis events, HumLog Suite utilizes a geographical information systems (GIS) map based on the real-world road, water and air infrastructure. This functionality is provided by the AnyLogic software and was customized to be manipulated according to disaster events, to block parts of the network as consequence of a flooding, for example. According to the mode of transportation, each transport unit needs to route within the GIS space, which results in realistic transport conditions.

A specific simulation scenario is defined through a set of input parameters, which affect the blood product demand, the blood donations, the facility placement, and their allocation to each other. The location and capacities of hospitals is seen as a long-

term. Yet, disaster events might affect the availability of a hospital due to for example power outages. Still, within a simulation study, is it typically not changing. The demand for blood products can be varied to compare best to worst case scenarios and can be subject to stochastic influences. Also, the blood testing facilities in South Africa are fixed to two major locations as described before. The amount, location, capacity, safety stocks and allocation of blood banks to hospitals can be varied to test different network configurations toward the objectives, e.g., the minimization of transport costs. The blood processing is defined alike. A set of stationary and mobile donation sites is provided to collect as much blood as needed to fulfil the demand. Given these input parameters, the simulation will assess the BSC performance under potential additional constraints like affected infrastructures or demand fluctuations. The simulation user interface is displayed in Figure 1.

HumLog Suite tracks each individual transport and the amounts of donated, tested, produced, stored and administered blood at the locations as well as the resource utilization over time. This data is used to conclude on key performance indicators for the BSC, which are defined in line with the objectives. Among others these are the utilization of blood and blood products, the amount of discarded blood products due to exceeded shelf life, the transport and overall supply chain costs and the utilization rates of the facilities and transport units. The analysis outcomes are aggregated and displayed via the web-interface, which can point practitioners towards current bottlenecks and improvement potentials to be tested in further simulation scenarios.

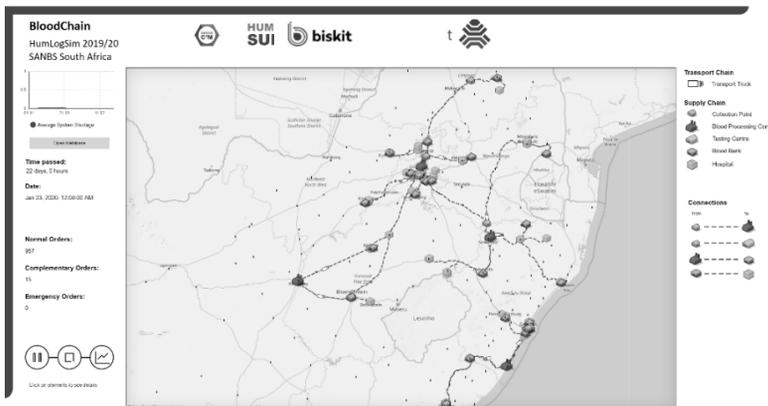


Figure 1: Simulation Run

3.2 Simulation and optimization

Following the VDI3633 Part 12 (Simulation and optimisation), two different ways to connect the simulation of the BSC with optimization-based decision support were implemented (c.f. Rabe and Clausen, 2015). To perform an optimization of strategic decisions, the simulation sequentially follows after an optimization of said decisions. Thus, strategic decisions, e.g. the location of facilities within the BSC, are first determined by a deterministic optimization, in this case using MIP models and the commercial CPLEX optimization software. They are then used as an input for the start of the simulation, which in turn can be used to evaluate the results of the optimization

considering stochastic influences such as demand variations and disruptions caused by external influences such as crises.

In contrast, operational decisions are supported through an optimization integrated into the simulation. At predefined points in time, the simulation model is paused and calls the execution of an optimization as a hierarchically subordinate element. This optimization considers the current state of the simulation, e.g., to perform operational planning tasks such as route planning or the allocation planning of the blood products to the different facilities within the BSC. These optimizations can either be performed using MIP models and the external CPLEX solver again, or by performing manually implemented heuristics for the respective problems. Depending on the analyzed timeframes and decision problems, tactical planning problems can be implemented as either of these optimization types. Figure 2 depicts a sequence diagram visualizing the two different optimization approaches.

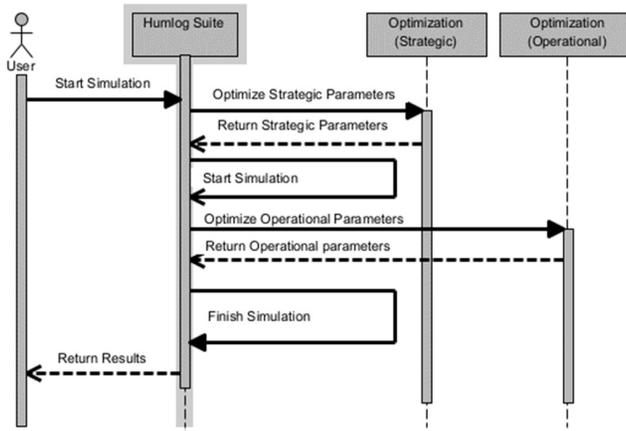


Figure 2: Simulation and Optimization

3.3 Application case: demand-dependent assignment of mobile blood donation facilities

In addition to stationary blood donation sites, mobile blood donation facilities (MBDF) are used to make it easier for potential blood donors to reach the donation sites and thus to increase their willingness to donate. In the South African BSC, MBDFs operate in a single allocation Hub-and-Spoke system. Therefore, operating locations for MBDFs are always assigned to only one hub and each MBDF returns to the hub after operations. Hubs are usually located at blood banks or blood processing centers, where the collected blood can be processed and/or forwarded. Due to their decentral nature, usually only whole blood components are collected by MBDF. The MBDFs visit their operating locations mostly within a fixed pre-planned sequence.

Considering the corona epidemic, two contrasting developments influenced the south African BSCs. On the one hand, the government's action to forbid the sales of tobacco and alcohol severely reduced the need for a number of blood products in the hospitals (e.g., through the reduction of violent behavior under the influence of alcohol). On the other hand, the limited availability of public transport, sicknesses and quarantine measures led to an understaffing of several facilities in the BSC, limiting the blood

supply even further. Thus, the existing sequence for operating the MBDFs is not aligned to the specific demand for blood products anymore. Furthermore, as not all MBDFs can be staffed, the efficiency of their utilization gains even more importance. Hence, one suggestion by domain experts of WCBS was to allocate the MBDFs considering the current stock and demand situation, taking the distribution of possible donors and their blood donation characteristics into account as well. Thus, a MBDF can be placed at a location that offers a concentration of blood donors that can provide a specific blood product (extracted from anonymized data from past donations), which is currently out of stock. The simulation-based optimization framework shall hereby be used to evaluate, this suggestions potential to reduce shortages within the BSC.

Keeping current staff availability in mind, the optimization preceding the simulation is used to determine the maximum number of available MBDFs, as well determining the hub locations. Therefore, a capacitated facility allocation problem is used (Melkote and Daskin, 2001). Depending on the operating radius of a MBDF, the possible operating locations near the hub location are enumerated and exclusively allocated to the MBDF. The maximum amount of MBDFs, their hub locations and their operating locations are then considered as an input for the simulation study itself.

During the simulation, aside from performing the usual transportation, transformation and storage processes within the BSC, it needs to be decided which operating locations are visited by the MBDFs on a daily basis. Therefore, the subordinate operational hierarchic optimization is called at the beginning of each day within the simulation. As an input for the optimization, the current demands and storage amounts of the different blood products weighted across all facilities in the BSC are used to determine which blood products are needed the most. Furthermore, the MBDFs, their possible operating locations and the expected blood donations per location are considered. The optimization model is then formulated as an MIP. Within the target function the demand fulfillment rate is being maximized. Calculating the expected donations using a specific combination of placements of the MBDFs and matching it to the existing demands, the fulfillment rates can be determined. To retain cost optimality, the operating costs of the MBDF are then subtracted from the demand fulfillment rate using respective weights. This modeling approach was chosen in contrast to the modeling using penalty costs, as the penalty cost factors itself are difficult to determine, i.e., cost factors for missing blood products that put patients at risk. The chosen locations for the operation of the MBDF are then returned to the simulation and used for the calculation of the blood donations and succeeding processes throughout the day. This is repeated until the simulations time horizon concludes.

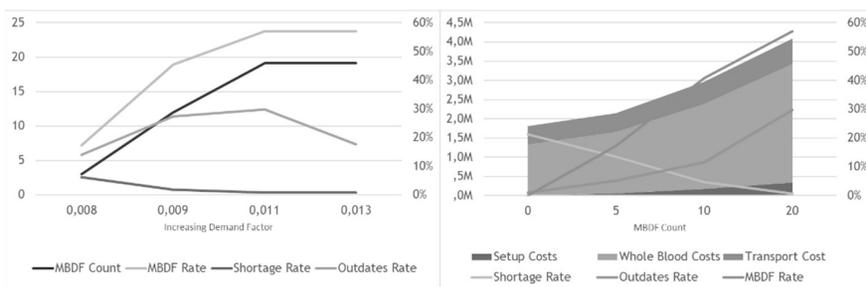


Figure 3: Simulation Results

First simulation studies were performed on the South African BSC based on the recorded data and statistics on blood type characteristics in the South African population. The objective of these first assessments was to validate the behavior of the combined simulation and optimization models and to identify first insights on the current performance of the BSC on highly aggregated values. The left-hand side of Figure 3 plots the BSC performance over an increasing demand factor across several simulation runs. With increasing demand, one can assume that in total more MBDFs are used, which was also observed. It has to be pointed out that the total of MBDFs at a given point in time was limited to 20 units, which peaked at 18.8 units used in average. The MBDF Rate depicts the rate of donations in relation to the total amount, which has a decreasing gradient, meaning that the next MBDF will add less donations due to the higher costs. While the increasing demand naturally decreases the shortages, the outages increased until no additional blood donations are produced and from then on, they decrease as well. This is reasoned by the overproduction of blood products of common blood types, which cannot be avoided, even when trying to optimize for donations matching blood type specific demands. With further increasing demand, shortages are expected to increase again, which was not recorded to this point. The right-hand side of Figure 3 plots additional performance indicators over an increasing MBDF count considering a fixed demand rate. Setup costs grow linearly as expected. Likewise, the linear transport costs are explained by the Hub-and-Spoke system. As MBDFs only collect whole blood, the whole blood costs of the network have an increasing gradient when more units are deployed. The MBDF Rate has a decreasing gradient again, yet not as significant, suggesting further analysis on that observation is needed in future research. Shortages and Outdates behave quite similar due to the increased production quantities. Yet, with a fixed demand, the outages are not expected to decrease as before as overproduction of blood types without demand will occur. 20 MBDFs seems to be the suggested limit for this demand scenario with close to zero shortages remaining. Both graphs illustrate a reliable and expectable behavior of the model supporting the validation of the combined simulation and optimization. Overall, the current analysis suggests, that a suchlike dynamic allocation of MBDFs is suitable to reduce shortages in the BSC especially under changing environmental influences. Limitations are the yet limited extend of simulation studies and the mixed use of WCBS and SANBS data with general population data, which can deviate from the individual organization's reality, necessitating further research.

4 Discussion and conclusion

BSCs are complex networks encompassing many legally independent actors, transporting a critical good with a very limited availability and shelf life. Nevertheless, decision support tools known in the general logistics domain are often not being applied or transferred to this domain. As such, commonly used tools such as simulation and optimization are not employed often, which is particularly detrimental in times of crises, where established action plans do not consider the changed environmental influences onto the BSC. We have proposed a simulation framework enabling an integration of optimization algorithms for strategic, tactical, and operational decision support, displaying their interactions and the influence of an actor's decision making onto the overall BSC. A suchlike system enables the BSC operators to adapt and prepare the BSC processes to the dynamically changing environmental influence, e.g., during a crisis. Furthermore, we gave the example case

of the dynamic operating location choice of MBDFs to illustrate the interactions between the different optimization types and the simulation. The simulation-based optimization enables the evaluation of the process to dynamically choose the operating locations of the MBDFS based on the current stocks of blood products across the BSC. Nevertheless, the lack of detailed case data does limit the possibilities to evaluate the overall system at the current point in time. Future work therefore needs to include studies containing more detailed case data from the individual BSC operators and facilities to enhance the evaluation of the system. Additionally, extended workshops and interviews with the stakeholders of the BSC operators need to be performed to define other relevant evaluation cases, which then can be implemented and tested also. Finally, performance analysis and improvements need to be done to ensure the applicability of the approach to support operations in practice.

Acknowledgement

This research was conducted within the project “Blood Information System for Crisis Intervention and Management” (BISKIT) funded by the German Federal Ministry of Research and Education in the context of its funding program for International Disaster and Risk Management (IKARIM).

References

- Alfonso, E.; Xie, X.; Augusto, V.; Garraud, O.: Modeling and simulation of blood collection systems. *Health care management science* 15 (2012) 1, pp. 63–78.
- Baesler, F.; Nemeth, M.; Martínez, C.; Bastías, A.: Analysis of inventory strategies for blood components in a regional blood center using process simulation. *Transfusion* 54 (2014) 2, pp. 323–330.
- Bashiri, M.; Ghasemi, E.: A Selective Covering-Inventory-Routing problem to the location of bloodmobile to supply stochastic demand of blood. *International Journal of Industrial Engineering & Production Research* 29 (2018) 2, pp. 147–158.
- Detzer, S.; Gurczik, G.; Widera, A.; Nitschke, A.: Assessment of Logistics and Traffic Management Tool Suites for Crisis Management. In: *European Transport Conference*, 2016,
- Ghasemi, P.; Khalili-Damghani, K.; Hafezalkotob, A.; Raissi, S.: Stochastic optimization model for distribution and evacuation planning (A case study of Tehran earthquake). *Socio-Economic Planning Sciences* 71 (2020), pp. 100745.
- Gunpinar, S.; Centeno, G.: An integer programming approach to the bloodmobile routing problem. *Transportation Research Part E: Logistics and Transportation Review* 86 (2016), pp. 94–115.
- Heidari, H.; Pasandideh, S.H.: Bi-objective optimization of a blood supply chain network with reliability of blood centers. *Journal of Industrial and Systems Engineering* 11 (2018) 3, pp. 190–213.
- Heidari-Fathian, H.; Pasandideh, S.H.: Modeling and solving a blood supply chain network: An approach for collection of blood. *International Journal of Supply and Operations Management* 4 (2017) 2, pp. 158–166.
- Karadağ, İ.; Keskin, M.E.; Yiğit, V.: Re-design of a blood supply chain organization with mobile units. *Soft Computing* 25 (2021) 8, pp. 6311–6327.

- Lang, J.C.: Production and inventory management with substitutions. Springer Science & Business Media 2009.
- Lowalekar, H.; Ravichandran, N.: Model for blood collections management. *Transfusion* 50 (2010) 12pt2, pp. 2778–2784.
- Melkote, S.; Daskin, M.S.: Capacitated facility location/network design problems. *European Journal of Operational Research* 129 (2001) 3, pp. 481–495.
- Osorio, A.F.; Brailsford, S.C.; Smith, H.K.: A structured review of quantitative models in the blood supply chain: a taxonomic framework for decision-making. *International Journal of Production Research* 53 (2015) 24, pp. 7191–7212.
- Osorio, A.F.; Brailsford, S.C.; Smith, H.K.; Forero-Matiz, S.P.; Camacho-Rodríguez, B.A.: Simulation-optimization model for production planning in the blood supply chain. *Health care management science* 20 (2017) 4, pp. 548–564.
- Rabe, M.; Clausen, U.: Verknüpfung von Simulation und Optimierung: Kategorien und Beispiele-Ein Bericht über die VDI Richtlinie 3633 Blatt 12 (2015).
- Rabbani, M., Aghabegloo, M., Farrokhi-A.H.: Solving a bi-objective mathematical programming model for bloodmobiles location routing problem. *International Journal of Industrial Engineering Computations* 8.1 (2017), pp 19-32.
- Ramezani, R.; Behboodi, Z.: Blood supply chain network design under uncertainties in supply and demand considering social aspects. *Transportation Research Part E: Logistics and Transportation Review* 104 (2017), pp. 69–82.
- Sachdev, S.; Singh, L.; Marwaha, N.; Sharma, R.R.; Lamba, D.S.; Sachdeva, P.: First report of the impact on voluntary blood donation by the blood mobile from India. *Asian journal of transfusion science* 10 (2016) 1, pp. 59.
- Şahinyazan, F.G.; Kara, B.Y.; Taner, M.R.: Selective vehicle routing for a mobile blood donation system. *European Journal of Operational Research* 245 (2015) 1, pp. 22–34.
- Stanworth, S.J.; New, H.V.; Apelseh, T.O.; Brunskill, S.; Cardigan, R.; Doree, C.; Germain, M.; Goldman, M.; Massey, E.; Prati, D.; others: Effects of the COVID-19 pandemic on supply and use of blood for transfusion. *The Lancet Haematology* (2020).
- Widera, A.; Konrad, C.; Böhle, C.; Hellingrath, B.: A multi-method simulation environment for humanitarian supply chains. In: 2017 4th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM), 2017a, pp. 1–8.
- Widera, A.; Lechtenberg, S.; Gurczik, G.; Bähr, S.; Hellingrath, B.: Integrated Logistics and Transport Planning in Disaster Relief Operations. In: 14th International Conference on Information Systems for Crisis Response and Management, 2017b
- Wise, R. D., Gibbs, M. W., & Louw, V. J.: Lockdown and our national supply of blood products. *SAMJ: South African Medical Journal*, 110(5), (2020).



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

System concept for semi-automated generation of layouts for simulation models based on point clouds

Systemkonzept zur teilautomatisierten Generierung von Layouts für Simulationsmodelle auf Basis von Punktwolken

Reinhard Zeiser¹, Felix Ullmann², Thomas Neuhäuser^{1,3}, Andrea Hohmann¹,
Johannes Schilp^{1,4},
reinhard.zeiser@igcv.fraunhofer.de, felix.ullmann@hoermann-rawema.de,
thomas.neuhaeuser@igcv.fraunhofer.de, andrea.hohmann@igcv.fraunhofer.de,
johannes.schilp@igcv.fraunhofer.de

¹Fraunhofer Institute for Casting, Composite and Processing Technology IGCV,
Augsburg (Germany),

²Hörmann Rawema Engineering & Consulting GmbH, Chemnitz (Germany),

³Technical University of Munich, Institute for Machine Tools and Industrial
Management (iwb), Garching b. München (Germany)

⁴University of Augsburg, Chair of Digital Manufacturing, Faculty of Applied
Computer Science, Augsburg (Germany)

Abstract: The turbulent market accounts for decreasing product lifecycles, which results in more frequent factory replanning processes. To support this process 2D and 3D simulation models are already widely used. Those virtual representations are usually generated once and never updated again. Factory planning in general needs a comprehensive data acquisition. This task usually takes up to around 50 % of the time required for planning projects. To reduce this, a continuous data management for continuous replanning is crucial. The combination of laser scanning a factory and 3D Computer Aided Design (CAD)-models of production facilities can help minimize some of the effort. Therefore, a system for semi-automated generation of layouts for simulation models based on point clouds is presented. The system consists of two modules, a model library and a module to register CAD-models of production facilities in the point cloud to obtain their position.

1 Introduction

With a share of 22.9 % of gross value added (as of 2020), the manufacturing sector is the second largest economic sector in Germany (Statistisches Bundesamt, 2019). However, companies in the sector face a turbulent market environment and are thus dependent on constant product and consequently process innovations (Abele and Reinhard, 2011; Pawellek, 2014). This means that existing production systems have to be replanned more frequently and decisions have to be made faster (Wiendahl et al., 2014). An important factor in factory replanning is the material flow. The material flow of a factory can account for 20 to 50 % of the factories operating costs, 30 % of these costs can be saved by efficient planning (Hawer et al., 2015). Besides material flow optimizations like utilization rate and order of production, optimizing the layout in regard to transport routes, accessibility and surface utilization is important. One way to support factory replanning is the use of a simulation model. For efficient replanning a location-accurate representation of the as-is condition of a factory is mandatory. In a survey, 76 % of respondents state that a detailed geometric representation plays an important role in the concept of the digital factory (Navvis, 2021). 62 % of the respondents even state that in the next two years methods for a continuous data management of the as-is condition must be created (Navvis, 2021).

A decisive basis for valid planning decisions is provided by comprehensive data acquisition and processing (VDI 5200 Blatt 1, 2011). However, this initial step is often difficult because plans are often not digital and only available in 2-dimensions (2D) and the complete information and data are distributed over several people (Sinnwell et al., 2018; Feldmann, K. and Reinhard, G. 2000). In addition, data is not continuously maintained and often does not reflect the actual as-is condition.

With recent developments in laser scanning, a location-accurate as-is condition can be determined in a relatively short time (Wunderlich and Wasmeier, 2013; Faro; Leica-Geosystems). Laser scans generate 3-dimensional (3D) point clouds in which the points have no semantics or relation to each other (Nurunnabi et al., 2012). Alongside laser scanning, automated guided vehicles (AGVs) are also developing in a vast manner (Rozsa and Sziranyi, 2018). AGVs often detect obstacles using light detection and ranging (Lidar) technology (Rozsa and Sziranyi, 2018) and thus are able to generate 3D maps of the environment. These point clouds can be used as input for continuous sensing of production systems (Neuhäuser et al., 2020) and therefore bridge the gap between factory replanning and continuous data management for location-accurate simulation models. Besides the comparatively simple task of scanning the factory, the manual effort required to prepare and make this data usable is considerable and requires a certain level of expertise. To make use of the point cloud, the positions of the production machines can be defined by measuring distances inside the point cloud and then the 3D Computer Aided Design (CAD)-models can be placed in a software tool. Some software tools support the loading and processing of point clouds. In these tools the production machines can be placed directly inside the point cloud. However, with an increasing number of production machines to be placed depending on the required accuracy, this can lead to an enormous time expenditure.

In addition, for factory replanning a certain set of tools is required to be able to analyze the as-is condition and based on the results integrate the new requirements and optimize the whole factory. Often these tools lack in interoperability and do not provide interfaces or common exchange formats, which can result in modeling the

same factory in multiple tools. For example, *plavis visTABLE* for calculating the floor utilization rate and transport routes and *Tecnomatix Plant Simulation* for analyzing the material flow. A generic approach can support the layout generation in both tools by extracting the position information of the individual machines from the point cloud. This information can then be automatically transformed for the different software tools to construct the layout for the simulation model, so the expert only has to add the software specific logic.

Therefore, in this paper, a system is presented, which is able to support the construction of the layout for a simulation model for different simulation tools. For this, the system extracts the position of production machines from a non-semantic and reference-free point cloud. This information supports the construction of a location-accurate simulation model without requiring any expertise in point cloud processing.

2 Related work

Denkena et al. (2019) describe an approach to generate a digital twin from laser scan data. Due to the complex data collection required for this approach, it is difficult to meet the requirements for frequent replanning as well as the integration with different planning tools. Stojanovic et al. (2018) address the continuous updating of a digital Building Information Modelling (BIM)-Model using point clouds. This approach uses location data extracted from the Industry Foundation Classes (IFC) file to align the point cloud and the BIM 3D geometry. This information cannot be extracted from CAD models, thus this approach is not possible. Melcher et al. classify objects based on images using a machine learning algorithm. However, the algorithm requires labeled data for training, making a generic approach difficult. In Braun and Borrmann (2019) a method is proposed for automatic image labeling based on registering a BIM-model and a point cloud generated by Photogrammetry. Bosché (2012) present a registration method consisting of a coarse and fine registration. This method is used to register a 3D model in the whole point cloud, instead of extracting the positions of objects inside a point cloud. Kim et al. (2013) also propose a method consisting of a combination of coarse and fine registration. The coarse registration is carried out using a principle component analysis (PC) and the fine registration is done by invoking the Levenberg-Marquardt iterative closest points (LM-ICP) Algorithm.

The aforementioned methods are all dealing with the automatic registration or detection of 3D models in point clouds and therefore give a good overview over different registration techniques in various domains. However, none of them is specifically concerned with supporting the factory replanning process and therefore dealing with an approach which is able to support the generation of factory layouts for different software tools. Thus a system concept for providing a generic approach to extract coordinates of objects inside a laser scan is presented, so that the creation of a location-accurate layout for simulation models of the current factory as-is condition can be achieved in different software tools.

3 System concept for semi-automated generation of layouts for simulation models based on point clouds

In this chapter the overall approach is presented to detect objects inside a point cloud, extract their positions and place their digital representatives in different software tools. The structure of this chapter is based on the process which is outlined in Fig. 1. First, the module *model library* is introduced followed by the module *layout extraction*. The *model library* is responsible for handling CAD-models, it is able to import, filter and export CAD-models to different software tools, e. g. the *layout extraction* module. This module handles the process of extracting positions from a point cloud by placing CAD-models in defined regions and gives an idea on how to construct the layout for a simulation model in different software tools. For this, the module is divided into three tasks: First, the *definition of regions*, second the four-step algorithm for the *registration of CAD-models* and third the *generation of the layout* for a simulation model.

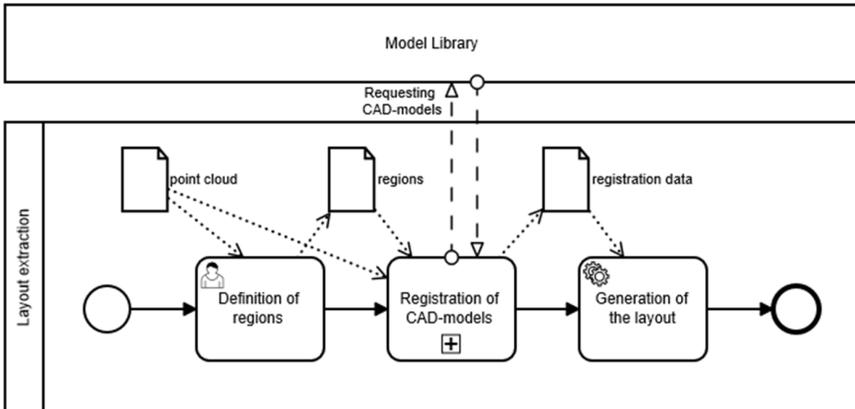


Figure 1: Process for registering CAD-models in the defined regions and extracting the factory as-is layout.

3.1 Model library

The model library is a standalone software written in the programming language python. It is designed as a central data management system for CAD-models and associated metadata to allow software tools to query specific models using an application programming interface (API).

The model library has three major functions. It is able to import new models, filter models and return a list of the filtered models and metadata.

To import a CAD-model it has to be placed in a specified directory on a file transfer protocol (FTP) server. Whilst importing, the model library automatically calculates metadata for the model based on the CAD object. The calculated metadata currently includes length, width, height and volume of the object bounding box, additionally the filename is saved. Supplementary, tags can be used to describe the object, e.g. the function of an object can be described as a “conveyor”.

The filtering function in the model library is required to reduce the computational effort in the automated matching process of CAD-models to point clouds (compare chapter 3.2.2). In order to avoid that every CAD-model is tested for a match, a preselection of possible CAD-models can be made, for example by transferring criteria like volume or maximum length of the defined region in the point cloud. In addition to the filter criteria, a relative tolerance value can be passed to compensate for differences between the CAD-models and the scanned objects but also to account for the rough bounding box. The communication with the model library is done via a Representational State Transfer (REST) Interface.

To receive a list of reference models, the filter has to be passed in a JavaScript Object Notation (JSON) format, which contains the filter criteria, the filter value and a relative tolerance. The model library applies the filter and returns the matching models as a JSON containing the absolute file paths of the CAD-model files and the associated metadata. The files can then be externally loaded and further processed.

3.2 Generation of layouts for simulation models

3.2.1 Definition of the regions

In the first step of generating the layout for a simulation model, the definition of the regions takes place in the point cloud. This task needs user interaction. A region in this context is defined as an area which encloses an object, also referred to as *target*, which has to be matched with a CAD-model. To create a region, a cuboid envelope, i.e. a bounding box, is drawn around the object. For the bounding box a rough indication of the region is sufficient. A bounding box can be drawn by just defining two points, e.g. the upper right corner and the opposite bottom left corner. This simple procedure allows a user to define many regions within a short time. An example of a defined region containing a production machine and noise can be seen in Fig. 2. After the definition of all regions, this information needs to be passed to the next step, the registration of CAD-models.

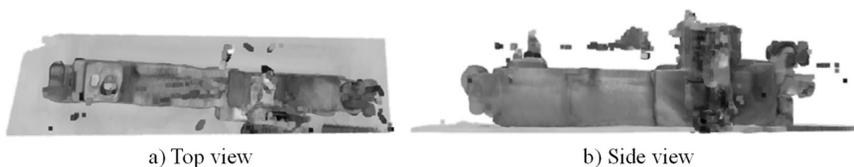


Figure 2: Defined region containing a production machine. a) displays the top view and b) is showing the machine from the side. Also the region is containing the floor and other noise.

3.2.2 Registration of CAD-models

In the CAD-model registration, the point cloud as well as the region information are imported from a local directory, then the pre-selection of the CAD-models takes place. First, the defined regions are cropped from the point cloud. Description criteria, such as width, length and height are calculated from these segments. Based on these criteria, the model library preselects suitable models by the calculated metadata. Optimally the model library returns only the matching model, but often those criteria

match multiple models and a final selection has to take place as last task in this process.

After loading, the selected CAD-models are registered in the region. For this, a two-stage algorithm is used. Since the point cloud may contain noise as well as interfering contours that do not belong to the target, the first stage is a rough placement. For this step, the Random Sample Consensus (RANSAC) is used. This algorithm is largely tolerant to noise (Schnabel et al., 2007), but the accuracy of the placement is often insufficient. With the help of the second stage, the fine registration, the model can be placed precisely in the region. For the fine registration, the Iterative Closest Points (ICP) algorithm is used. The ICP algorithm is susceptible to noise and interfering contours and thus depends on the quality of the rough registration. An example is shown in Figure 3. The left figure a) shows the registration after the RANSAC took place and the right figure b) shows the registration result after execution of RANSAC and ICP. The black solid lines are for visual indication of the target.

Based on the result of the fine registration, the Mean Squared Error (MSE) is calculated to quantify the registration quality. At this stage of development it is assumed that the region contains the best-scored model. For this CAD-model, the metadata is exported. The metadata contains the transformation information as well as the information returned from the model library, which for example contains the type of the model, e.g. “conveyor”, which can be used in some simulation software to spawn a fitting object.

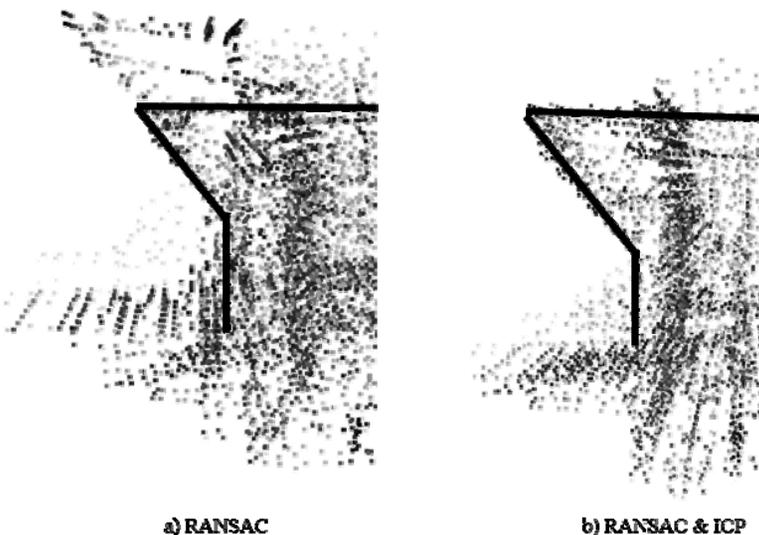


Figure 3: Example for the registration of a source CAD-model in the defined region. The left image a) shows the registration result after the application of the RANSAC-Algorithm and the right figure b) shows the registration result after execution of RANSAC and ICP. The black solid lines are for visual indication of the target.

3.2.3 Generation of the layout for a simulation model

The exported transformation and metadata of the registration process for the detected CAD-models inside the point cloud can be used to generate the layout for a simulation model. Therefore, the data has to be transformed to fit the requirements of the specific simulation tool. Below the process for transforming the data for plavis visTABLE and Tecnomatix Plant Simulation is outlined.

For plavis visTABLE the positions have to be transformed in a block layout and exported in a Microsoft Excel format (XSLX). Therefore, the CAD-models are loaded in a python program and positioned according to the transformation matrix. This representation is then used to calculate the important parameters for the block layout. Besides the x, y, z-Position in millimeters, this is the rotation of the block, width, length, height and volume of the object. Name, surface type, layer and color can be defined by the user or loaded from the model library.

Importing the Layout in Tecnomatix Plant Simulation can be achieved by writing a method in Simtalk. The method loads the metadata from disk, containing the calculated position, its rotation and the type of object to spawn, e.g. a “conveyor”. With this information, the type of object, which is loaded in form of a string has to be mapped to the internal classes of functional elements to get the class type references. Using the class type references the objects can be instantiated and transformed to the correct position and rotation.

4 Discussion

The described procedure delivers a possibility to semi automatically extract positions and information of objects out of point clouds with the goal of generating the layout for a simulation model. The defined regions are cropped from the point cloud, to increase the accuracy and speed of the registration process. Therefore, the calculations are done on a subsample. Additionally the different models returned by the model library are registered in parallel. To register and extract the positions of four identical machines with a width of 4 meters, length and height of 2.5 meters each it took under one Minute on a standard office laptop with no graphics card. This time includes registering every returned CAD-model in the region. In this specific case, 5 different CAD-models were returned by the model library for each region. Thus, the efficiency of the registering workflow depends mainly on the pre-selection of the CAD-models. Currently, the pre-selection can be optimized by precisely calculating the filter criteria or having very few similarities in the reference data. Using additional filtering criteria, e.g. key point features or characteristic faces, could also counter these problems.

In most cases the method is able to reliably place the models, but if real objects are too similar in size and shape it happens that the wrong models are placed inside a region. This can be explained by the current registration process selecting the CAD-model with the lowest MSE as the correct model. By handing this decision over to a user, two benefits, besides a growing user acceptance, can be achieved:

1. If there is no fitting CAD-model for the target, a dummy can be placed and manually be enriched with metadata, like the functionality.
2. If the placement of a CAD-model is not accurate enough, the user can trigger a recalculation of the registration process or manually improve the placement.

In terms of an automated generation of simulation models, only the structure of machines and components are captured by the laser scan data. The resulting layout is a location-accurate representation of the as-is layout. There is no automatic method to provide functionality. Nonetheless the generated layout contains enough information for calculating key performance indicators (KPIs) like the floor utilization rate and transport routes in certain simulation tools. For functionality comprehensive simulation models, additional information like material flows, movement sequences, transport sequences, simulation parameters have to be added manually by an expert.

5 Summary and outlook

The presented system concept is able to generate the layout for a simulation model based on point cloud data acquired from laser scans. The layout is generated in a user-supported process. Therefore, regions are defined in the originally context-free point cloud. In these regions, facilities are located, which in turn are matched with CAD-models via a two-stage registration algorithm and afterwards evaluated by calculating the MSE. The information obtained is then used to construct a location-accurate layout for simulation models. Besides the task of generating a layout for a simulation model, the implemented model library can be used to reduce effort, as CAD-models can be stored centrally and provided with metadata. Nevertheless, the overall benefit of the system has yet to be quantified. The presented process has laid the foundation for further work aiming at automated model generation and continuous data management in the context of continuous factory replanning.

One way to achieve this is to update existing simulation models with changes in the real production environment. As stated, replanning factories is done in ever-shorter cycles (Wiendahl et al., 2014). By using the point clouds created from AGVs as an input and additionally implementing an automatic point cloud segmentation (Nurunnabi et al., 2012), it should be possible to register the 3D CAD-models in a background process. Those results can then be used to continuously update the layout for a simulation model according to the current, real environment. Also the system can be the basis for an automatic transport time calculation depending on the mode of transport or a future collision analysis in case of a change of the plant configuration.

Acknowledgement

A grateful acknowledgement goes to the Bavarian Research Foundation (BFS), which made these research activities possible within the research project “BIMPro - Bestandsdigitalisierung zur Nutzung von Building Information Modeling in bestehenden Produktionsumgebungen”

References

- Abele, E.; Reinhard, G.: *Zukunft der Produktion: Herausforderungen, Forschungsfelder, Chancen.* s.l.: Carl Hanser Fachbuchverlag 2011.
- Bosché, F.: Plane-based registration of construction laser scans with 3D/4D building models. *Advanced Engineering Informatics* 26 (2012) 1, pp. 90–102.

- Braun, A.; Borrmann, A.: Combining inverse photogrammetry and BIM for automated labeling of construction site images for machine learning. *Automation in Construction* 106 (2019), pp. 102879.
- Denkena, B.; Stobrawa, S.; Dittrich, M.-A.; Stjepandic, J.: Automated Generation of a Digital Twin Using Scan and Object Detection for Data Acquisition (2019).
- Faro: Produktübersicht FARO FOCUS. Online verfügbar unter <https://www.faro.com/de-de/produkte/bausektor-bim-cim/faro-focus/>, zuletzt geprüft am 22.04.2021.
- Feldmann, K.; Reinhart, G. (Hrsg.): Simulationsbasierte Planungssysteme für Organisation und Produktion: Modellaufbau, Simulationsexperimente, Einsatzbeispiele. Berlin, Heidelberg, s.l.: Springer Berlin Heidelberg 2000.
- Hawer, S.; Ilmer, P.; Reinhart, G.: Klassifizierung unscharfer Planungsdaten in der Fabrikplanung. *ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb* 110 (2015) 6, pp. 348–351.
- Kim, C.; Son, H.; Kim, C.: Fully automated registration of 3D data to a 3D CAD model for project progress monitoring. *Automation in Construction* 35 (2013), pp. 587–594.
- Leica-Geosystems: Produktübersicht Laserscanner Leica Geosystems. Online verfügbar unter <https://leica-geosystems.com/de-de/products/laser-scanners/scanners>, zuletzt geprüft am 22.04.2021.
- Melcher, D.; Küster, B.; Stonis, M.; Overmeyer, L.: Optimierung von Fabrikplanungsprozessen durch Drohneneinsatz und automatisierte Layoutdigitalisierung, 2018,
- Navvis: Digital Factory Survey 2021: A time of change for manufacturing industries. Hg. v. Navvis, 2021, zuletzt geprüft am 29.04.2021.
- Neuhäuser, T.; Chen, Q.; Rösch, M.; Hohmann, A.; Reinhart, G.: Building Information Modeling im Fabriklebenszyklus. *ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb* 115 (2020) special, pp. 66–69.
- Nurunnabi, A.; Belton, D.; West, G.: Robust Segmentation in Laser Scanning 3D Point Cloud Data. In: International Conference on Digital Image Computing Techniques and Applications (DICTA), 2012, Fremantle, Australia, 12/3/2012 - 12/5/2012, 2012, pp. 1–8.
- Pawellek, G.: Ganzheitliche Fabrikplanung. Berlin, Heidelberg: Springer Berlin Heidelberg 2014.
- Rozsa, Z.; Sziranyi, T.: Obstacle Prediction for Automated Guided Vehicles Based on Point Clouds Measured by a Tilted LIDAR Sensor. *IEEE Transactions on Intelligent Transportation Systems* 19 (2018) 8, pp. 2708–2720.
- Schnabel, R.; Wahl, R.; Klein, R.: Efficient RANSAC for Point-Cloud Shape Detection. *Computer Graphics Forum* 26 (2007) 2, pp. 214–226.
- Sinnwell, C.; Haße, A.; Fischer, J.; Aurich, J.C.: Kollaborative Produktionssystemplanung unter Verwendung eines modellbasierten, PLM-gestützten Entwicklungsprozesses. *at - Automatisierungstechnik* 66 (2018) 5, pp. 406–417.
- Statistisches Bundesamt, 2019: Anteil der Wirtschaftssektoren an der Bruttowertschöpfung* in Deutschland im Jahr 2019. Online verfügbar unter <https://de.statista.com/statistik/daten/studie/36846/umfrage/anteil-der-wirtschaftsbereiche-am-bruttoinlandsprodukt/>, zuletzt aktualisiert am 01.01.2020, zuletzt geprüft am 17.02.2021.

- Stojanovic, V.; Richter, R.; Döllner, J.; Trapp, M.: Comparative visualization of BIM geometry and corresponding point clouds. *International Journal of Sustainable Development and Planning* 13 (2018) 01, pp. 12–23.
- VDI-Gesellschaft Produktion und Logistik: VDI 5200 Blatt 1, Februar 2011.
- Wiendahl, H.-P.; Reichardt, J.; Nyhuis, P.: *Handbuch Fabrikplanung: Konzept, Gestaltung und Umsetzung wandlungsfähiger Produktionsstätten*. München: Hanser 2014.
- Wunderlich, D.-I.T.; Wasmeier, D.-I.P.: *Objektivierung von Spezifikationen Terrestrischer Laserscanner – Ein Beitrag des Geodätischen Prüflabors der Technischen Universität München*. Blaue Reihe des Lehrstuhls für Geodäsie (2013) 20.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Exploiting the potential of human-machine work systems: Cycle-time reduction through simulation-based analysis and optimized task allocation

***Nutzung des Potenzials von Mensch-Maschine-Arbeitssystemen:
Taktzeitreduzierung durch simulationsbasierte Analyse und optimierte
Aufgabenverteilung***

Titanilla Komenda, Maximilian Schelle, Felix Kamhuber, Fraunhofer Austria
Research GmbH, Wien (Austria), titanilla.komenda@fraunhofer.at,
maximilian.schelle@fraunhofer.at, felix.kamhuber@fraunhofer.at
Sebastian Schlund, Technische Universität Wien, Wien (Austria),
sebastian.schlund@tuwien.ac.at

Abstract: Even though human-machine work systems seem to be a promising approach in closing the economic efficiency gap between manual production and automation, the installation rate of such kind of systems is still low. This is based on the general opinion that cycle time is only achievable on the basis of the machine's velocity. However, the velocity of a robot is often limited due to prevailing safety guidelines. In the present work, it is therefore investigated what system parameters influence cycle time in a human-machine work system and how transparency can be created so as to increase planning reliability. Based on a hybrid model approach, it is shown how the modelling of human-machine work systems can be carried out taking into account system parameters that influence cycle time. The application of the developed hybrid cycle time model is shown and validated on a circuit board packaging process.

1 Motivation and Problem Formulation

The global demand for variable batch sizes and customised product variants call for flexible production systems capable of meeting these requirements (Monostori et al. 2016). Human-machine work systems can be an enabler to this, by taking advantage of both, the robots' and humans' respective capabilities (Ranz et al. 2017). According to the International Federation of Robotics (IFR 2020), as an international umbrella organisation of all national industrial and research robotics associations, the implementation of industrial robotics systems has increased by an average of 11 % annually since 2014 (with the exception of 2019). Compared to industrial robot

systems, the share of collaborative robot systems is just under 5 %, although an increase in the installation rate of 15 % was identified in the years 2017-2019. According to a recent survey by Karlsruhe University of Applied Sciences, the use of collaborative robots in the context of work systems with human-robot collaboration is mainly suitable for reducing ergonomically stressful or monotonous work steps – an improvement in cycle time cannot be achieved (Hornung 2021). In this context, the lack of improvement in cycle time is cited as one of the major disadvantages of work systems with human-robot collaboration, which is also considered as one of the main reasons for the inhibited introduction of the concept of human-machine collaboration and thus of collaborative robots, in addition to the lack of knowledge on effective implementation.

However, the concept of human-machine collaboration would close the economic efficiency gap between purely manual production and robot- or machine-assisted production and thus make production in high-wage countries more competitive (Matthias und Ding 2013). What is unclear, are the quantifiable marginal lot sizes for the economic use of work systems with human-machine collaboration. Added to this is the lack of knowledge about the influence of system parameters. In particular, cycle time determination is usually based on the safety mode of the machine – and thus on its safety-compliant or permissible movement speed (ABB 2014). It is therefore assumed that cycle time determination is a two-dimensional problem whose solution leads to the achievement of the minimum productivity requirements.

However, by closely linking humans and machines, workflow organisation, cycle time prediction and cost estimation get more complicated due to mutually influencing system parameters and with this, trade-offs in the specified objectives. Hence, a simple to use simulation-based optimisation framework is required to effectively exploit potentials of human-machine work systems by enabling an automated comparison of system configurations based on different system parameters influencing cycle time.

The research question at hand focuses on how the planning reliability with regards to cycle time determination in human-machine work systems can be increased by creating transparency of system parameters influencing cycle time at the same time.

2 State of the Art in Cycle Time Determination

Due to spatial separation of humans and machines and the associated separation of tasks, the design of automated work systems and the design of manual work systems were considered separately. As a result, methods for determining cycle times for manual and automated tasks were also developed separately and are different. In the following, state of the art methods of cycle time determination are first listed on the basis of planning methods in the areas mentioned, in order to finally refer to methods in the field of research that already include aspects of a holistic consideration of humans and machines in a single approach.

Wloka (2013) determines cycle times of automated tasks by modelling and simulating the robot system's behaviour. For this purpose, the structure of the robot manipulator, the control behaviour as well as the work system design including the arrangement of peripheral devices are described in a model. The cycle time determination model is therefore composed of a kinematics model, a geometry model, a control model and a

simulation model. The geometry model contains all information regarding the external shape of the robot and peripheral devices in the work system. The kinematics model describes the spatial position of the individual joints of the robot during the execution of a specific robot path. The control model is used for the characteristic description of the robot's movement behaviour, such as maximum travel ranges, speeds and accelerations. In addition to communication interfaces, control properties are described in an algorithmic model component. The simulation model ultimately includes further simulation-specific model components, such as the visualisation of gripping processes.

Bokranz and Landau (2012) determine cycle times of manual tasks through process time modules, so-called Methods-Time Measurement (MTM) modules. MTM was developed in America in the 1950s and is based on empirically determined studies of manual work performance, in which the determined process times for specified activities were summarised in the form of tables. In the application of MTM, all movements performed by humans are traced back to certain basic movements (such as grasping, joining, releasing and walking) for which the required times are specified. Building on the MTM method, condensed methods have also been derived that combine the basic movements (e.g. "pick up and place", consisting of "reach, grasp, bring, join and release") and thus enable a faster and often also sufficiently accurate analysis of manual work processes. The most common condensed systems are MTM-UAS (Universal Analysing System) for series production and MTM-MEK (MTM in individual and small series production). According to Kuhlmann (2015), MTM can also be applied to transfer partial operations that have led to an overall negative assessment of manual work to machine activities – leading to an approach to task sharing in the context of collaborative activities.

Schröter et al. (2016) defined an approach for describing robot movements based on the MTM process time modules. These so-called RTM (Robot Time and Motion) building blocks contain five categorical elements, i.e. movement (reaching, moving, orienting), measuring (stopping the movement due to a sensor signal or force overshoot, information about touch or camera), grasping (picking up and placing), delay (process-specific delays or waiting times) and movement type (linear, point-to-point). With the help of the RTM modules, it is now possible to define tasks in collaborative human-machine work systems already in the planning phase with the help of a description formalism. Due to the similarity to the MTM building blocks, different task division variants can also be quantifiably compared with each other. However, the accuracy of the RTM time blocks is not satisfactory, especially for point-to-point movements of the robot, as it is strongly dependent on the control model of the respective robot system.

Gombolay et al. (2013) considered a mutual influence of temporal and spatial constraints in human-robot work systems. They demonstrate real-time task assignment and scheduling in collaborative human-machine work systems using a custom algorithm called Tercio. Tercio takes a set of tasks, temporal interval constraints, a number of agents and an objective function as input parameters. The algorithm then first computes an optimal agent assignment by solving a mixed integer problem (MIP) that contains constraints for balancing per-agent workload. Based on the agent assignment and task structure, Tercio then allocates tasks using an analytical test such that all time constraints are satisfied. Once the schedule satisfies the time constraints, agent and space resource sequencing constraints are added to the problem.

Pellegrinelli and Pedrocchi (2018) extended this approach to include the influence on the robot's motion behaviour when humans are in temporal and spatial proximity. They assume that the robot changes its movement behaviour as soon as the human enters the workspace or comes too close to the robot. In this case, the robot will not necessarily stop its movement, but will avoid the human and thus dynamically adapt its planned movement path to the local proximity of the human. For this reason, cycle time estimation of machine behaviour in the planning phase is no longer possible with conventional methods, as it depends on the specific system and resource conditions. Pellegrinelli and Pedrocchi now followed the approach of a workspace segmentation taking into account the volume of space occupied by the human and the robot during the movement. This segmentation is then used to define a set of Markov chains that describe the human-robot interaction and allow the estimation of the robot execution time.

Bänziger et al. (2018) take a similar approach to Gombolay et al. (2013). However, they use a genetic algorithm for task assignment and scheduling and describe the tasks based on MTM building blocks extended by the number of repetitions within a process. Furthermore, the model also contains information on the ergonomics of work execution. Based on a two-dimensional simulation, the local relationships of individual resources are visualised and necessary times for the resulting movement paths are calculated. To determine the optimal task assignment, the minimum cycle time is then calculated as a function of process times, waiting times and movement times based on distances to be covered, which are provided with weighting factors.

State of the art cycle time determination approaches focus on system parameters such as motion speed or task execution speed of humans and robots, their spatial relation as well as task allocation patterns but do not consider the mutual influence of those parameters in one holistic approach. Especially the dynamic change of the robot's motion speed dependent on the temporal and spatial proximity between human and robot (in order to apply with current safety regulations) are not considered. In addition, conventional planning methods also do not consider individual performance levels of humans which can be an essential benefit of the adaptive robot control with regards to human-centric planning approaches. Therefore, a cycle time determination model must consider the mutual influence of system parameters and their impact on cycle time.

3 Cycle Time Determination Model

A human-machine work system can be considered as a hybrid model consisting of both discrete and continuous elements. While the motion behaviour of humans and machines can be considered as a continuous process, specific properties changing for example the robot's velocity can be considered as discrete events. This way, a human-machine work system can be modelled as a hybrid system model with discrete and continuous model parts that influence each other depending on the specified conditions. Thus, a hybrid system model approach was used to simulate the motion behaviour of the individual resources, i.e. humans and machines, in the human-machine work system. The motion behaviour then results in execution times of individual tasks which finally leads to a cycle time.

Zeigler et al. (2000) defined a modular, hierarchical formalism for modelling and analysing discrete event systems (Discrete Event System Specification), which also

includes a formulation of continuous systems (Differential Equation Specified System) and a description formalism for hybrid systems, the so-called DEV&DESS formalism (Discrete Event and Differential Equation Specified System). The model concept includes the combination of a discrete (DEVS) and a continuous (DESS) component, which can influence each other. Pawletta et al. (2006) eventually implemented this description formalism in a toolbox in MATLAB®, which allows hybrid models to be simulated on the basis of modified discrete models with additional continuous model behaviour. This modified discrete event simulator solves ordinary differential equations between individual events to generate the continuous model behaviour (Heinzl 2020). The toolbox applies an ODE wrapper approach with the MATLAB® integrated ODE45 solver based on Runge Kutta with variable step size.

In addition, the MatlabDEVS toolbox uses the object-oriented programming of MATLAB® both for the implementation of the DEVS simulation environment and for the definition of the simulation model. The simulation model can be composed of sub models (objects) of different types (classes). By using a discrete simulator and a mapping of model elements to objects (instances), the information of the hierarchical model structure is preserved during the simulation and thus enables the simulation of the dynamic system behaviour (Deatcu and Pawletta, 2012). Thus, the hybrid system model of the human-machine work system was implemented as a hybrid simulation model consisting of individual hybrid resource models, i.e. sub models, of different types, i.e. humans and machines, which are instantiated by a specific resource class but with individual attributes and properties. The motion behaviour of those hybrid resource models was then modelled based on a uniformly accelerated motion with a trapezoidal velocity profile with an acceleration phase, a deceleration phase of equal size as well as a phase with constant velocity. Consequently, the cycle time is determined based on

$$t = \sum_{\tau=1}^m \sum_{r=1}^n x_{\tau r} t_{\tau r} \quad (1)$$

where m is the number of tasks τ and n is the number of resources r , while $x_{\tau r}$ defines the task allocation to a specific resource. In addition, one task can only be allocated once to a resource while all tasks have to be allocated to at least one resource

$$x_{\tau r} \in \{0,1\} \quad \forall \tau = 1, \dots, m, \quad \forall r = 1, \dots, n \quad (2)$$

$$\sum_{r=1}^n x_{\tau r} = 1 \quad \forall \tau = 1, \dots, m \quad (3)$$

while the task execution time by a specific resource is given by

$$t_{\tau r} = \frac{v_r^2 + a_r s_{\tau r}}{a_r v_r} \quad (4)$$

which can be determined on the basis of the resource velocity v_r and the resource acceleration a_r , where the travelled distance $s_{\tau r}$ to perform the task τ by the resource r is defined by the distance travelled in the acceleration and deceleration phase s_b and the constant speed phase s_{vk}

$$s_{\tau r} = s_b + s_{vk} + s_b = 2s_b + s_{vk} \quad (5)$$

If the task execution time determined this way is set in relation to the task execution time without an acceleration phase $\tilde{t}_{\tau r}$ – e.g. applied in the isolated model – the model quality of the hybrid model with a trapezoidal velocity profile can be determined in comparison to that simplified motion model

$$\frac{t_{\tau r}}{\tilde{t}_{\tau r}} = \frac{\frac{v_r^2 + a_r s_{\tau r}}{a_r v_r}}{\frac{s_{\tau r}}{v_r}} = \frac{\frac{v_r^3}{a_r} + v_r s_{\tau r}}{v_r s_{\tau r}} = 1 + \frac{v_r^2}{a_r s_{\tau r}} \quad (6)$$

The same applies to the comparison with higher-order s-curve models.

4 Critical system parameters

A crucial part in human-machine work systems is the technical work organisation as it opens up a complex field of planning issues due to the constant probabilistic interaction scenarios of humans and machines. Those interaction scenarios influence the spatial and temporal relation between humans and machines, thus the motion behaviour of those entities and thus influence cycle time.

Thus, one critical system parameter is the layout of the work system. By spatially dividing the work areas of humans and machines, the distances between humans and machines can be increased, which increases safety (in terms of duration of exposure to a potential hazard) as well as cycle time since longer distances have to be covered to execute potential collaborative tasks or handover-tasks. Furthermore, different types of interaction scenarios require different safety principles which again influence cycle time. They can go from stopping a machine if a human is in a specific distance to the machine, to dynamically adapting the machine's velocity as a function of the human's proximity, towards determining a safety-rated velocity of the machine so as to not exceed pre-defined biomechanical force and pressure limits during a possible collision between humans and machines. The different safety requirements also lead to different requirements in modelling accuracy with regards to the physical and dynamic behaviour of humans and machines.

Besides safety, the definition of interaction scenarios also refers to the task or resource allocation problem. Here, not only the temporal and spatial relation between humans and machines changes with different task allocation patterns but the execution time of a task can also change dependent on the allocated resource. This is not only dependent on the resource's distance at the time of task allocation but also on its individual properties, e.g. humans with different performance levels or dynamically changing performance levels due to fatigue.

5 Model Application and Validation

The cycle time determination model was applied on the design of a human-machine work system for the implementation of a PCB packaging process consisting of six different tasks, i.e. (1) scanning, (2) packaging, (3) tray handling, (4) lid handling, (5) container handling and (6) labelling. Those tasks were allocated to either one or two resources resulting in different system variants with six different task allocation patterns from fully manual (pattern A) to semi-automated (pattern B-E) to fully automated (pattern F) task execution. Furthermore, humans with five different performance levels were considered as well as robots with three different safety modes. Thus, three different system parameters, i.e. task allocation, human performance level as well as robot safety mode, alone led to 68 theoretically possible work system variants. The application of the cycle time determination model shows different cycle time values for the identified 68 system variants (Fig. 1).

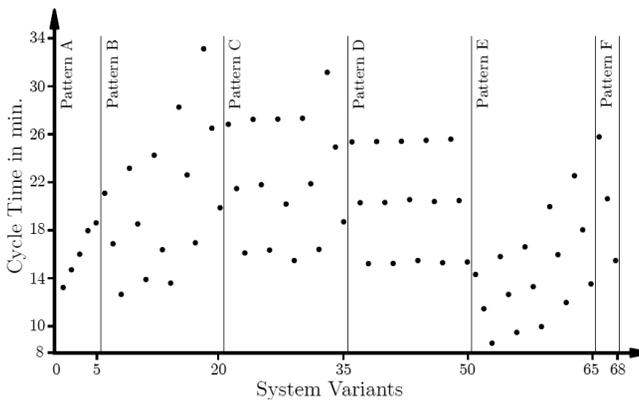


Figure 1: Cycle time values for different system variants of the PCB packaging process

The cycle time was calculated based on a packaging process of a full container containing 216 PCBs. The cycle time values reach from a minimum cycle time of 8.58 min. to a maximum cycle time of 33.12 min. which results in an average cycle time of 20.85 min. \pm 12.27 min. corresponding to a scattering of \pm 59 % of the overall cycle time values.

Tab. 1 shows the comparison of the cycle time values for one real system variant and the values determined with the methods MTM, a conventional simulation model with isolated models of humans and machines as well as with the hybrid model approach considering the mutual behaviour influence of resources.

Table 1: Comparison of cycle time values based on different determination methods

Real System	Hybrid Model	Conventional Model	MTM
15 min.	15.80 min.	16.17 min.	18 min.
100 %	105.3 %	107.8 %	120 %

6 Conclusion and Outlook

The state of the art demonstrates that there are many ways to consider human-machine interaction and collaboration and thus determine cycle time in the planning phase of human-machine work systems. The hybrid model approach based on the DEV&DESS-formalism seems to be more accurate than currently applied methods as the consideration of discrete and continuous components in the simulation model leads to a more realistic system behaviour as it considers dynamically changing human-machine interaction. The applied approach with a simplified motion behaviour representation of humans and machines based on trapezoidal velocity profiles already showed an impact. Compared to methods such as MTM or conventional simulation models with isolated model behaviour, an increased accuracy of 5.3 % in terms of the determined cycle time could be achieved. With a targeted production rate of 2.7 Mio. pieces per year this could make a difference of around 50,000 PCBs or 231 containers per year.

Even though, the needed accuracy of the cycle time is highly dependent on the implemented motion behaviour model of the simulated entities, a simulation-based optimization could help in an automatic comparison of different system variants in terms of cycle time. Accordingly, we are working on building a scalable hybrid simulation model that can consider multiple resources with different motion characteristics. Thus, the PCB packaging process might also be conducted by multiple resources such as humans, mobile robots and stationary robots with individual motion characteristics. System variants could go from manual and rigidly linked work system designs to flexible and reconfigurable layouts (Fig. 2). Especially for the reconfigurable system variants, the simulation defines an initial state for humans and machines, where the preparation and post-processing times are set accordingly generous (worst case), while allocated tasks can be executed consecutively or independently of each other (in parallel) under certain conditions. Reconfigurable system variants make the optimization problem dependent on the spatial resolution of the defined collaborative workspace. However, an initial solution is optimized step by step by a GA, which on the one hand can change the task assignment (from human to machine and vice versa) and on the other hand can reschedule the start times for each task. We are currently in the process of building and validating the simulation in combination with the GA for the presented system variants.

The results of our research show that there is need for further investigation in modelling hybrid human-machine work systems. As discussed, there is strong interconnection between the timing, trajectory, humans' individual behaviour and the layout in a human-machine work system. With the development of a simulation-integrated optimization tool, a sensitivity analysis can be achieved identifying the most crucial system parameters when it comes to specific system requirements. Furthermore, optimization might help to identify the most resilient work system variant where changing individual entities have little impact on the cycle time or other objectives.

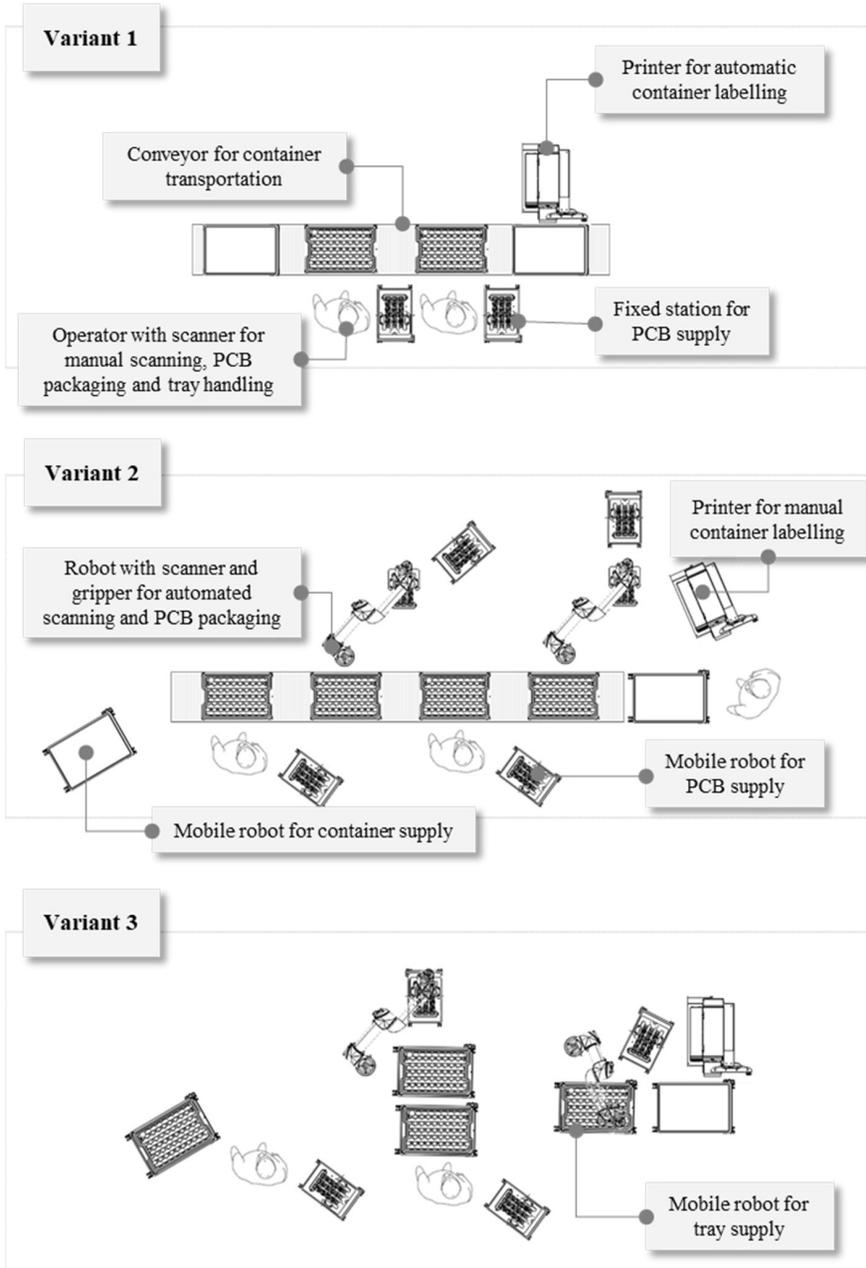


Figure 2: Scalable system variants for the PCB packaging process

Literature

- ABB Corporate Research: ABB. Industrial Safety Requirements for Collaborative Robots and Applications, 2014. https://eu-robotics.net/cms/upload/euRobotics_Forum/ERF2014_presentations/day_2/Industrial_HRC_-_ERF2014.pdf, letzter Zugriff. 17.05.2015.
- Bänziger, T.; Kunz, A.; Wegener, K.: Optimizing Human-Robot Task Allocation Using a Simulation Tool Based on Standardized Work Descriptions. *Journal of Intelligent Manufacturing* (2018).
- Bokranz, R.; Landau, K.: *Handbuch Industrial Engineering. Produktivitätsmanagement mit MTM*. Schäffer-Poeschel-Verlag, Stuttgart 2012.
- Deatcu, C.; Pawletta, T.: A Qualitative Comparison of Two Hybrid DEVS Approaches. *SNE Simulation Notes Europe 22* (2012) 1, S. 15-24.
- Gombolay, M.; Wilcox, R.; Shah, J.: *Fast Scheduling of Multi-Robot Teams with Temporospatial Constraints*. Robotics Science and Systems Foundation (2013).
- Heinzl, B.: *Methods for Hybrid Modeling and Simulation-based Optimization in Energy-Aware Production Planning*. Dissertation TU Wien, Fakultät für Mathematik und Geoinformation, 2020.
- Hornung, L.: Hochschule Karlsruhe. *Ergebnisse der Umfrage Mensch-Roboter-Kollaboration, 2021*. <https://tourings.eu/>, letzter Zugriff 25.04.2021.
- IFR: International Federation of Robotics. IFR Press Conference, 2020. https://ifr.org/downloads/press2018/Presentation_WR_2020.pdf, letzter Zugriff 27.04.2021.
- Matthias, B.; Ding, H.: Die Zukunft der Mensch-Roboter Kollaboration in der industriellen Montage. In: *Internationales Forum Mechatronik (ifm), Business Upper Austria – OÖ Wirtschaftsagentur GmbH, Linz, 2013*.
- Monostori, L.; Kadar, B.; Bauernhansl, T.; Kondoh, S.; Kumara, S.; Reinhart, G.; Sauer, O.; Schuh, G.; Sihn, W.; Ueda, K.: Cyber-physical systems in manufacturing. *CIRP Annals* 65 (2016) 2, S. 621–641.
- Pawletta, T.; Deatcu, C.; Hagedorf, O.; Pawletta, S.; Colquhoun, G.: DEVS-Based Modeling and Simulation in Scientific and Technical Computing Environments. In: Hamilton, D. (Hrsg.): *Proceedings of DEVS Integrative M&S Symposium (DEVS'06) – Part of the 2006 Spring Simulation Multiconference (SpringSim'06)*, Society for Modeling and Simulation International (SCS), Huntsville, S. 151-158.
- Pellegrinelli, S.; Pedrocchi, N.: Estimation of Robot Execution Time for Close Proximity Human-Robot Collaboration. *Integrated Computer-Aided Engineering* 25 (2018), S. 81-96.
- Ranz, F.; Hummel, V.; Sihn, W.: Capability-based task allocation in human-robot collaboration. *Procedia Manufacturing* 9 (2017), S. 182–189.
- Schröter, D.; Kuhlmann, P.; Finsterbusch, T.; Kuhrke, B.; Verl, A.: Introducing Process Building Blocks for Designing Human Robot Interaction Work Systems and Calculating Accurate Cycle Times. *Procedia CIRP* 44 (2016), S. 216-221.
- Zeigler, B.P.; Praehofer, H.; Kim, T.G.: *Theory of Modeling and Simulation. Integrating Discrete Event and Continuous Complex Dynamic Systems*, Academic Press, London, 2000.
- Wloka, D.W.: *Robotersimulation*. Springer 2013.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Modelling of microstructures during in-situ alloying in additive manufacturing for efficient material qualification processes

Modellierung von Mikrostrukturen während des in-situ Legierens in der additiven Fertigung für effiziente Materialqualifizierungsprozesse

Patrick Zimbrod, Johannes Schilp, Universität Augsburg, Augsburg (Germany),
patrick.zimbrod@informatik.uni-augsburg.de,
johannes.schilp@informatik.uni-augsburg.de

Abstract: In this work, a numerical simulation framework is presented based on the Phase Field Method that is able to capture the evolution of heterogeneous metallic microstructures during solidification. The involved physics can prove especially useful when studying not only systems undergoing thermal gradients, such as in homogeneous systems, but also in conditions that exhibit stark spatial gradients, i.e. when these inhomogeneities are present even on a mesoscopic scale. To illustrate the capabilities of the model, in-situ alloying of a High Entropy Alloy during Laser Powder Bed Fusion is investigated as an exemplary use case. The resulting digital twin is expected to shorten development times of new materials as well as cut down on experimental resource needs considerably, therefore contributing to efficient material qualification processes.

1 Introduction

The phenomenological modelling of phase evolution dates back to over 70 years and has been used widely in computational material science ever since (Cahn and Hilliard 1958). The method has especially proven its worth in the analysis of alloy solidification which has been well established for binary alloys (J. A. Warren and Boettinger 1995). With increasing computational resources, this approach has also been adapted to more demanding problems such as the solidification of polycrystalline materials (James A. Warren et al. 2003). Since then, phase field modelling has undergone a considerable rise in popularity within material science such that it is almost impossible to extensively cover all applications of this method (Steinbach 2009). Because of the valuable insight that modelling highly dynamic systems in thermodynamic non-equilibrium provides, the analysis of microstructures within additive manufacturing has also been subject to extensive efforts by various authors (DebRoy et al. 2018). Within powder bed fusion additive manufacturing, especially

the micro-scale coalescence behaviour of the present phases is of interest and has therefore been investigated (Shinagawa 2014). Beyond that, because of the stark thermal gradients typically present in those processes, the morphology of the evolving microstructure is often of particular interest, limiting the mechanical performance of the as-built alloys (Ji, Chen, and Chen 2017).

Within the normal process of powder bed fusion though, pre-alloyed powders are fed into the process and molten by a laser or electron source, yielding a comparably homogeneous distribution of the alloying elements. In contrast, in-situ alloying relies on mixing the pure elemental powder in the melting process itself (Katz-Demyanetz, Koptug, and Popov 2020). This method naturally causes local gradients with respect to the alloy species that have to be included in the modelling approach. Therefore, additional sub-models are required in the form of partial differential equations (PDE) in order to be able to solve a closed system. Additional information on this topic is given in section 3.

Hence, we develop a computational framework that covers the physical effects of these gradients in addition to the nonetheless present polycrystalline solidification. It is anticipated that this will provide a valuable framework for preliminary case studies in material qualification. This can greatly assist the aim of digitally supplemented discovery of new materials, which has already been discussed in the literature (Megahed et al. 2019). Through in-situ alloying, costly powder alloying steps can potentially be further reduced and elemental powders more universally been used (see Fig. 2). Consequently, the experimental effort could be shifted to simulative studies to some extent, hence reducing the use of valuable resources in production. This in turn has also been previously identified as a core need for modern material qualification in additive manufacturing (Seifi et al. 2016).

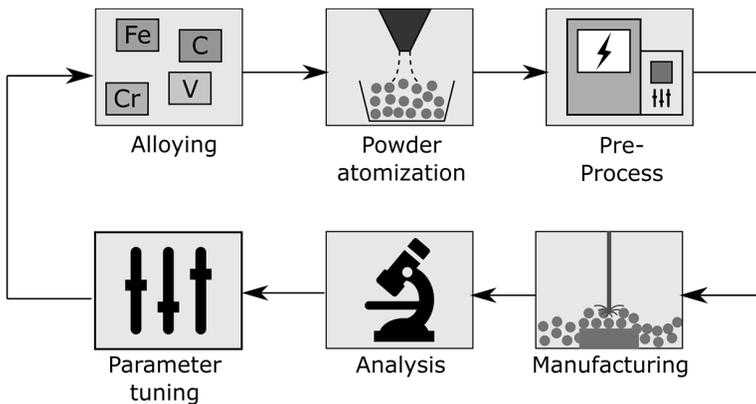


Figure 1: A simplified illustration of the typically strongly empirical and iterative material qualification process within powder bed additive manufacturing. Several pieces of specialized equipment along with substantial process knowledge is required for reliable operation. In-situ alloying typically renders the first step of the process obsolete as alloying occurs during the manufacturing step. The total amount of iterations of this process for a new material is expected to be reduced by adding the proposed digital twin approach.

In summary, in-situ alloying can be considered a logical next step regarding flexible materials discovery. By providing a computational framework to simulate microstructure evolution in these conditions, we aim to support and extend the digital material qualification by an enhanced digital twin that enables a fast discovery of novel candidates for high-performance alloys.

2 Previous Works

In-situ alloying as a means of accelerated parts qualification and increased flexibility in production has spiked interest within additive manufacturing in the recent years (Katz-Demyanetz, Koptuyug, and Popov 2020). As described previously, this approach introduces additional variability in the process that has to be addressed.

Ma et al. have conducted preliminary studies on how local inhomogeneities of this type affect microstructures in titanium using tungsten arc welding (Ma et al. 2015). Simonelli et al. investigated resulting microstructures from different kinds of powder blending. They found that a suitable arrangement of particle size distribution (PSD) between the alloying elements can impact the mechanical performance positively (Simonelli et al. 2018).

Especially the development of completely new alloys has been making use of in-situ alloying as a cost-effective technique for Design of Experiments type studies. A prominent example is the group of High Entropy Alloys (HEA) that have picked up considerable attention within additive manufacturing (Torralba and Campos 2020). Among those, several studies have been conducted using in-situ alloying, solidifying its worth in material qualification (P. Chen et al. 2020; Kim, Yang, and Lee 2020; Pegues et al. 2020; Sun et al. 2020; Gao et al. 2020).

The modelling and simulation in turn has mainly been focusing on predicting the microstructures of homogeneous powder blends. Using the phase field method in powder bed fusion, several authors provided a foundation for simulating grain evolution, mostly being coupled to mesoscale simulations to obtain the necessary information about the temperature fields (Liu and Shin 2020; Staroselsky, Acharya, and Cassenti 2020; Zielinski et al. 2020; D. Chen, Wang, and Zhang 2021; Wang et al. 2021). Therefore, it is considered a valuable addition to the current research to introduce spatial gradients into the analyses with the aim of modelling the complex physical conditions present in in-situ alloying. In particular, the investigation of inhomogeneous fields dictates the use of methods that track additional quantities in the phase field model. One way to incorporate this complexity is by using species-specific order parameters that are solved simultaneously. This so called multiphase field method has been applied successfully to solidification problems (Schmitz and Prahl 2016). However, information on orientation can hardly be tracked as each order parameter would be coupled to an individual orientation variable. Furthermore, diffusion between the phases cannot be considered which has proven to be important in form of interstitial crystal formation (Moghaddam et al. 2020). Hence, a more refined approach to this problem is needed.

3 Phase field modelling and the CALPHAD method

The starting point of a phase field model generally is the construction of a functional that describes the physical system in terms of an extensive thermodynamic variable.

This expression must commonly follow a minimum principle in order to represent an equilibrium state and can consequently be used in order to obtain equations for the spatial and temporal evolution of the system.

One can construct these equations by taking the variational derivatives of the functional in terms of the dependent variables. A characteristic feature of the phase field method is the use of a so-called phase variable or order parameter, typically denoted with ϕ . This has proven to be a useful approximation for systems undergoing solid phase transformations, as solids tend to be characterized by internal ordering of some degree (Provatas and Elder 2010). Here, we will consider the Helmholtz free energy functional \mathcal{F} which has been a popular choice besides the entropy functional (Provatas and Elder 2010). This expression for this particular application depends on the phase variable ϕ , crystallographic orientation θ as well as the species concentrations C_j of the alloying elements.

The formulation of the Helmholtz functional that is based on the works of Warren et al is given by (James A. Warren et al. 2003):

$$\mathcal{F} = \iiint_{\Omega} \left\{ f(\phi, C_j, T) + \frac{\alpha^2}{2} \Gamma(|\nabla\phi|, \theta - \psi) + sg(\phi)|\nabla\theta| + \frac{\epsilon^2}{2} h(\phi)|\nabla\theta|^2 \right\} dV \quad (1)$$

We now substitute the free energy density f with a more specific thermodynamic formulation that incorporates the variation of alloy species into the model:

$$f(\phi, C_j, T) = \sum_{j=1}^N C_j \left[\mu_j(\phi, T) + RT \ln\left(\frac{C_j}{\rho}\right) \right] \quad (2)$$

In this way, we can capture the influence of changing species concentrations as well temperature dependence through the chemical potentials in the domain. Though, we now need the respective chemical potentials μ_j as a function of phase and temperature as inputs. Furthermore, we have to track N more quantities, namely the respective species concentrations. From this point, we can compute the variational derivatives $\delta\mathcal{F}$ in terms of the dependent variables and obtain the governing equations for the problem. For sake of brevity, the details of this formulation will not be discussed here but instead further detailed in a future publication. The equations for an uncoupled system containing the variables given above are given in the literature (James A. Warren et al. 2003; Guyer et al. 2004). Computing the variational derivatives $\delta\mathcal{F}/\delta\phi$ and $\delta\mathcal{F}/\delta C_j$ using the free energy formulation in Eq. 2 allows to track the evolution of interstitial and substitutional species in the solidified alloy system, which is highly relevant for in-situ alloying, as discussed in previous sections.

In order to obtain the chemical potentials μ_j that serve as inputs in Eq. 2, one can employ the computational calculation of phase diagrams approach (CALPHAD). Although being heavily reliant on thermodynamic databases, this style of modelling is capable of calculating equilibrium states and quantities at given process conditions. The basic idea can be illustrated by constructing a mixed Gibbs free energy formulation by using the pure components $G_{A/B}$, an interaction term L_{AB} and the

thermodynamic conditions, e.g. temperature T and compositions $x_{A/B}$ (Czichos, Saito, and Smith 2006):

$$G_m = x_A G_A + x_B G_B + x_A x_B L_{AB} + RT(x_A \ln x_A + x_B \ln(x_B)) \quad (3)$$

The calculation of other thermodynamic quantities follows a similar, straightforward approach and is well-documented in the literature (Czichos, Saito, and Smith 2006).

4 Methodology

The phase field model in the present formulation contains various physical and numerical parameters that are listed in Table 1. Further details on the reasoning behind these constants as well as the numerical choice are given in the original publication by Warren et al. (James A. Warren et al. 2003). These among others control the physical effects that are present during solidification, e.g. grain rotation, anisotropy leading to dendritic growth, grain coalescence and coarsening or neck growth kinetics.

Table 1: Choice of mostly non-alloy-specific modelling parameters for the present phase field simulation (James A. Warren et al. 2003)

Quantity	Description	Value	Unit
L	Volumetric latent heat	2 e9	[J/m ³]
a	Double well potential height	31,6 e3	[(J/m ³) ^{0,5}]
ϵ	ϕ - θ interaction energy coefficient	44,7 e3	[(J/m ³) ^{0,5}]
α	ϕ gradient energy coefficient	83,9 e3	[(J/m ³)]
s	θ gradient energy coefficient	4,69 e9	[J/m ³]
c	Additional heat source coefficient to stabilize domain temperature	0,1	[1/s]
N	Order of crystal anisotropy	4	[-]
τ_ϕ	Kinetic phase scaling factor	3 e-4	[-]
τ_θ	Kinetic orientation scaling factor	3 e5	[-]
D_T	Thermal diffusivity coefficient	2,25	[m ² /s]
κ_1	Free energy phase scaling coefficient	0,9	[-]
κ_2	Free energy temperature scaling coefficient	20	[-]

The proposed digital qualification process follows the procedure depicted in Fig. 2. After information about composition, temperature and thermodynamics are gathered from CALPHAD and Computational Fluid Dynamics calculations, the phase field model tracks the evolution of the microstructure being investigated in terms of the order parameter, crystallographic orientation and species concentration.

As a result, we obtain a two-dimensional phase and orientation field that can be colour-mapped to graphically encode the grain orientation and boundaries. The morphology of the grains can then be compared to Inverse Pole Figures (IPF) from experimental data.

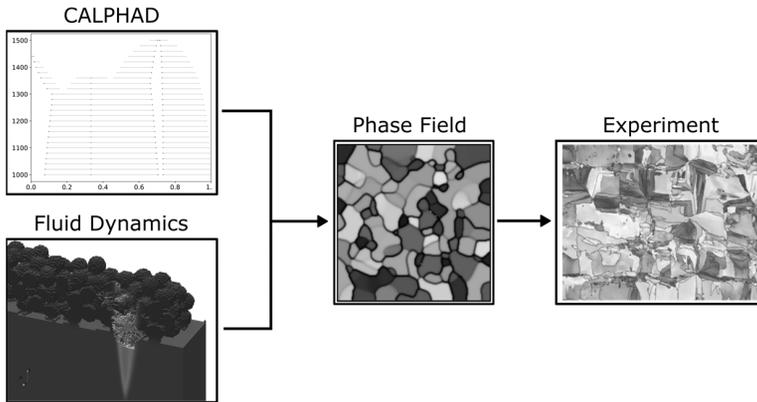


Figure 2: Illustration of the proposed computational materials engineering workflow: Thermodynamic calculations alongside with thermo-fluid-dynamics modelling provide inputs for microstructure evolution models that can be directly evaluated.

Using the computational methods presented in the theory, we create the relevant chemical potentials for the investigated quantities using the CALPHAD approach. Within this study, we used the *pyCALPHAD* library alongside with the openly available thermodynamic database for nickel-based alloys for our calculations (Otis and Liu 2017; Povoden-Karadeniz 2016).

In order to solve the phase field problem numerically, we employ the finite volume library *fipy* alongside with the *PETSc* numerics suite and its respective Python bindings (Guyer, Wheeler, and Warren 2009; Balay et al. 2019; Dalcin et al. 2011). We use simple upwinding as the flux function and solve the problem using a Jacobi-preconditioned linear solver based on the generalized minimal residual method (GMRES) for its numerical robustness and reasonable weak scaling among many processors.

5 Results and Discussion

As outlined in section 4 and illustrated in Fig. 2, we need thermodynamic data in order to appropriately incorporate the influence of species into the simulation. Otherwise, the first contribution to the integral in Eq. 1 would not be representative. The calculated chemical potentials that serve as input for the phase field model (Eq. 1) are given in Figure 3. These results represent the numerical values of $\mu_j(\phi, T)$ that were introduced in Eq. 2 and can be used directly to compute the evolution of the species concentrations.

Regarding the validity of CALPHAD calculations, it has to be noted that this method is of purely data-driven nature. The framework in itself only provides rules to interpolate given values for pure elements in a thermodynamically consistent manner. It is therefore crucially important to rely on valid data in order to create consistent results (Otis and Liu 2017). For this reason, we chose the open database offered by Povoden-Karadeniz as it contains large amounts of consolidated data from well over 200 references (Povoden-Karadeniz 2016). Despite these obvious limitations,

CALPHAD has proven to be a highly useful and relied on tool to create thermodynamic quantities for arbitrary alloys that are valid within the defined boundaries of the database considered.

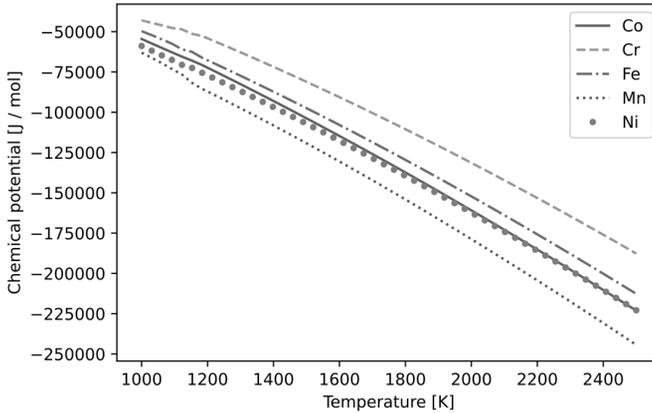


Figure 3: Chemical potentials μ_j of the investigated alloying elements between 1000 K and 2500 K. Calculations were performed using the CALPHAD method

Going on, we take the constant values given in Table 1 and assemble the remainders of the model. In this case, we simplify the input regarding boundary conditions to a constitutional undercooling of 100 K that is often found in the typically harsh process conditions during powder bed fusion (DebRoy et al. 2018). Additionally, we impose Neumann boundary conditions of the form

$$\partial u(x_i) = 0, \quad \forall x_i \in \partial\Omega \quad (4)$$

on all walls $\partial\Omega$ to create an adiabatic system. Within the simulated domain of $100 \times 100 \mu\text{m}$, we seed 100 grains at a random location and orientation in order to approximate the real grain growth behavior. For this problem, we use a grid size of $0.25 \mu\text{m}$, resulting in a total 160.000 cells.

As the process of nucleation itself is of stochastic nature both with respect to the sites and to the crystallographic orientation, the quantities output by the phase field simulation will inherently not match a given manufactured microstructure precisely. Instead, we will be looking at the grain morphologies, i.e. the shape and size distribution for a given domain. These characteristics are of vital importance in order to estimate the mechanical properties of an alloy system.

A resulting microstructure at varying timesteps is given in Fig. 4. The snapshots clearly show the originally anisotropic, dendritic growth of the initialized grains that over time coalesce and impinge with other grains. The morphology shown is not uniform owing to the present anisotropy and yields grains of noticeably different size. The dendritic arms that stretch out of the initially circular grains show the imposed four-fold anisotropy. Some of the solidified grains end up with an elongated morphology that would deteriorate fracture elongation considerably in a printed part.

These phenomena can also be observed in the printed specimens manufactured by Chen et al.

In Fig. 4 d), we observe several different grain sizes and aspect ratios, i.e. the ratio between length and width. Comparing these irregularly shaped grains with the region of interest in Fig. 4 e), we see that the structures share indeed similar characteristics regarding size and morphology. The manufactured structure by Chen et al. contains highly irregular grains that (I) have a very high aspect ratio with some grains having about 3x more length than width and (II) exhibit an almost rectangular shape. Both effects can be observed in Fig. 4 d) as well. Grains of such trapezoidal shape (II) can e.g. be observed in the lower right edge of the domain in Fig. 4 d) as well as in the lower left corner for high aspect ratio grains (I). Furthermore, one can also observe typically smaller grains with an aspect ratio much closer to unity, i.e. globular grains in the simulated structure. We therefore conclude that due to these similarities in morphology, the model captures the impingement behaviour of the nuclei in an appropriate manner.

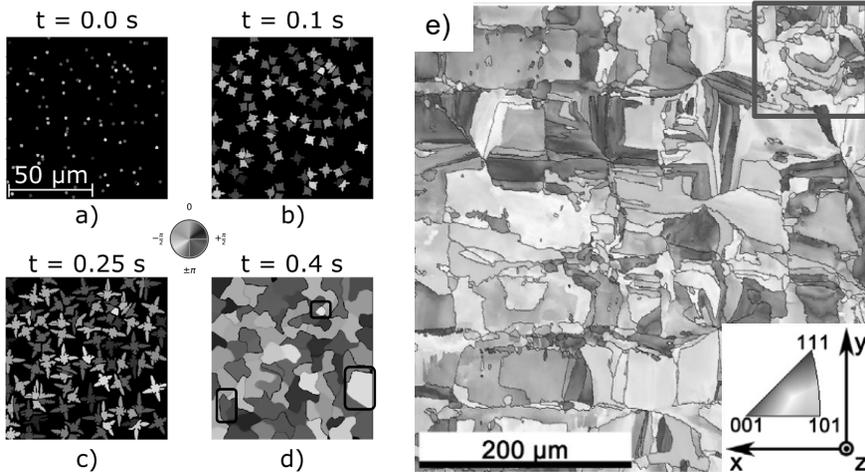


Figure 4: Results of the phase field simulation (a – d) and microstructure of the in-situ alloyed CoCrFeMnNi HEA (e) investigated by Chen et al (P. Chen et al. 2020), where a region of interest matching the simulation domain size is marked in red. One can clearly observe the originally small, circular grains (a) evolving with a dendritic growth pattern (b and c) as well as the subsequent grain impingement that forms the final microstructure (d). Grains of noteworthy morphology are indicated by black squares for illustrative purposes.

It should be noted that even at the very small scales considered in this study, the computational effort necessary to solve such models is rather large. Negligible resources are needed to obtain the physical quantities from CALPHAD, which for this study took time in the order of few minutes on a four-core mobile processor (Intel® Core i7-8550U). The calculations necessary comprise mainly light I/O operations and well-implemented simple arithmetics, as outlined by Eq. 3. The aforementioned amount of computational expense instead stems from the coupled nature of the

governing equations that arise from Eq. 1. On a workstation equipped with an 18-core/36-thread Intel® Xeon W-2295 paired with 128 GB of DDR4 RAM, the calculations up to a time of 0.4 s took approximately 200 minutes using OpenMPI. Therefore, this type of simulation using the Finite Volume Method is prohibitively expensive for conducting large parameter studies, especially because the parameter space regarding the alloying elements is typically rather large. The same rationale applies for simulating entire layers of a build process since the domain would likely be around two orders of magnitude larger in each direction. One possible next step to remedy this shortcoming could be to conduct these calculations on more parallel architectures such as state-of-the-art GPUs. It has been shown that solving the system of linear equations on GPU architectures can produce faster solutions than on pure CPU architectures (Kuo et al. 2011). However, this speedup diminishes rapidly as soon as the amount of CPU resources grows up to about 48 cores or more. This can be attributed to the rather large MPI communication overhead necessary within the Finite Volume Method. One alternative route would be to transfer the presented phase field method to more efficient numerical schemes regarding parallelization, such as the Discontinuous Galerkin Finite Element Method (DG-FEM). Using this method, much more drastic improvements regarding computational effort can be achieved that can reach up to two orders of magnitude (Kirby and Mavriplis 2020).

6 Conclusion

We have proposed a modelling approach alongside a numerical scheme and implementation of modelling in-situ alloying during powder bed fusion additive manufacturing. The model consists of an extended phase field formulation for polycrystalline materials that adds the influence of concentrations and chemical potentials to the free energy formulation. The necessary but typically hard to capture physical quantities are determined using the CALPHAD method. Simulation outputs qualitatively resemble real EBSD graphs that can be captured after processing in-situ alloyed High Entropy Alloys via laser powder bed fusion. This digital twin can therefore be considered a valuable addition to the discovery of new high-performance materials that can be processed via additive manufacturing. Further studies can and should be conducted upon the initial parameter choice, as these impact the quantitative occurrence of grain rotation, coarsening etc. considerably. It has been furthermore outlined that this specific numerical implementation of the phase field method has room for improvement regarding the use in high-throughput parameter studies. As an outlook, two pathways of improving the computational efficiency in the form of GPU parallelization and developing of more advanced numerical schemes have been outlined.

References

- Balay, S, S Abhyankar, M Adams, J Brown, P Brune, K Buschelman, L Dalcin, et al. 2019. 'PETSc Users Manual'. Argonne National Laboratory.
- Cahn, John W., and John E. Hilliard. 1958. 'Free Energy of a Nonuniform System. I. Interfacial Free Energy'. *The Journal of Chemical Physics* 28 (2): 258–67. <https://doi.org/10.1063/1.1744102>.

- Chen, Dongxu, Junsheng Wang, and Chi Zhang. 2021. 'Coupling Phase-Field Model and CFD for Hot Cracking Predictions of Al-Li Alloys'. *Computational Materials Science* 192 (May): 110361. <https://doi.org/10.1016/j.commatsci.2021.110361>.
- Chen, Peng, Chao Yang, Sheng Li, Moataz M. Attallah, and Ming Yan. 2020. 'In-Situ Alloyed, Oxide-Dispersion-Strengthened CoCrFeMnNi High Entropy Alloy Fabricated via Laser Powder Bed Fusion'. *Materials & Design*, January, 108966. <https://doi.org/10.1016/j.matdes.2020.108966>.
- Czichos, Horst, Tetsuya Saito, and Leslie Smith, eds. 2006. *Springer Handbook of Materials Measurement Methods*. Springer Handbooks. Berlin: Springer.
- Dalcin, Lisandro D., Rodrigo R. Paz, Pablo A. Kler, and Alejandro Cosimo. 2011. 'Parallel Distributed Computing Using Python'. *Advances in Water Resources* 34 (9): 1124–39. <https://doi.org/10.1016/j.advwatres.2011.04.013>.
- DebRoy, T., H. L. Wei, J. S. Zuback, T. Mukherjee, J. W. Elmer, J. O. Milewski, A. M. Beese, A. Wilson-Heid, A. De, and W. Zhang. 2018. 'Additive Manufacturing of Metallic Components – Process, Structure and Properties'. *Progress in Materials Science* 92 (March): 112–224. <https://doi.org/10.1016/j.pmatsci.2017.10.001>.
- Gao, Xiaoyu, Zejiang Yu, Weihua Hu, Yi Lu, Zeyu Zhu, Yu Ji, Yunzhuo Lu, Zuoxiang Qin, and Xing Lu. 2020. 'In Situ Strengthening of CrMnFeCoNi High-Entropy Alloy with Al Realized by Laser Additive Manufacturing'. *Journal of Alloys and Compounds*, January, 156563. <https://doi.org/10.1016/j.jallcom.2020.156563>.
- Guyer, J. E., W. J. Boettinger, J. A. Warren, and G. B. McFadden. 2004. 'Phase Field Modeling of Electrochemistry. II. Kinetics'. *Physical Review E* 69 (2): 021604. <https://doi.org/10.1103/PhysRevE.69.021604>.
- Guyer, J. E., D. Wheeler, and J. A. Warren. 2009. 'FiPy: Partial Differential Equations with Python'. *Computing in Science Engineering* 11 (3): 6–15. <https://doi.org/10.1109/MCSE.2009.52>.
- Ji, Yanzhou, Lei Chen, and Long-Qing Chen. 2017. 'Understanding Microstructure Evolution During Additive Manufacturing of Metallic Alloys Using Phase-Field Modeling'. In *Thermo-Mechanical Modeling of Additive Manufacturing*, edited by Michael Gouge and Pan Michaleris, 93–116. Oxford; Cambridge, MA: BH Butterworth-Heinemann an imprint of Elsevier. <https://doi.org/10.1016/B978-0-12-811820-7.00008-2>.
- Katz-Demyanetz, A., A. Koptug, and V. V. Popov. 2020. 'In-Situ Alloying as a Novel Methodology in Additive Manufacturing'. In *2020 IEEE 10th International Conference Nanomaterials: Applications Properties (NAP)*, 02SAMA05-1-02SAMA05-4. <https://doi.org/10.1109/NAP51477.2020.9309652>.
- Kim, Young-Kyun, Sangsun Yang, and Kee-Ahn Lee. 2020. 'Compressive Creep Behavior of Selective Laser Melted CoCrFeMnNi High-Entropy Alloy Strengthened by in-Situ Formation of Nano-Oxides'. *Additive Manufacturing* 36 (January): 101543. <https://doi.org/10.1016/j.addma.2020.101543>.
- Kirby, Andrew C., and Dimitri J. Mavriplis. 2020. 'GPU-Accelerated Discontinuous Galerkin Methods: 30x Speedup on 345 Billion Unknowns'. <http://arxiv.org/pdf/2006.15698v2>.
- Kuo, Fang-An, Matthew R. Smith, Chih-Wei Hsieh, Chau-Yi Chou, and Jong-Shinn Wu. 2011. 'GPU Acceleration for General Conservation Equations and Its Application to Several Engineering Problems'. *Computers & Fluids*, 22nd

- International Conference on Parallel Computational Fluid Dynamics (ParCFD 2010), 45 (1): 147–54. <https://doi.org/10.1016/j.compfluid.2010.10.007>.
- Liu, Shunyu, and Yung C. Shin. 2020. ‘Integrated 2D Cellular Automata-Phase Field Modeling of Solidification and Microstructure Evolution during Additive Manufacturing of Ti6Al4V’. *Computational Materials Science* 183 (January): 109889. <https://doi.org/10.1016/j.commatsci.2020.109889>.
- Ma, Yan, Dominic Cuiuri, Nicholas Hoye, Huijun Li, and Zengxi Pan. 2015. ‘The Effect of Location on the Microstructure and Mechanical Properties of Titanium Aluminides Produced by Additive Layer Manufacturing Using In-Situ Alloying and Gas Tungsten Arc Welding’. *Materials Science and Engineering: A* 631 (April): 230–40. <https://doi.org/10.1016/j.msea.2015.02.051>.
- Megahed, Mustafa, Hans-Wilfried Mindt, Jörg Willems, Paul Dionne, Lars Jacquemeton, James Craig, Piyush Ranade, and Alonso Peralta. 2019. ‘LPBF Right the First Time—the Right Mix Between Modeling and Experiments’. *Integrating Materials and Manufacturing Innovation* 8 (2): 194–216. <https://doi.org/10.1007/s40192-019-00133-8>.
- Moghaddam, Ahmad Ostovari, Nataliya A. Shaburova, Marina N. Samodurova, Amin Abdollahzadeh, and Evgeny A. Trofimov. 2020. ‘Additive Manufacturing of High Entropy Alloys: A Practical Review’. *Journal of Materials Science & Technology*, November. <https://doi.org/10.1016/j.jmst.2020.11.029>.
- Otis, Richard, and Zi-Kui Liu. 2017. ‘Pycalphad: CALPHAD-Based Computational Thermodynamics in Python’. *Journal of Open Research Software* 5 (January): 1. <https://doi.org/10.5334/jors.140>.
- Pegues, Jonathan W., Michael A. Melia, Raymond Puckett, Shaun R. Whetten, Nicolas Argibay, and Andrew B. Kustas. 2020. ‘Exploring Additive Manufacturing as a High-Throughput Screening Tool for Multiphase High Entropy Alloys’. *Additive Manufacturing*, January, 101598. <https://doi.org/10.1016/j.addma.2020.101598>.
- Povoden-Karadeniz, Erwin. 2016. *MatCalc Ni Database* (version 2.034). Wien: TU Wien. https://www.matcalc.at/images/stories/Download/Database/mc_ni_v2.034.tdb.
- Provatas, Nikolas, and Ken Elder. 2010. *Phase-Field Methods in Materials Science and Engineering*. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA. <https://doi.org/10.1002/9783527631520>.
- Schmitz, Georg J., and Ulrich Prahl. 2016. *Handbook of Software Solutions for ICME*. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA. <https://doi.org/10.1002/9783527693566>.
- Seifi, Mohsen, Ayman Salem, Jack Beuth, Ola Harrysson, and John J. Lewandowski. 2016. ‘Overview of Materials Qualification Needs for Metal Additive Manufacturing’. *JOM* 68 (3): 747–64. <https://doi.org/10.1007/s11837-015-1810-0>.
- Shinagawa, Kazunari. 2014. ‘Simulation of Grain Growth and Sintering Process by Combined Phase-Field/Discrete-Element Method’. *Acta Materialia* 66 (January): 360–69. <https://doi.org/10.1016/j.actamat.2013.11.023>.
- Simonelli, Marco, Nesma T. Aboulkhair, Philip Cohen, James W. Murray, Adam T. Clare, Chris Tuck, and Richard J. M. Hague. 2018. ‘A Comparison of Ti-6Al-4V in-Situ Alloying in Selective Laser Melting Using Simply-Mixed and Satellited Powder Blend Feedstocks’. *Materials Characterization*, Metal Additive

- Manufacturing: Microstructures and Properties, 143 (September): 118–26. <https://doi.org/10.1016/j.matchar.2018.05.039>.
- Staroselsky, Alexander, Ranadip Acharya, and Brice Cassenti. 2020. ‘Development of Unified Framework for Microstructure, Residual Stress, and Crack Propensity Prediction Using Phase-Field Simulations’. *International Journal of Computational Methods and Experimental Measurements* 8 (2): 111–22. <https://doi.org/10.2495/CMEM-V8-N2-111-122>.
- Steinbach, Ingo. 2009. ‘Phase-Field Models in Materials Science’. *Modelling and Simulation in Materials Science and Engineering* 17 (7): 073001. <https://doi.org/10.1088/0965-0393/17/7/073001>.
- Sun, Zhongji, Xipeng Tan, Chengcheng Wang, Marion Descoins, Dominique Mangelinck, Shu Beng Tor, Eric A. Jägle, Stefan Zaefferer, and Dierk Raabe. 2020. ‘Reducing Hot Tearing by Grain Boundary Segregation Engineering in Additive Manufacturing: Example of an AlxCoCrFeNi High-Entropy Alloy’. *Acta Materialia*, November, 116505. <https://doi.org/10.1016/j.actamat.2020.116505>.
- Torralba, José M., and Mónica Campos. 2020. ‘High Entropy Alloys Manufactured by Additive Manufacturing’. *Metals* 10 (5): 639. <https://doi.org/10.3390/met10050639>.
- Wang, X., L. Yuan, L. Li, C. O. Yenusah, Y. H. Xiao, and L. Chen. 2021. ‘Multi-Scale Phase-Field Modeling of Layer-by-Layer Powder Compact Densification during Solid-State Direct Metal Laser Sintering’. *Materials & Design*, February, 109615. <https://doi.org/10.1016/j.matdes.2021.109615>.
- Warren, J. A., and W. J. Boettinger. 1995. ‘Prediction of Dendritic Growth and Microsegregation Patterns in a Binary Alloy Using the Phase-Field Method’. *Acta Metallurgica et Materialia* 43 (2): 689–703. [https://doi.org/10.1016/0956-7151\(94\)00285-P](https://doi.org/10.1016/0956-7151(94)00285-P).
- Warren, James A., Ryo Kobayashi, Alexander E. Lobkovsky, and W. Craig Carter. 2003. ‘Extending Phase Field Models of Solidification to Polycrystalline Materials’. *Acta Materialia* 51 (20): 6035–58. [https://doi.org/10.1016/S1359-6454\(03\)00388-4](https://doi.org/10.1016/S1359-6454(03)00388-4).
- Zielinski, Jonas, Guillaume Boussinot, Gottfried Laschet, Markus Apel, and Johannes Henrich Schleifenbaum. 2020. ‘From Melt Pool to Microstructure to Mechanical Properties: A Simulative Approach to L-PBF Processed Material Behaviour’. *BHM Berg- Und Hüttenmännische Monatshefte* 101 (January): 697. <https://doi.org/10.1007/s00501-020-00965-4>.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Dispo 4.0 | Simulationsbasierte Optimierung von Bestelllosgrößen in der verbrauchsgesteuerten Materialdisposition der Investitionsgüterindustrie

Dispo 4.0 | Simulation-based optimisation of purchase orders in consumption-based material disposition of the capital good industry

Alexander Schmid, Thomas Sobottka, Fraunhofer Austria Research GmbH, Wien (Austria), alexander.schmid@fraunhofer.at, thomas.sobottka@fraunhofer.at

Magdalena Lielacher, Wilfried Sihm, TU Wien, Wien (Austria), e1526704@student.tuwien.ac.at, wilfried.sihm@tuwien.ac.at

Abstract: With materials planning today facing a volatile and complex market environment, the trend towards automated processes in procurement is the key to market competitiveness. Digitisation in procurement and especially algorithms for lot-sizing problems can support decisions and achieve major cost savings. Although simulation-based optimisation, based on lot-sizing algorithms, offers great potential, these digital tools are not yet widely used in industry. This paper develops a simulation-based optimisation of purchase orders and presents a case-study evaluation in the capital good industry, using 13 different lot-sizing algorithms. The results point to a significant cost savings potential for the developed approach.

1 Einleitung

1.1 Motivation & Simulationsbezug

Die Materialdisposition bzw. der operative Einkauf befindet sich in einem volatilen, komplexen Marktumfeld im Wandel. Zunehmende Globalisierung und Entwicklung des Marktes hin zu einem Individualmarkt mit kundenspezifischer Gestaltung von Produkten, ständiger Verfügbarkeit von Waren und Informationen (Wischmann und Hartmann 2018) sowie durch die fortschreitende Digitalisierung der Arbeitswelt treiben die Komplexität der Beschaffung. Eine Studie des deutschen und des österreichischen Bundesverbands Materialwirtschaft, Einkauf und Logistik zeigt einen Trend zur Automatisierung operativer Prozesse im Einkauf auf (Bogaschewsky und Müller 2018). Die HTWK Leipzig zeigt in ihrer Studie, dass durch Digitalisierung der Beschaffung Einsparungen von Prozesskosten im Unternehmen um fast 50% möglich sind (Müller 2017).

Trotz des erkannten großen Nutzenpotentials der Digitalisierung und der Entwicklung einer Vielzahl von Modellen und Methoden zur Berechnung optimaler Bestelllosgrößen in den letzten Jahrzehnten (siehe Abbildung 1), wird in vielen Unternehmen noch „mit der Hand am Arm“ disponiert und über Bestelltermine und -mengen, ohne Nutzung von Verfahren zur Bestelllosgrößenberechnung, manuell entschieden (Kemmer und Sames 2015). Optimierungsverfahren und Algorithmen unterstützen Materialdisponenten/innen bei der Dispositionsentscheidung und bilden die Grundlage für eine Digitalisierung und Automatisierung der dispositiven Prozesse. Laut Minner und Stößlein (Claus et al. 2015) ist simulationsbasierte Optimierung besonders für die Optimierung und Parametrisierung von Dispositionsregeln geeignet. Durch eine simulationsbasierte Optimierung lassen sich Systemdynamik wie auch Informationsdynamiken vielfach angemessen abbilden und darstellen. Mit einer simulationsbasierten Optimierung können Arbeitsweisen und Erfahrungswerte der Materialdisponenten/innen abgebildet, bewährte Algorithmen allgemein nutzbar gemacht und die Nachvollziehbarkeit von Entscheidungen sichergestellt werden. Simulation kann dabei entweder als Bewertungsfunktion einer Optimierung aufgerufen werden, oder die Optimierung kann als optimierter Algorithmus in den Ablauf der Simulation integriert werden (März et al. 2011). Die Simulation ist dabei, im Kontext von Abläufen in Produktion und Logistik, typischerweise eine dynamische Discrete-Event Simulation, z.B. kombiniert mit einer Optimierung mittels Heuristiken und Metaheuristiken (Kamhuber et al. 2019), oder einer Form statischer Simulation, bzw. Terminierung auf Basis determinierten Systemverhaltens.

Dieser Anwendungsbeitrag stellt die Entwicklung eines digitalen Planungswerkzeugs für den operativen Einkauf vor, das eine produktspezifische optimierte Disposition ermöglicht. Dafür wird eine Optimierung auf Basis einer statischen Simulation als Bewertungsfunktion entwickelt. Das Nutzenpotenzial wird in einer Fallstudie evaluiert.

1.2 Zielsetzung und Forschungshypothese

Zielsetzung dieses Beitrags ist es, kostenoptimale Bestellmengen und Bestellzeitpunkte auf Artikelebene auf Grundlage zukünftig prognostizierter Bedarfe und einer Variation dispositiver Parameter zu ermitteln.

Die Entwicklung und der Einsatz einer digitalen Planungsmethode der Materialdisposition mit simulationsbasierter Optimierung und Nutzung von State-of-the-art Verfahren der Bestelllosgrößenbestimmung, unter Einbezug einer „*Total Landed Cost*“-Betrachtung (alle Kosten in Zukaufartikelbeschaffung), kann erhebliche Kostenvorteile im Praxiseinsatz gegenüber den heute praktizierten Bestelllosgrößenermittlungen in den Unternehmen der Investitionsgüterindustrie erzielen.

2 Überblick Bestelllosgrößenplanung

In einer Literaturanalyse wurde zunächst eine Übersicht der verfügbaren Losgrößenverfahren ermittelt, die Algorithmen charakterisiert und die Einsatzmöglichkeiten im betrieblichen Umfeld der Investitionsgüterindustrie evaluiert. Die Abbildung 1 zeigt das Ergebnis der ermittelten Verfahren. Die

dunkelgrau markierten Verfahren wurden nach einer Häufigkeitsanalyse in der Literatur als die meist-gebrauchten Verfahren ausgewählt und in der entwickelten Planungsmethode berücksichtigt. Diese Algorithmen finden teilweise schon Anwendung in ERP-Systemen. Jedoch fehlt Verantwortungsträgern/innen in den Unternehmen eine Entscheidungsgrundlage, welche der Bestelllosgrößenverfahren für des Unternehmen die geeignetsten und wie diese zu parametrisieren sind.

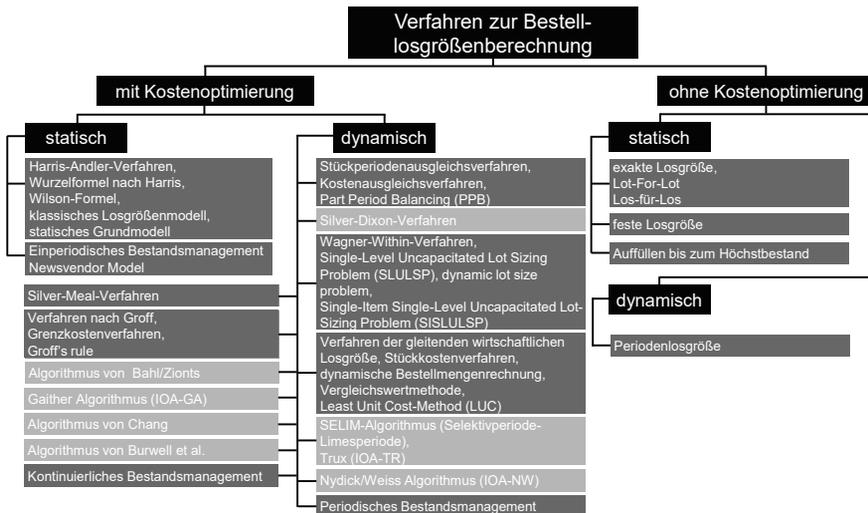


Abbildung 1: Übersicht Verfahren zur Bestelllosgrößenberechnung – eigene Darstellung als Ergebnis der Literaturanalyse (Schuh und Schmidt 2014; Zoller und Robrade 1987; Chang 2013; Brabänder 2020; Heiserich et al. 2011; Singhal und Singhal 2009; Robrade 2013; Prabhu et al. 2013; Grabner 2019)

3 Charakterisierung der Fallstudie

Der Aufbau der simulationsbasierten Optimierung gliedert sich in die folgenden drei wesentlichen Schritte:

1. Datencharakterisierung und Datenaufbereitung
2. Auswahl optimaler Bestelllosgrößenverfahren durch simulationsbasierte Optimierung
3. Optimierte Parametrisierung und prognosebasierte Simulation von Bestellmengen und -zeitpunkten

Die auf historischen Daten beruhende simulationsbasierte Optimierung und die darauffolgende, auf prognostizierenden Bedarfen folgende, Simulation wurden in einem VBA-basierten MS-Excel-Tool umgesetzt. Zielsetzung war es, dass Anwender/innen (Materialdisponent/innen, operativer Einkäufer/innen) die Planung für Artikel auf Basis einer „Total Landed Cost“-Betrachtung selbstständig durchführen können, ohne eigenes Expertenwissen im Bereich Simulation. Die Fallstudie wurde mit den dispositionsrelevanten Daten eines Unternehmens der Investitionsgüterindustrie [Branche: Erzeugung von Armaturen und Ventilen

(Branchencode: 25990 - Herstellung von sonstigen Metallwaren a.n.g.) durchgeführt. Das in einen Konzern eingebettete Unternehmen der Fallstudie hat ca. 115 Mitarbeiter, einen Jahresumsatz von 22,3 Mio.€, mit 346 Kunden aus 51 Ländern sowie ein jährliches Einkaufsvolumen von 11,5 Mio.€, 1.780 aktive Lieferanten aus 61 Ländern, und nutzt ein ERP als zentrales IT-System. Für die Fallstudie müssen die Inputdateien in einer standardisierten Form aus den verwendeten IT-Systemen des Unternehmens vorliegen und werden über eine Schnittstelle eingelesen.

4 Inputdaten und Datenaufbereitung

Im Folgenden wird, nach einer grundsätzlichen Einführung in den *Total Landed Cost* Ansatz, die Vorgehensweise der Datenerhebung sowie die Datenstruktur und Ergebnisse beschrieben.

4.1 „Total Landed Cost“ Betrachtung

Um die gesamtheitlichen Kosten des Beschaffungsprozesses für die Wahl optimaler Bestelllosgrößen ermitteln zu können, wurde das *Total Landed Cost* (TLC) Konzept herangezogen. Neben dem Kaufpreis werden ebenso Transport-, Wareneingangs-, anschließende Lager- sowie Kapital- und eventuelle Unterbestandskosten berücksichtigt. Der Vorteil von TLC ist hierbei auch das Offenlegen von versteckten und indirekten Kosten, die ohne detaillierte Analyse auf der Basis von umfassenden Daten vor der Beschaffung nicht sichtbar gewesen wären (Pumpe und Vallée 2015; Blecker et al. 2015). TLC ist ein Teilgebiet des *Total Cost of Ownership* Ansatzes, das zusätzlich zu den Kosten der Beschaffung auch jene der Nutzung und Entsorgung eines Produktes erfasst.

4.2 Vorgehensweise und Datenbeschreibung

In einem ersten Schritt werden auf Grundlage der 13 ausgewählten Bestelllosgrößenverfahren (siehe Abbildung 1) alle für die Anwendung der Algorithmen relevanten Daten (Material-, Dispositionsstammdaten, Verbrauchsdaten, Kosten etc.) identifiziert und aus den IT-Systemen des/der Anwenders/in aus der Investitionsgüterindustrie bezogen. Dabei werden insgesamt zehn verschiedene Dateien in drei unterschiedlichen Dateiformaten (.xlsx, .csv und .txt) mittels eigens entwickelter Importlogik in dem VBA-basierten MS-Excel-Tool zusammengeführt, nach Artikelnummer sortiert sowie für die weitere Nutzung aufbereitet.

4.3 Datenstruktur und Ergebnisse

In der folgenden Abbildung sind die für die simulationsbasierte Optimierung sowie prognosebasierte Simulation notwendigen Daten dargestellt. Zusätzlich zu den aufgelisteten Angaben wird die Artikel- bzw. Lieferantenummer bei jeder Datei für die korrekte Sortierung importiert.

Die Aufteilung der Daten innerhalb der Dateien sowie deren Bezeichnung erfolgte vom Unternehmen selbst auf Basis der verfügbaren Exportmöglichkeiten des ERP-Systems.

Input für simulationsbasierte Optimierung			
Dateninput 1 -Preismenge -Wareneingangszeit -Standardpreis -Höchstbestand	Dateninput 2 -Auswahl fester Lieferant -Lieferant	Dateninput 3 -Warenbewegungen vergangene 5 Jahre auf Wochenbasis	Dateninput 4 -Bezeichnung des Lieferanten
Dateninput 5 -Disponent -Warengruppe -Dispomerkmal -Materialkurztext -Basismengeneinheit -Sicherheitsbestand -Lagerbestand -Feste Losgröße / Rundungsmengen -Meldebestand aus SAP	Dateninput 6 -Gelieferte Menge -Bestellmenge -Bestellzeitpunkt auf Wochenbasis -Lieferdatum	Dateninput 7 -Lieferant -Nettopreis -Planlieferzeit -Einkäufer -Höchstbestellmenge -Bestellbelege -Währung	Dateninput 8 -Lieferantenseitige fixe Bestellkosten -Variable Stückkosten -Prozesskosten Einkauf und Materialdisposition -Unterbestandskosten -Buchhaltungskosten -Wareneingangskosten -Lagerkosten -Kapitalbindungskosten
Input für prognosebasierte Simulation			
Dateninput 9 -Prognosewerte der zukünftigen 12 Monate		Dateninput 10 -Lieferzeit Modell	

Abbildung 2: Inputdaten

5 Auswahl optimaler Bestelllosgrößenverfahren durch simulationsbasierte Optimierung

Innerhalb des zweiten Schrittes der Fallstudie wird die Auswahl optimaler Bestelllosgrößenverfahren durch simulationsbasierte Optimierung durchgeführt. Die Auswahl erfolgt dabei auf Basis der minimalen Kosten.

5.1 Verwandte Arbeiten

Beck et al. (2015) führten eine Performanceanalyse von dynamischen Losgrößen durch. Dabei wurden die *Verfahren nach Groff*, *Silver-Meal*, *Stückperiodenausgleich* sowie *gleitende wirtschaftliche Losgröße* angewendet. Als Datenbasis diente ein Simulationsaufbau mit 5100 Datensätzen und 300 Perioden. Das Verfahren nach Groff und Silver-Meal zeigten sich als am geeignetsten für die größte Datenbasis. Diese Arbeit bietet allerdings nicht die Anwendung von realen Daten eines konkreten Unternehmens. Des Weiteren wurden nur vier Verfahren angewendet und das mögliche Einsparungspotential wurde nicht betrachtet.

Ein weiterer Artikel von (Alharkan et al., 2020) beschäftigt sich mit simulationsbasierter Optimierung mit dem Ziel die Gesamtkosten über die Optimierung der Bestellkosten zu verringern. Des Weiteren existiert noch ein Anwendungsfall mittels Randomized Ordering Policy (Ahmadi et al., 2018). Auch hier werden nicht mehrere Verfahren gegenübergestellt.

5.2 Vorgehensweise und Modellbeschreibung

Nach der Datenaufbereitung werden aus den vorhandenen historischen Daten für jeden Artikel relevante logistische Kenngrößen (Gesamtlieferzeit, Meldebestand, Ermittlung des Hauptlieferanten etc.) für die Optimierung berechnet. Auf Grundlage der vergangenen Verbräuche und des durchgeführten Bestellverhaltens der Materialdisponenten/innen werden dann die 13 Bestelllosgrößenverfahren anhand einer „Total Landed Cost“ (TLC) Betrachtung rückwirkend auf Artekebene mit dem IST-Bestellverhalten vergleichend angewendet, bewertet und jeweils das kostenoptimale Verfahren für alle Artikel bestimmt. Da statische Bestelllosgrößenverfahren nur optimale Bestellmengen liefern und im Gegensatz dazu dynamische Bestelllosgrößenverfahren optimale Bestellmengen und optimale Bestellzeitpunkte ausgeben, wird eine statische Simulation für die Bewertung ausgeführt, um dynamische und statische Bestelllosgrößenverfahren miteinander vergleichen zu können. Die simulationsbasierte Optimierung verfügt dabei über folgende Stellgrößen: Wahl des Prognosehorizonts von 1 bis 53 Wochen, das Aussortieren von Verfahren und das Einstellen von Raffungsparametern für das Verfahren *Periodenlosgröße*. Die Abbildung 3 stellt die Vorgehensweise bei der simulationsbasierten Optimierung dar.

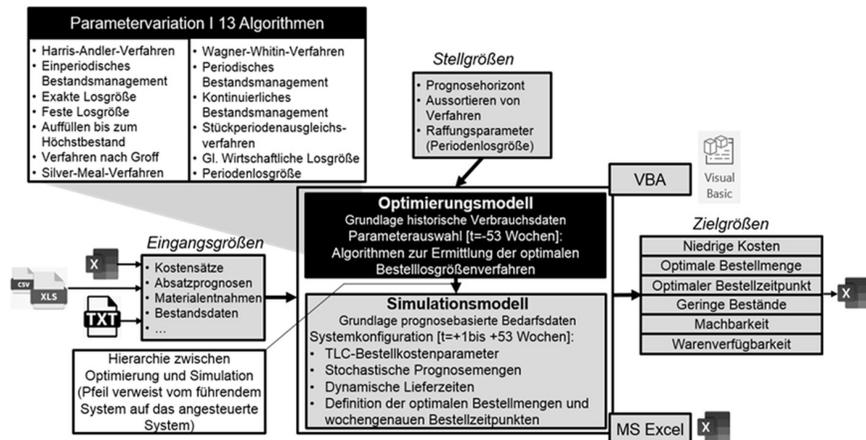


Abbildung 3: Simulationsbasierte Optimierung von Bestelllosgrößen in der verbrauchsgesteuerten Materialdisposition -eigene Darstellung in Anlehnung an: (März et al. 2011)

Als Input für das Simulationsmodell dienen Haupteingangsgroessen, wie Kostensätze, Absatzprognosen, Materialentnahmen und Bestandsdaten. Als Zielgrößen werden niedrige Kosten, optimale Bestellmenge und -zeitpunkt, geringe Bestände, Machbarkeit und Warenverfügbarkeit definiert.

5.3 Umsetzung und Ergebnisse

Die Anwendung der beschriebenen, simulationsbasierten Optimierung dispositiver Systemdynamiken, in ihrer Implementierung als MS-Excel-Tool, wurde für eine Fallstudie in der Investitionsgüterindustrie evaluiert. Mit einer Datenbasis von 4.066 Artikeln konnten für 1.733 optimierbare Artikel Kosteneinsparungen von über 55% erreicht werden. Für Artikel ohne ausreichende Datenbasis bzw. ohne zukünftige Bestellbedarf konnte in der Fallstudie keine simulationsbasierte Optimierung durchgeführt werden. Die Ergebnisse zeigen, dass ein optimiertes Verfahren der Bestelllosgrößenermittlung für jeden Zukaufartikel individuell ermittelt werden muss, um das Kostenoptimierungspotential auszuschöpfen. Durch die Anwendung der simulationsbasierten Optimierung dispositiver Systemdynamiken über alle Artikel der Fallstudie kann eine potenzielle Einsparung von 419.606€ für den Betrachtungszeitraum von einem Jahr generiert werden.

In der folgenden Tabelle 1 ist zu sehen, in welcher Häufigkeit die Verfahren für die untersuchten Artikel als optimaler Algorithmus ermittelt wurden. Zwar zeigen sich gewissen Losgrößenverfahren als häufig anwendbar, insgesamt ist aber zu erkennen, dass nicht ein Algorithmus für alle Artikel optimal ist, sondern jedes Material, je nach Bedarfsverlauf und Kostenstruktur, seinen jeweiligen besten Algorithmus bedingt.

Tabelle 1: Prozentuelle Verteilung der Losgrößenverfahren

Verfahren nach Groff	50,72%	Periodisches Bestands-management	0,98%
gleitende wirtschaftliche Losgröße	17,20%	Periodenlosgröße	0,69%
Anderl-Verfahren	11,71%	Kontinuierliches Bestands-management	0,35%
Wagner-Whitin Verfahren	8,31%	feste Losgröße	0,23%
Stückperiodenausgleichs-verfahren	4,10%	Einperiodisches Bestandsmanagement	0,00%
exakte Losgröße	4,04%	Auffüllen bis zum Höchstbestand	0,00%
Silver-Meal-Verfahren	1,67%		

Die zwei Verfahren *Einperiodisches Bestandsmanagement* sowie *Auffüllen bis zum Höchstbestand* ergaben keine Treffer, da ersteres nur auf verderbliche Artikel angewendet werden kann und zweiteres einen hinterlegten Höchstbestand erfordert, dessen Datenbasis beim angewendeten Unternehmen allerdings fehlte.

6 Optimierte Parametrisierung und prognosebasierte Simulation von Bestellmengen und -zeitpunkten

6.1 Verwandte Arbeiten

Larasati et al. verwenden zur Simulation die optimale Bestellmenge ermittelt mit dem EOQ-Modell (Larasati et al. 2020). In Rekek et al. (2017) wird ein Newsvendor Modell mit sich veränderlicher Nachfrage betrachtet, um die Optimierung der Bestelllosgröße unter Berücksichtigung von Judgement-Replenishment Entscheidungen zu berechnen. Es fehlen jeweils konkrete Anwendungen mit mehreren Algorithmen sowie die Parametrisierbarkeit der Lieferzeit.

6.2 Vorgehensweise und Modellbeschreibung

Auf Grundlage des identifizierten optimalen Bestelllosgrößenverfahrens je Artikel, der stochastischen Absatzprognosen der zukünftigen Bedarfe, der vorhandenen Bestände und der bereits getätigten offenen Bestellungen, wird eine Simulation für die zukünftigen Bestellungen ausgeführt, um die optimalen Bestellmengen und Bestellzeitpunkte je Artikel zu ermitteln. Die stochastischen Absatzprognosen werden als Input-Daten der Simulation zu Verfügung gestellt. Das Simulationsmodell wird parametrisiert, indem ein Bedarfszeitraum von +1 bis +53 Wochen gewählt und die Gesamtlieferzeit aus den hinterlegten Systemdaten oder vergangenen Bestellungen anhand eines statistischen Modells für die Simulation herangezogen werden. Dabei werden Materialdisponent/innen unterstützt, auf Articlebene kostenoptimale Bestellmengen und optimale Bestellzeitpunkte für die zukünftigen Bedarfe zu wählen.

6.3 Umsetzung und Ergebnisse

Die prognosebasierte Simulation von Bestelllosgrößen wurde wiederum, in ihrer Implementierung als MS-Excel-Tool, für eine Fallstudie in der Investitionsgüterindustrie evaluiert. Dabei wurden die Ergebnisse der simulationsbasierten Optimierung herangezogen und auf Basis der Absatzprognosen für maximal 53 zukünftigen Wochen die Losgrößen je nach berechnetem optimalem Verfahren je Artikel ermittelt. Dies wurde für die optimierbare Datenbasis von 1.733 Artikeln durchgeführt. Aufgrund der Parametrisierbarkeit der Lieferzeit (Lieferzeit aus dem ERP-System oder berechnete Lieferzeit) ergeben sich zwei mögliche Output-Dateien mit unterschiedlichen Bestelllosgrößen und -zeitpunkten. Damit lässt sich die prognosebasierte Simulation zum Einen mit den der Disposition zu Verfügung stehenden Lieferzeiten aus dem IT-System und zum Zweiten mit optimierten Lieferzeiten, welche aus der Belieferungshistorie berechnet werden, durchführen.

7 Schlussfolgerung und Ausblick

Das Ergebnis zeigt das hohe Nutzenpotenzial kostenoptimaler Anwendung von Bestelllosgrößenverfahren. Trotz einmalig erhöhtem Aufwand für die Umsetzung des Planungswerkzeugs und Einbindung in die ERP Systeme sowie des Aufwands für die (Daten-) Pflege des Werkzeugs überwiegt die potenzielle, jährliche Kostenersparnis diesen Aufwand deutlich.

Danksagung

Diese Forschung wurde von der Österreichischen Forschungsförderungsgesellschaft (FFG), Förderungsnummer 872787, finanziert. Die Autoren danken allen Fallstudienpartnern für ihren Beitrag.

Literatur

Ahmadi, T.; Mahootchi, M.; Ponnambalam, K.: Optimal randomized ordering policies for a capacitated two-echelon distribution inventory system. *Computers & Industrial Engineering* 124 (2018), S. 88–99.

- Alharkan, I.; Saleh, M.; Ghaleb, M.; Farhan, A.; Badwelan, A.: Simulation-Based Optimization of a Two-Echelon Continuous Review Inventory Model with Lot Size-Dependent Lead Time. *Processes* 8 (2020) 9, S. 1014.
- Beck, F.G.; Ries, J.M.; Urnauer, C.: Dynamische Losgrößenplanung in ERP-Systemen: Eine Performanceanalyse von Heuristiken der Fertigungs- und Bestellmengenplanung für unterschiedliche Bedarfsstrukturen. *Publications of Darmstadt Technical University, Institute for Business Studies (BWL)* (2015) 103748.
- Blecker, T.; Kersten, W.; Ringle, C.M.: *Proceedings of the Hamburg International Conference of Logistics (HICL)/Operational Excellence in Logistics and Supply Chains: Efficiency, Data-driven Approaches and Security-related Insights*. Berlin: epubli GmbH 2015.
- Bogaschewsky, R.; Müller, H.: BME-Barometer „Elektronische Beschaffung“. Würzburg, Leipzig: Bundesverband Materialwirtschaft, Einkauf und Logistik e.V.(BME) 2018.
- Brabänder, C.: *Stochastisches Bestandsmanagement: Grundmodelle für Betriebswirte*. Wiesbaden: Springer Fachmedien Wiesbaden GmbH 2020.
- Chang, H.-C.: A note on an economic lot size model for price-dependent demand under quantity and freight discounts. *International Journal of Production Economics* 144 (2013) 1, S. 175–179.
- Claus, T.; Herrmann, F.; Manitz, M.: *Produktionsplanung und –steuerung: Forschungsansätze, Methoden und deren Anwendungen*. Berlin, Heidelberg: Springer-Verlag 2015.
- Grabner, T.: *Operations Management*. Wiesbaden: Springer Fachmedien Wiesbaden 2019.
- Heiserich, O.-E.; Helbig, K.; Ullmann, W.: *Logistik: Eine praxisorientierte Einführung*. Wiesbaden: Gabler Verlag 2011.
- Kamhuber, F.; Sobottka, T.; Heinzl, B.; Sihm, W.: An Efficient Multi-Objective Hybrid Simheuristic Approach for Advanced Rolling Horizon Production Planning. *2019 Winter Simulation Conference (WSC)* (2019), S. 2108–2118.
- Kemmner, G.-A.; Sames, G., 2015: *DISPOSITION 4.0 FÜR DIE FABRIK 4.0*, Abels & Kemmner GmbH - Supply Chain Management Consultants, 2015. Online verfügbar unter <https://docplayer.org/78078351-Disposition-4-0-fuer-die-fabrik-4-0.html>, zuletzt geprüft am (13.05.2021).
- Larasati, R.H.; Khabibah, S.; Aziz, A.: An optimization model of economic order quantity with financial constraints and market tolerance in ud plastikq. *Journal of Physics: Conference Series* 1524 (2020), S. 12097.
- März, L.; Krug, W.; Rose, O.; Weigert, G.: *Simulation und Optimierung in Produktion und Logistik: Praxisorientierter Leitfaden mit Fallbeispielen*. Berlin, Heidelberg: Springer Berlin Heidelberg 2011.
- Müller, H., 2017: *Freies Online-Tool ermittelt Einsparpotenzial im indirekten Einkauf*, HTWK Leipzig. Online verfügbar unter https://www.htwk-leipzig.de/no_cache/forschen/aktuelles/detailansicht/artikel/787/, zuletzt geprüft am 13.05.2021.



- Prabhu, V.; Taisch, M.; Kiritsis, D.: *Advances in Production Management Systems. Sustainable Production and Service Supply Chains*. Berlin, Heidelberg: Springer Berlin Heidelberg 2013.
- Pumpe, A.; Vallée, F.: *Total Landed Cost for International Sourcing Decisions: Applying Importance-Performance Analysis to Evaluate Information Systems*. 2015 International Conference on Logistics, Informatics and Service Sciences (LISS), 2015.
- Rekik, Y.; Glock, C.H.; Syntetos, A.A.: *Enriching demand forecasts with managerial information to improve inventory replenishment decisions: Exploiting judgment and fostering learning*. *European Journal of Operational Research* 261 (2017) 1, S. 182–194.
- Robrade, A.D.: *Dynamische Einprodukt-Lagerhaltungsmodelle bei periodischer Bestandsüberwachung*. *Physica-Schriften zur Betriebswirtschaft*, 2013 (34).
- Schuh, G.; Schmidt, C.: *Produktionsmanagement: Handbuch Produktion und Management 5*. Berlin, Heidelberg: Springer-Verlag 2014.
- Singhal, J.; Singhal, K.: *On the noniterative multiproduct multiperiod production planning method*. *Operations Research Letters* 37 (2009) 1, S. 65–66.
- v. d. Heuvel, W.; Wagelmans, A.P.: *A holding cost bound for the economic lot-sizing problem with time-invariant cost parameters*. *Econometric Institute Report*, 2008.
- Wischmann, S.; Hartmann, E. (Hrsg.): *Zukunft der Arbeit: Eine praxisnahe Betrachtung*. Berlin, Germany: Springer Vieweg 2018.
- Zoller, K.; Robrade, A., 1987: *Dynamische Bestellmengen- und Losgrößenplanung: Verfahrensübersicht und Vergleich*. *OR Spektrum* 1987, 1987 (9), S. 219–233. Online verfügbar unter <https://doi.org/10.1007/BF01719833>.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Zuweisung bester Abstellplätze im LIFO-Lager unter Berücksichtigung der Kommissionierungsreihenfolge mittels genetischem Algorithmus

***Allocation of best storage locations in the LIFO warehouse taking into
account the picking sequence using genetic algorithm***

Dominik Kuhn, Jan Adelsbach, Attique Bashir, Leenhard Hörauf, Rainer Müller,
ZeMA gGmbH, Saarbrücken (Germany), d.kuhn@zema.de, j.adelsbach@zema.de,
a.bashir@zema.de, l.hoerauf@zema.de, rainer.mueller@zema.de

Abstract: We present a practical approach for storage optimization of a LIFO (Last-In-First-Out) topology warehouse using complementary brute-force and genetic algorithms in a simulated environment using real-time data from the real-life counterpart in order to assist personal in allocation planning. Both the optimizations for placement and removal of storage units are being discussed.

1 Einleitung, Motivation und Ausgangssituation

Bodenlager dienen in der Intralogistik als Puffer zwischen beispielsweise einer Produktion und dem Warenausgang. Je nach Layout des Lagers und örtlichen Gegebenheiten unterliegen diese spezifischen Einlagerungs- und Auslagerungsprinzipien, wie bspw. FIFO (First-In-First-Out) oder LIFO (Last-In-First-Out). Je nach Produktionsmenge und verfügbarer Lagerkapazität kann eine manuell geplante Einlagerung mit Hinblick auf die Kommissionierung suboptimal sein. Bemerkbar macht sich dieses Problem, wenn eine Zugänglichkeit zu einem zu kommissionierenden Produkt (Palette) nicht ohne weiteres ermöglicht wird.

Getränkehersteller produzieren ihre Produkte in der Regel mit einer bestimmten Vorlaufzeit und organisieren ihre Lagereinheiten in Paletten. Dabei wird eine Lagereinheit an dem Zeitpunkt der Herstellung nicht zwangsläufig mit einem Kommissionsauftrag in Verbindung gebracht (Keine Auftragspezifische Produktion). Bei einer zufälligen (ungünstigen) Positionierung von Paletten können bei der Kommissionierung Aufträge einander den Zugang blockieren, wenn bspw. ein später platzierter Auftrag später kommissioniert werden soll. Daher werden Abstellplätze durch einen Mitarbeiter der Lagerlogistik so ausgewählt, dass Paletten sich entsprechend der geplanten Kommissionssequenz nicht blockieren. Gerade bei hohen Produktionsmengen ist die Erstellung eines Belegungsplans unüberschaubar

und für den Lagerleiter nicht mehr zuverlässig zu erstellen. Dies führt zu vermehrten Blockaden im Lager. Folglich ist der Lagerist mit Umplatzierungsaufwänden konfrontiert, sodass eine weitere unplanmäßige Umplatzierung zu weiterer Eskalation führt.

Um die Umplatzierungsaufwände für den Lageristen zu verringern, wird im vorliegenden Paper eine Methode beschrieben, die bei bekannter Kommissionierungsreihenfolge einer oder mehreren fertiggestellten Paletten einen Platz zuweist. Zur Anwendung kommt hierbei als Metaheuristik ein genetischer Algorithmus, der zwar nicht die optimale Lösung garantiert, aber meistens zu einer annehmbaren Lösung kommt, da dieser die Lösung gezielt hinsichtlich einer Bewertungsfunktion zu minimieren versucht.

2 Stand der Technik

Die Zuweisung bester Abstellplätze ist ein bekanntes Problem, dem unter dem Begriff „Storage Location Assignment Problem - SLAP“ nachgegangen wird. (Kofler et al., 2010; Battista et al., 2011) Gängige Methoden versuchen das Problem mathematisch zu beschreiben und mittels Linearer Programmierung (LP) zu lösen. Weitere Ansätze nutzen Metaheuristiken, wie die genetischen Algorithmen oder Simulated Annealing, um beste Lagerplätze zu ermitteln. Die Anwendung des Kuhn-Munkres Algorithmus ist ebenfalls eine Optimierungsmethode zur Zuweisung von Produkt oder Bauteil auf Lagerplätze. Viele Arbeiten berücksichtigen jedoch nicht explizit die Restriktionen beim Einlagern in Lager mit einem LIFO Betriebskonzept. Beim LIFO (Last-In-First-Out) Prinzip können Paletten nur in der umgekehrten Einlagerreihenfolge wieder ausgelagert werden.

3 Lagerbeschreibung

Der Typ des Fertiglagers im betrachteten Fall (siehe **Fehler! Verweisquelle konnte nicht gefunden werden.**) wird als Bodenlager in Blockform bezeichnet. Lagereinheiten, sind in diesem Fall Paletten, die direkt und lückenlos nebeneinander, hintereinander und paarweise übereinandergestapelt werden. Der Zugriff auf die jeweiligen Lagereinheiten folgt dem LIFO-Prinzip, sodass einzelne Reihen (Spalten) eines Blocks nur in einer Richtung befüllt oder geleert werden können. Reihenübergreifend lässt das Lager beliebige Abstellplätze zu. Einzelne Reihen oder Blöcke sind in der Regel rechtwinklig oder auch schräg zur eigentlichen Fahrgasse angeordnet.

Bei dem Kommissionierungslager handelt es sich ebenfalls um ein Bodenlager in Blockform. Hier werden Lagereinheiten auftragsbezogen für die jeweiligen Kunden zusammengestellt. Der Lagertyp des Kommissionierungslagers orientiert sich in Grundzügen an dem FIFO-Prinzip. Aus dem Fertiglager entnommene Lagereinheiten werden ausschließlich in einer Richtung in den jeweiligen Reihen gestapelt, wobei die Beladung der LKWs aus genau entgegengesetzter Richtung erfolgt.

Softwaretechnisch lassen sich die beiden Lager tabellenförmig mittels Matrizen modellieren. Die jeweiligen Einträge spiegeln die reale Belegung der jeweiligen Lagerplätze wieder, ein Nulleintrag steht für einen unbelegten Platz. Die Kopplung der Einträge mit den dazugehörigen Varianten sowie Länderkennung stellt ein

notwendiges eindeutiges Unterscheidungsmerkmal dar, welches bei der Bestimmung der Lagerplätze von großer Bedeutung ist. Im konkreten Anwendungsfall wird jeder Eintrag durch einer sogenannten Paletten-ID beschrieben. Analog zu dem realen Fertiglager werden die Reihen der Matrix durch eingehende Aufträge von vorne befüllt. Lagereinheiten, die zur Kommission bereitgestellt werden, werden durch eine Entnahme zuerst mit einem Nulleintrag von vorne in den Spalten versehen. Ähnlich verhält es sich mit der softwaretechnischen Beschreibung des Kommissionierungslagers. Die Modellierung des Kommissionierungslagers spielt im konkreten Optimierungsproblem eine untergeordnete Rolle, aus diesem Grund wird in diesem Paper auf diesen Lagerteil im Folgenden nicht näher eingegangen.

4 Problemformulierung

Sei R ein Lager mit N heterogenen Reihen $R = \langle R_1, R_2, \dots, R_n \rangle$ bestehend aus Lagereinheiten. Letztere sind definiert als $p \in \mathbb{N} \cup \{\emptyset\}$, welches entweder zu einer Paletten-ID oder \emptyset als nicht-besetzt entspricht. Der Produkttyp einer Palette $p \neq \emptyset$ ist definiert als $P_t(p)$, das Herstellungsdatum als $P_d(p)$. Als Kurzhandform wird für eine Lagereinheit in Reihe n und Tiefenposition m auch $p_{n,m}$ verwendet.

Bei der Einlageroptimierung sollen zu allen Paletten in einer gegebenen Liste E die bestmöglichen Positionen nach einem Bewertungsschema gefunden werden. Zur Bewertung der Lösung wird bei der Einlageroptimierung der kumulativ höchste normierte Wert für das gesamte Lager $s = \frac{1}{|R|} \sum_i^N s_i$ aus der Summe von mehreren im späteren beschriebenen normierten Bewertungsverfahren B für jede Reihe $s_n = \sum_{b \in B} s_{b,n}$ gesucht.

Bei der Auslageroptimierung sollen für eine Kommissionierungssequenz O passende Palettenpositionen im Lager gefunden werden. Hierbei wird versucht die Anzahl von Umlagerungen für das gesamte Lager zu minimieren.

Die Optimierung soll anhand des tatsächlichen Lagerbestandes simulieren, wie die aktuell produzierten Paletten am besten eingelagert werden können bzw. welche Paletten des Lagers ausgelagert werden sollen. Das System soll dabei den Lagerplaner in seiner Arbeit unterstützen. Die Optimierung wird dabei über eine Schnittstelle aus dem proprietären ERP-System des Getränkeherstellers heraus aufgerufen. Als Laufzeitumgebung wurde die Python Programmiersprache gewählt und zur Umsetzung der Algorithmik genutzt.

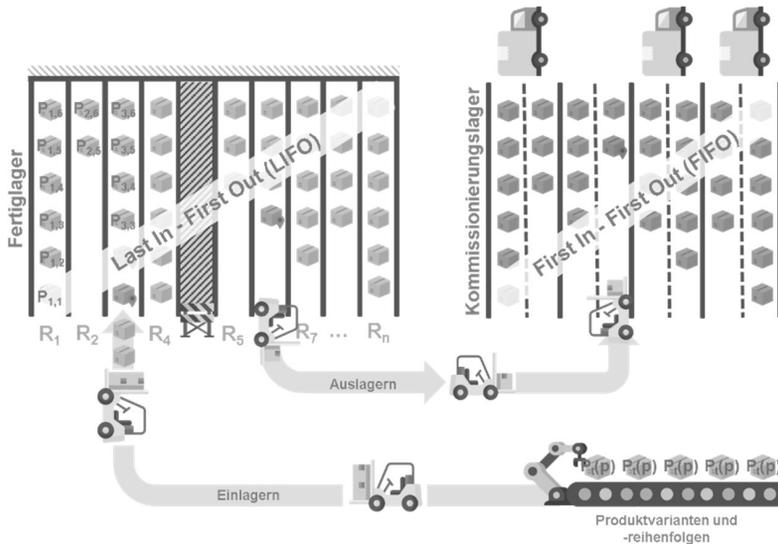


Abbildung 1: Lagermodell und Problembeschreibung

5 Genetischer Algorithmus

5.1 Biologisches Vorbild sowie Grundidee

Der genetische Algorithmus stellt eine Metaheuristik aus der Klasse der evolutionären Optimierungsverfahren dar. Dabei orientiert sich dieser Algorithmus an der Evolutionstheorie von Charles R. Darwin. Nach Darwin erfolgt im Zuge der biologischen Evolution eine natürliche Selektion von Individuen mit Eigenschaften, die diesem eine höhere Überlebenschance bieten. Diese Art von Optimierung wurde ursprünglich unter anderem in Turing (1950) beschrieben.

Analog dazu werden bei genetischen Algorithmen Populationen von Individuen betrachtet und somit eine Vielzahl von möglichen Lösungen in einem gegebenen Suchraum. Diese Individuen werden in Anlehnung zur Genetik häufig auch als Chromosomen bezeichnet. Eine Stelle oder Sequenz eines Chromosoms, also eine Entscheidungsvariable des betrachteten Problems, wird wiederum als Genom bezeichnet. Für die Charakterisierung der Lösungsgüte bezüglich der betrachteten Problemstellung wird die sogenannte Fitnessfunktion betrachtet. Diese beschreibt die Güte eines Individuums mit einem reellen Wert.

Entsprechend der biologischen Mechanismen der Kreuzung, Mutation und der Selektion wird auch bei den genetischen Algorithmen verfahren, um eine Weiterentwicklung der Population der Individuen zu ermöglichen.

5.2 Aufbau des genetischen Algorithmus

Der moderne genetische Algorithmus entspricht den Konzepten, die von Bremermann beschrieben wurden, siehe Anderson und Conrad (1995) und besteht grundlegend aus den folgenden Phasen, die dieser teilweise iterativ durchläuft:

1. Initialisierung einer zufälligen Ausgangspopulation
2. Errechnen der Fitness für jedes Individuum
3. Selektion
4. Rekombination/ Crossover
5. Mutation
6. Wiederholung ab Schritt 2, bis ein Abbruchkriterium erfüllt ist.

Die initiale Kodierung der Individuen beim genetischen Algorithmus wird häufig in Form von Permutationen vorgenommen. Beispielsweise bei der Ermittlung eines optimalen Weges, indem verschiedenen Wegpermutationen, über einer Fitnessfunktion, Zielkriterien, wie zum Beispiel der Zeit, zugeordnet werden. Für eine zufällig initialisierte Startpopulation wird anschließend über eine geeignete Fitnessfunktion die Güte der einzelnen Lösungen bestimmt. Die Selektion dient dazu, aus der Population die vermeintlich fittesten Individuen zur Übernahme in die Folgepopulation auszuwählen. Dabei gibt es verschiedene Methoden. Der Großteil der Methoden haben gemein, dass sie bei der Wahl der fittesten Individuen zusätzlich eine stochastische Komponente berücksichtigen. Die bekanntesten Selektionsmethoden sind (Mehmeti und Amrein et al., 2014):

- Monte-Carlo Selektion (Roulette-Wheel-Selektion)
- Rangbasierte Selektion
- Tournament-Selektion

Eine weit verbreitete Methode ist die Tournament-Selektion. Dabei werden zufällige Individuen ausgewählt und jenes mit dem höchsten Fitness-Wert von diesen selektiert. Dabei können die Selektionen auch Elitär verfahren, wobei nur die Individuen mit dem höchsten Fitness-Wert einbezogen werden. Aus zwei selektierten Individuen werden bei der Rekombination durch das sogenannte Crossover neue Individuen generiert. Aus den neu entstandenen Individuen werden mit einer geringen Wahrscheinlichkeit Kandidaten für Mutationen ausgewählt, um die Exploration des Suchraums zu stärken. Bei der Mutation wird ein zufälliges Allel geändert. (Mehmeti und Amrein et al., 2014) Für die Zusammensetzung der neuen Generation werden wieder insgesamt N Individuen aus der alten Generation und den neuen Individuen ausgewählt.

6 Umsetzung der Algorithmik

6.1 Einlageroptimierung

6.1.1 Bewertungsverfahren

Für die Einlageroptimierung werden mehrere Bewertungsverfahren $s_{t,n}$, $s_{l,n}$, $s_{d,n}$, $s_{a,n}$ und $s_{u,n}$ für eine Reihe n (normierte Werte $[0,1]$) verwendet, welche folgend im Detail abgehandelt werden, diese werden anhand von definierbaren Koeffizienten c_k summiert

$$s_n = \frac{1}{5} (c_t s_{t,n} + c_l s_{l,n} + c_d s_{d,n} + c_a s_{a,n} + c_u s_{u,n}) \quad (1)$$

Die kumulative Bewertung für das gesamte Lager ist ein wie folgt abgeleiteter Normierter Wert $s = \frac{1}{N} \sum_i^N s_i$ dieser wird zu maximieren versucht.

Das Typen-Clustering leitet die Typenreinheit anhand der Anzahl von Paletten k ab, die den gleichen Typ haben und zu Beginn der Reihe stehen, sowie die Anzahl der restlichen Paletten in der Reihe l :

$$s_{t,n} = \frac{\max(0, k - 1) - l}{|R_n|} \quad (2)$$

Gegeben das arithmetische Mittel der Produktionsdaten aus allen besetzten Lagerplätzen in einer Reihe, $\overline{P_{d,n}} = \min(V_n) + \frac{1}{2}[\min(V_n) + \max(V_n)]$ für $V_n = \{P_d(p) : p \in R_n \wedge p \neq \emptyset\}$, sowie das arithmetische Mittel über das gesamte Lager $\overline{P_d} = \frac{1}{N} \sum_i^N \overline{P_{d,i}}$ wird der Produktionsdaten-Clustering Bewertung als:

$$s_{d,n} = \max\left(0, \frac{\overline{P_{d,n}} - \overline{P_d}}{\overline{P_{d,n}}}\right) \quad (3)$$

errechnet. Beim Leer-Clustering wird errechnet, wie viele Paletten fehlen, um eine gegebene Belagerung c_u zu erreichen. Hierbei wird gezählt, wie viele Paletten in der Reihe frei sind in Relation zu der Anzahl von Lagerplätzen in der Reihe:

$$s_{l,n} = |\{z \in R_n \mid z = \emptyset\}| / |R_n| - c_u \quad (4)$$

Der Belagerungskoeffizient kann entweder manuell definiert werden, oder für das gesamte Lager kann durch das arithmetische Mittel der einzelnen Reihenbelegungen errechnet werden:

$$c_u = \frac{1}{N} \sum_i^N s_{l,i} \Big|_{c_u=0} \quad (5)$$

Beim Auftrags-Clustering wird anhand einer Kommissionssequenz O gezählt, wie viele Paletten in der aktuellen Reihe auf diese passen, bzw. nicht passen. Die Reihenfolge der Paletten in der Kommission spielt hierbei keine Rolle, da es in den gegeben Einsatzszenario irrelevant ist, in welcher Reihenfolge diese auf Kommissionsplätze verschoben werden.

$$s_{a,n} = |\{z \in R_n \mid z \in O\}| / |R_n| \quad (6)$$

Hierbei wird jedes Element in O nur maximal einmal gezählt. Bei der Umlagerbewertung wird, gegeben eine Liste an auszulagernden Paletten $a_{n,k} \in A$ errechnet, wie viele Umlagerungen benötigt werden, um jede von diesen Paletten zu erreichen.

$$s_{u,n} = 1 - \max|\{z \mid z_m \in R_n \wedge m < k\}| / |R_n| \quad (7)$$

Hierbei werden, wenn Paletten des gleichen Typs an einer vorherigen Position in der Reihe auftreten jedoch nicht zum Auslagern markiert sind, diese anstatt der eigentlich markierten Paletten in A zum Errechnen dieser Bewertung genommen.

6.1.2 Brute-Force Methode

Bei $|E| = 1$ wird eine Brute-Force Methode genutzt. Dabei werden die beschriebenen Bewertungen für alle Reihen errechnet, und die Reihe mit der höchsten Bewertung herausgesucht. Sollten mehrere diesen teilen, so wird eine zufällige dieser ausgewählt.

6.1.3 Genetischer Algorithmus

Für das Einlagern von mehreren Paletten $|E| > 1$ hingegen stellt sich das Einlagern als Kombinatorisches Problem mit einer Anzahl von Möglichkeiten nach dem Binomialkoeffizienten da. Dieses Problem wird mit einem genetischen Algorithmus gelöst.

Zuerst werden alle freien Lagerplätze im Lager in einen Vektor T herausgesucht, dieser enthält die Indices von Reihen mit freien Lagerplätzen. Dabei repräsentiert das mehrfache Auftreten eines Index eine entsprechende Anzahl von freien Lagerpositionen in dieser Reihe, diese sind jedoch als Optimierung auf die Anzahl von einzulagernden Paletten limitiert $\min(|E|, |\{z \in R_n \mid z = \emptyset\}|)$, siehe **Fehler! Verweisquelle konnte nicht gefunden werden..**

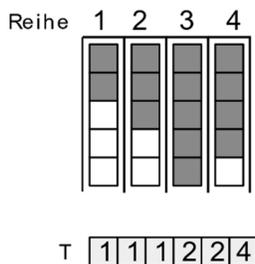


Abbildung 2: Generation des Positionsvektors T

Der Individuums-Vektor I welcher dieselbe Anzahl von Elementen wie T enthält und mit diesem korrespondiert, enthält die Paletten-IDs der einzulagernden Paletten oder \emptyset . Hierbei tritt jede Paletten-ID nur einmal auf. Zum Rückwärtsauflösen wird der Individuen-Vektors I sequentiell durchlaufen, leere Elemente übersprungen und die entsprechenden Paletten in die Reihe entsprechen des Indexes im Vektor T eingelagert, siehe **Fehler! Verweisquelle konnte nicht gefunden werden..**

Für das Initialisieren, werden in einem leeren Vektor entsprechend der Anzahl von einzulagernden Paletten zufällige Positionen gewählt und die Paletten diesen zugeordnet. Dabei kann optional eine strikte Reihenfolge der Paletten eingehalten werden. Beim Mutieren wird zufällig eine von drei Methoden, Twors, RSM (Reverse-Sequence-Mutation) und CIM (Center-Inversion-Mutation) wie beschrieben unter anderem in Abdoun et al. (2012) mit einer gleichen Wahrscheinlichkeit gewählt. Bei der Twors Mutation werden zufällige Positionen vertauscht. Bei der RSM, entsprechend Abdoun et al. (2012) werden zwei Punkte $1 \leq a \leq b \leq |I|$ per Zufall gewählt und alle Elemente zwischen diesen in der Reihenfolge umgekehrt. Bei der CIM wird ein Punkt $1 \leq a \leq |I|$ zufällig gewählt und die Subsequenzen davor und danach in der Reihenfolge umgekehrt.

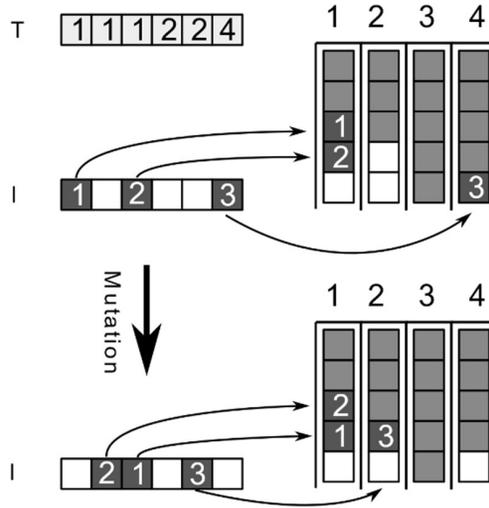


Abbildung 3: Rückwärtsauflösung der Vektoren I und T

Die Rekombination wird durch einen Ordered Crossover (OX) nach Abdoun und Abouchabaka (2012) durchgeführt. Um eine Standard-Implementierung der OX Rekombination zu ermöglichen, werden alle nicht gesetzten Werte der beiden Eltern-Individuen durch absteigende negative Integer ersetzt, sodass alle Elemente distinktive Werte beinhalten. Letzteres setzt voraus, dass es keine negativen Paletten-IDs gibt. Zur Selektion wird eine Tournament Selektion wie in Goldberg und Deb sowie Blickle und Thiele (1996) benutzt. Der Fitnesswert wird berechnet, in dem die Paletten des Individuums-Vektor auf das Lager temporär angewendet werden und die bereits beschriebene Bewertung für das gesamte Lager s errechnet wird. Hierbei werden die Bewertungen für einzelne Reihen neu errechnet.

6.2 Auslageroptimierung

Bei der Auslageroptimierung wird jeder Eintrag o_k auf einer gegebenen Kommissionierungssequenz $O = \langle o_1, o_2, \dots, o_k \rangle$ mit der Position einer Palette des entsprechenden Typs im Lager zu befüllt. Hierbei gilt es die Positionen der Paletten so zu wählen, dass diese mit möglichst wenig Umlagerungen aufgegriffen werden können. Hierbei kann zwischen „direkt“ und „indirekt“ freien Paletten, sowie „blockierten“ Paletten unterschieden werden: Direkte Paletten stehen in erster Position einer Reihe. Indirekte Paletten werden durch andere Paletten, die ebenfalls in der Kommissionierungssequenz zur Auslagerung zugeordnet worden freigelegt, $p_{n,m}$ wenn $\{p_{n,1}, p_{n,2}, \dots, p_{n,m-1}\} \subset O$. Blockierte Paletten werden durch andere Paletten blockiert, die zwingend umgelagert werden müssen, $\exists p \notin O$ für $p \in \{p_{n,1}, p_{n,2}, \dots, p_{n,m-1}\}$.

6.2.1 Brute-Force Methode

Bei der Brute-Force Methode wird die gegebene Kommissionierungssequenz sequentiell durchlaufen und es wird versucht für jeden geforderten Palettentyp eine direkt- oder indirekt freistehende Palette des Lagers zu finden. Um indirekt

freistehende Paletten zu finden, werden bereits zugewiesene Paletten aus dem Lager entfernt. Paletten, die nicht auf diese Weise zugewiesen werden können $O_f \subseteq O$ und dadurch blockiert sind werden daraufhin mit dem im Folgenden beschriebenen genetischen Algorithmus ausgewählt bzw. optimiert.

6.2.2 Genetischer Algorithmus

Sei O_f das Subset an nicht direkt- oder indirekt freistehende Paletten und gegeben das Lager ohne die bereits durch den Brute-Force Algorithmus zugewiesenen Paletten. Hierbei wird versucht die Anzahl von Umlagerungen s_v zu minimieren. Diese ist hierbei die Summe der Maxima der Tiefenpositionen von den zu auslagernden Paletten im Individuums-Vektor I definiert.

$$s_{v,n} = \sum_i^N \max\{m : p_m \in I \wedge p_m \in R_i\} - 1 \quad (8)$$

Bei der Mutation wird ein zufälliges Element des Individuums-Vektors $x \in I$ gewählt. Danach wird aus einer Liste von Paletten selbigen Typs, die sich im Lager befinden, jedoch nicht bereits an einer anderen Stelle im Individuen-Vektors zugewiesen sind $y \in \{p \in P \mid P_t(p) = P_t(x) \wedge \nexists z \in I\}$ eine zufällige gewählt und die Position von dieser mit x vertauscht. Sollten sich keine weiteren Paletten eines Typs im Lager befinden, so wird sofern vorhanden, ein weiteres Element des Individuum-Vektors vom gleichen Typ ausgewählt, $y \in \{p \in I : P_t(p) = P_t(x)\}$ und das Element x mit diesem vertauscht. Für die Rekombination werden die beiden Eltern-Vektoren I_1, I_2 in Sub-Vektoren nach Palettentyp t partitioniert $I_{k,t} = \{z \mid z \in I_k \wedge z \in P_t\} \forall t$. Dies ist nötig, da nur Paletten des gleichen Typs rekombinierbar sind. Die beiden Sub-Partitionen jedes Typs der Eltern $I_{1,t}, I_{2,t}$ werden daraufhin mit einem Cyclic Crossover (CX) wie beschrieben in Abdoun und Abouchabaka (2012) rekombiniert. Im Anschluss werden die rekombinierten Sub-Partitionen, entsprechend der originalen Indexe zurück zu zwei Kinder-Vektoren zusammengefügt. Für den Fitness-Wert, wird die oben beschriebene Kostenfunktion $s_{v,n}$ für das Lager entsprechend der Paletten im Individuums-Vektor errechnet.

7 Diskussion und Ausblick

Die vorgestellte Einlageroptimierung sucht eine ideale Position von einzulagernden Paletten anhand des Ist-Zustandes des Lagers und der Kommissionsequenz zu dem Zeitpunkt der Optimierung. Dabei werden in diese Optimierungsstrategie lediglich die Anzahl der Paletten einbezogen, welche sich im Puffer zwischen Abfüllung und Lager befinden. Zukünftige Herstellungs- und Kommissionsequenzen können durch diese Methode nicht berücksichtigt werden, was eine gewisse Limitierung des Optimierungspotentials darstellt. Der Einsatz von probabilistischen Modellen könnten an dieser Stelle das gegebene Potential um eine gewisse Vorausschau erweitern. Die Verwendung von reihenbasierenden Bewertungskriterien ist limitiert in der Anzahl von Metriken, die abgeleitet werden können. Mit Hilfe von Machine Learning (ML) Algorithmen könnte zudem der Lagerbestand in einem definierten zukünftigen Zeitraum im Voraus optimiert werden.

Die Auslagermethode ist limitiert, da sie nur versucht definierte Aufträge in einer definierten Reihenfolge zu vollfüllen. Ein weiterer Ansatz hierbei könnte sein die Kommissionierungssequenz selber anhand von dem Ist-Zustandes des Lagers, der Verfügbarkeit von Kommissionierungsplätzen und anhand von den Auslieferzeiten zu optimieren. Dabei könnten Kommissionierungen, die zeitlich verschiebbar sind so umgeplant werden, dass es fast nie zu Umlagerungen kommt. Der Ansatz, zuerst die Paletten anhand von einer Brute-Force Methode zuzuweisen und nur die nicht durch diese Methode zuweisbaren Paletten durch einen GA zu optimieren ist aufgrund von Zykluszeiten getroffen worden. Dies hat jedoch den trade-off, dass der GA nur die bestmögliche Lösung relativ zu dem Ergebnis der Brute-Force Methode findet. Ein anderer Ansatz hierbei wäre es die gesamte Auslageroptimierung durch den GA zu optimieren.

Aktuell befindet sich der beschriebene Ansatz in der Validierung im laufenden Produktivbetrieb. Hierbei wird observiert ob dieser Vorteile gegenüber der manuellen Planung hervorbringt. Offline Simulationen mit ca. 1/5 so großen Lagerdaten, wie im konkreten Anwendungsfall, zeigten deutlich die Funktionsweise dieser Methode in der Einlagerungsoptimierung.

Literatur

- Abdoun, O.; Abouchabaka, J.: A Comparative Study of Adaptive Crossover Operators for Genetic Algorithms to Resolve the Traveling Salesman Problem (2012).
- Abdoun, O.; Abouchabaka, J.; Tajani, C.: Analyzing the Performance of Mutation Operators to Solve the Travelling Salesman Problem. IJES (2012).
- Anderson, R.W.; Conrad, M.: Hans J. Bremermann: A pioneer in mathematical biology. *Biosystems* 34 (1995) 1-3, S. 1–10.
- Battista, C.; Fumi, A.; Giordano, F.; Schiraldi, M.: Storage Location Assignment Problem: implementation in a warehouse design optimization tool. In: Breaking down the barriers between research and industry, Padua (Italy), 14-16.09.2011, 2011,
- Blickle, T.; Thiele, L.: A Comparison of Selection Schemes Used in Evolutionary Algorithms. *Evolutionary Computation* 4 (1996) 4, S. 361–394.
- Goldberg, D.E.; Deb, K.: A Comparative Analysis of Selection Schemes Used in Genetic Algorithms. In: *FOGA 1990*, S. 69–93.
- Kofler, M.; Beham, A.; Wagner, S.; Affenzeller, M.; Reitingner, C.: Reassigning Storage Locations in a Warehouse to Optimize the Order Picking Process. In: 22nd European Modeling and Simulation Symposium EMSS 2010, Morocco, 13.10.2010, 2010,
- Mehmeti, Q.; Amrein, V.; Kulms, M.; Macin, N.; Loenser, C.; Bogonos, D.; Waidhas, M.: *Genetische Algorithmen: Wirtschaftsmathematisches Projekt zur Numerik*. Hg. v. Technische Universität Dortmund, Fakultät für Mathematik; Angewandte Mathematik und Numerik Dortmund, 2014,
- Turing, A.M.: I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind* LIX (1950) 236, S. 433–460.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Interoperabilität realer und simulierter Produktionssysteme mittels OPC UA

Interoperability of real and simulated production systems using OPC UA

Elisabeth Schmidl, Matthias Wenk, OTH Amberg-Weiden, Amberg (Germany),
e.schmidl@oth-aw.de, m.wenk@oth-aw.de

Eva Russwurm, Jörg Franke, FAU Erlangen-Nürnberg, Erlangen (Germany),
eva.russwurm@faps.fau.de, jörg.franke@faps.fau.de

Abstract: Simulation-based tools support developers in the engineering process, as they allow decision-relevant investigations to be carried out simulatively in advance. Some of these simulation studies require models that are as close to reality as possible. For their creation, data from real plants are necessary, which is associated with an increased manual acquisition and implementation effort. Furthermore, the different available simulation tools are suitable for different levels of detail. In order to use the synergies of the different tools, they have to be coupled, for which there are often no standardized interfaces yet. This paper presents a concept based on the OPC UA communication standard, which enables automated data transfer from real to simulated production plants, as well as flexible data exchange between different simulation tools.

1 Motivation

Die zunehmende Individualisierung von Produkten und kürzere Innovationszeiten erfordern neue leistungsfähigere Produktionssysteme. Diese müssen über eine wesentlich höhere Flexibilität und Wandlungsfähigkeit (Nyhuis et al. 2009) verfügen, bei gleichbleibend hoher Produktqualität und Anlagenverfügbarkeit, was zu einer deutlichen Zunahme der Komplexität im Engineeringprozess führt. Um diesen auch in Zukunft beherrschen zu können, wird der Einsatz simulationsbasierter Hilfsmittel vorangetrieben. Die Simulation ermöglicht entscheidungsrelevante Untersuchungen bereits vor der Realisierung der Anlage. Es können beispielsweise Planungsvarianten bewertet, Risikofaktoren identifiziert und die Anlage virtuell in Betrieb genommen werden (Wenk 2014). Außerdem können auf Basis der Simulation mittels moderner maschineller Lernverfahren verschiedene Bereiche, wie der Materialfluss (Jaensch et al. 2018), der Energiebedarf (Schmidl et al. 2021) oder die Fehlererkennung (Sobie et al. 2018) optimiert werden.

Voraussetzung ist die Erstellung geeigneter Simulationsmodelle, die das Verhalten der Anlage möglichst realitätsnah abbilden. Fertige Bibliothekselemente für unterschiedliche Komponenten unterstützen bei der Erstellung der Simulationen. Allerdings beinhalten diese Bibliothekselemente oft kein detaillierteres Verhalten, welches für Anlagenoptimierungen nötig wäre. Auch die Werte aus den Datenblättern der betreffenden Komponenten sind hierfür meist zu ungenau und lückenhaft. Somit müssen aufwändige Messungen an der realen Anlage erfolgen, welche dann häufig manuell in das Simulationsmodell integriert werden.

Je nach benötigtem Detaillierungsgrad existieren für die Simulation von Produktionssystemen entsprechende Simulationstools. Für Optimierungsaufgaben ist jedoch sowohl eine hohe Realitätsnähe als auch das Zusammenwirken mehrerer Produktionsanlagen entscheidend, was zu einer immensen Erhöhung des Rechenaufwandes führt. Deshalb ist es notwendig, detaillierte Komponenten- und Produktionsanlagensimulationen mit gröberen Fabriksimulationen zu koppeln, welche auf unterschiedlichen Rechnern ablaufen. Auch in Abhängigkeit der spezifischen Eigenschaften des zu simulierenden Verhaltens gibt es unterschiedliche zu bevorzugende Simulationssysteme. In dem Beitrag von Tur et al. (2019) wird beispielsweise gezeigt, dass Energieflüsse zeitkontinuierlich simuliert werden sollten, damit ein geringerer Informationsverlust entsteht. Da Produktionsprozesse hinsichtlich des Materialflusses überwiegend zeitdiskret simuliert werden, ist für eine ganzheitliche Betrachtung die Kopplung der unterschiedlichen Systeme notwendig. Da verschiedene Simulationstools oft nicht miteinander kompatibel sind, müssen hierfür geeignete Schnittstellen und Datenaustauschkonzepte entwickelt werden.

In diesem Artikel soll auf Basis von OPC UA ein Ansatz vorgestellt werden, welcher es ermöglicht, die Interoperabilität zwischen realen und simulierten Produktionssystemen zu erhöhen. Dadurch wird die automatisierte Optimierung von Simulationsmodellen sowie die flexible Kopplung von Simulationsmodellen mit unterschiedlichem Detaillierungsgrad erleichtert.

2 Stand der Technik

Es gibt bereits verschiedene Ansätze, das Verhalten von Simulationskomponenten realitätsnah abzubilden. In dem Beitrag von Seow und Rahimifard (2011) wird der Energiebedarf ermittelt und simuliert, indem Daten von Geräteherstellern mit eigenen empirischen Messstudien kombiniert werden. Kohl et al. (2014) zeichnen die Lastprofile während der verschiedenen Betriebszustände auf und stellen die Ergebnisse mittels Datentabellen zur Verfügung. Allerdings werden in beiden Fällen einige der Schritte zur Datenerhebung und der darauf aufbauenden Erstellung eines realitätsnahen Simulationsmodells manuell durchgeführt.

Das Gebiet der Co-Simulation beschäftigt sich damit, unterschiedliche Simulationswerkzeuge miteinander zu koppeln. Einen wichtigen Ansatz stellt hier das Functional Mock-up Interface (FMI) dar, welches eine standardisierte Schnittstelle zum Austausch dynamischer Modelle liefert (Blochwitz et al. 2012). Allerdings muss bereits zu Beginn der Modellierung entschieden werden, welche Werkzeuge als Functional Mock-up Unit (FMU) exportiert werden und welche Werkzeuge wiederum diese FMU's importieren, um miteinander zu interagieren, was zu einer geringeren Flexibilität während der Simulation führt. Außerdem unterstützen einige Simulationstools nicht oder nur teilweise den FMI-Standard (fmi-standard 2021).

Darüber hinaus ist die High-Level Architecture (HLA), eine vom US-amerikanischen Verteidigungsministerium definierte Architektur, zu nennen (IEEE Standards Association 2010). HLA ist bereits 2000 zu einem internationalen Standard geworden und wurde entwickelt, um die verteilte Modellierung und Simulation zu ermöglichen. Straßburger (2019) zeigt in seinem Beitrag die Vorteile von HLA auf, wie die dedizierte Synchronisationsunterstützung für Simulationssysteme und die ausgefeilte Datenverteilung. Allerdings wird hier auch die mangelnde Akzeptanz in der Fertigungsindustrie erwähnt, weshalb der Standard von kaum einem in diesem Bereich verfügbarem Simulationstool unterstützt wird.

Eine weitere Möglichkeit zur Kopplung von Simulationstools bietet die plattformunabhängige und service-orientierte Open Platform Communication Unified Architecture (OPC UA) (OPC Foundation 2010). Kommunikation findet zwischen einem OPC UA-Server und einem OPC UA-Client statt, wobei auch mehrere Clients gleichzeitig auf einen Server zugreifen können. Hierbei werden Security Aspekte unterstützt. Zudem bietet OPC UA die Möglichkeit, Daten und Methoden in einem Informationsmodell semantisch zu beschreiben und zu strukturieren. Hensel et al. (2016) stellen ein Konzept vor, in dem jedes Simulationswerkzeug mit einem Adapter kommuniziert, welcher über einen OPC UA-Server und einen OPC UA-Client verfügt. Über diesen Adapter registrieren sich die Werkzeuge an einem zentralen Server, welcher die Co-Simulation koordiniert. Der zentralisierte Ansatz sowie die Notwendigkeit des Servers führen jedoch zu einer Minderung der Flexibilität und zusätzlichem Entwicklungsaufwand.

Ziel dieser Arbeit ist die Erhöhung der Interoperabilität realer und simulierter Daten auf Basis standardisierter und in der Produktion sowie Produktionssimulation etablierter Schnittstellen. Derzeit wird OPC UA von den meisten hier verwendeten Simulationstools unterstützt, weshalb sich die im Folgenden vorgestellte Methode auf diese Kommunikationsarchitektur beschränkt. Darauf aufbauend wird ein dezentrales System vorgestellt, welches es ermöglicht, flexibel verschiedene Simulationstools und reale Anlagen zu koppeln.

3 Konzept der dezentralen Kommunikation

Mittels OPC UA-Server können alle realen und simulierten Systeme ihre Daten den anderen Teilnehmern zur Verfügung stellen (Abb. 1). Um mit dem gewünschten Detaillierungsgrad auf die Daten der realen Anlagen zugreifen zu können, werden OPC UA-Server in den Steuerungen der unterschiedlichen Ebenen der Automatisierungspyramide integriert. Auf der Feldebene sind dies einzelne Komponenten, die mit der entsprechenden Intelligenz ausgestattet werden (1). Moderne speicherprogrammierbare Steuerungen (SPS) der Steuerungsebene haben derzeit meist einen OPC UA-Server integriert und können dadurch die Daten eines Anlagenmoduls bereitstellen (2). Auch viele Softwareprogramme für übergeordnete SCADA- und MES-Systeme beinhalten mittlerweile OPC UA-Server, wodurch unter anderem Produktionsplanungsdaten ausgelesen werden können (3). Die Simulationswerkzeuge werden ebenfalls unterschiedlichen Ebenen zugeordnet und stellen ihre Daten mit dem entsprechenden Detaillierungsgrad über einen OPC UA-Server bereit. Dazu gehören einzelne physikalische Komponentensimulationen (4), Produktionsanlagensimulationen (5) und Fabriksimulationen (6).

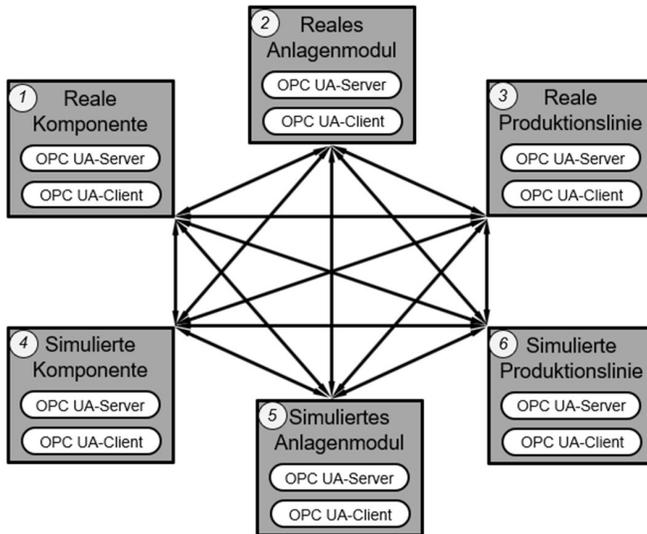


Abbildung 1: Kommunikation zwischen realen und simulierten Systemen

Jedes System mit implementiertem OPC UA-Client kann auf einen OPC UA-Server eines beliebigen anderen Systems zugreifen und so mit diesem interagieren. Neue Systeme können flexibel ergänzt und nicht mehr benötigte entfernt werden, ohne dabei die Kommunikation anderer Systeme zu gefährden.

Für Steuerungen, Softwareprogramme und Simulationstools, die über keinen OPC UA-Server verfügen, können Softwareadapter erstellt werden, die über eine interne Kommunikation Daten austauschen und diese über einen OPC UA-Server zur Verfügung stellen. Ist kein OPC UA-Client vorhanden, kann auch hier mittels Adapter Abhilfe geschaffen werden. Abbildung 2 zeigt die verschiedenen Konstellationen.

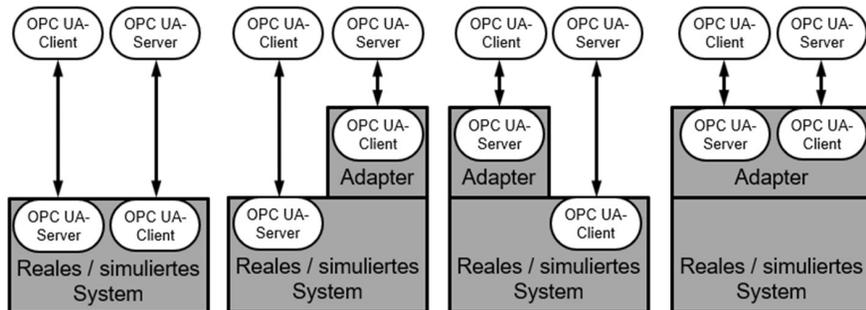


Abbildung 2: Übersicht der Kommunikation ohne und mit Adapter

Auch zusätzliche Software, wie z.B. ein Datenverarbeitungssystem auf einem Edge Device, kann für weitere Anwendungen mittels OPC UA-Client auf die OPC UA-Server zugreifen, um Daten auszulesen oder Methoden auszuführen.

Die Verwendung eines einheitlichen Informationsmodells bei der Strukturierung der Namensräume in den OPC UA-Servern ist wichtig, damit die OPC UA-Clients automatisiert auf die richtigen Daten zugreifen können. In diesem Beitrag wird zwischen dem Datenaustausch zur automatisierten Optimierung von Simulationsmodellen und der Co-Simulation unterschieden. Im Folgenden wird für diese beiden Fälle jeweils der Kommunikationsworkflow beschrieben und eine Methodik vorgestellt, den Namensraum zu modellieren.

3.1 Datenaustausch zur automatisierten Modelloptimierung

Für die Optimierung der Simulationsmodelle liest ein Edge-Device mittels eines OPC UA-Clients die Daten aus dem OPC UA-Server einer Informationsquelle während des Produktivbetriebs aus (Abb. 3). Als Informationsquellen dienen beispielsweise reale Anlagen oder detaillierte Modelle anderer Simulationstools. Um noch weitere Daten aufzunehmen, kann das Edge-Device die Informationsquelle (außerhalb des Produktivbetriebs) mittels OPC UA-Methoden in verschiedene Zustände schalten. Die ausgelesenen Werte werden in einer Datenbank abgespeichert. Auf Basis dieser Datensammlung können analytisch detaillierte Simulationswerte und -kurven ermittelt werden. Die Ergebnisse können ebenfalls in der Datenbank abgespeichert und dadurch der Simulation zur Verfügung gestellt werden. Alternativ kann ein Modellgenerator dazu verwendet werden, ein optimiertes Simulationsmodell zu erstellen.

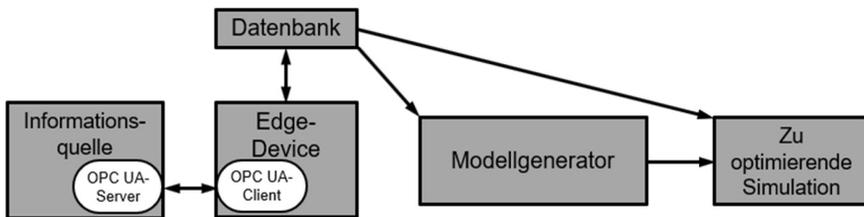


Abbildung 3: Kommunikationsübersicht zur automatisierten Modelloptimierung

Um einen automatisierten Zugriff des Edge-Devices auf die Informationsquellen zu gewährleisten, müssen die Namensräume der OPC UA-Server einheitlich modelliert werden. Dabei wird wie nachfolgend beschrieben vorgegangen.

Falls für den Anlagentyp bereits ein standardisiertes Informationsmodell (Companion Specification) existiert, wird auf diesem aufgesetzt. Ansonsten wird die Hierarchie der Anlage ermittelt (z.B. Linie, Modul, Technologiefunktionen, Komponenten) und darauf aufbauend die Struktur der anfallenden Daten (Variablenknoten) aufgebaut. Die Werte der Messgeräte werden entsprechend dem jeweiligen Messpunkt zugeordnet. Da die Übertragungsgeschwindigkeit von OPC UA meist größer einer Millisekunde ist, werden sich schnell ändernde Anlagenwerte zusammen mit den Zeitwerten als historische Daten oder in Arrays zwischengespeichert und gebündelt zur Verfügung gestellt. Um die Anlage in verschiedene Zustände versetzen zu können, erhalten die schaltbaren Komponenten OPC UA-Methoden (Methodenknoten), die von einem OPC UA-Client aufgerufen werden können. Über

Referenzen werden Beziehungen zwischen den Knoten hergestellt. Dadurch werden z.B. Abschalt- und Anschaltabhängigkeiten schaltbarer Komponenten dargestellt.

3.2 Co-Simulation

Mit Hilfe der Co-Simulation ist es möglich, Simulationen auf unterschiedliche Tools (je nach Detaillierungsgrad und spezifischen Eigenschaften) aufzuteilen, die miteinander gekoppelt sind. Dadurch kann der Rechenaufwand optimiert und der Informationsverlust minimiert werden.

Simulationstools mit OPC UA-Client können auf die Daten der OPC UA-Server anderer Simulationen zugreifen und mit diesen interagieren (Abb. 4). Dadurch können Daten ausgelesen, aber auch Ereignisse anderer Simulationen getriggert werden.

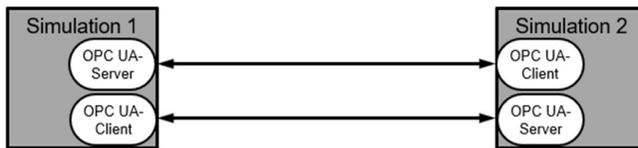


Abbildung 4: Kommunikationsübersicht zur flexiblen Co-Simulation

Die eigenständige Kopplung und Organisation zwischen den Tools führt zu einer hohen Flexibilität. Um dies zu ermöglichen, werden die Namensräume der OPC UA-Server identisch strukturiert.

Zuerst werden Variablenknoten modelliert, welche Ereignisse anderer Simulationen auslösen können. Ihnen untergeordnet finden sich Variablenknoten, die die jeweiligen Ereignisse mit weiteren Daten anreichern. Danach folgen die Knoten, welche die simulierten Daten des Modells zur Verfügung stellen. Diese können, wie in Kapitel 3.1 dargestellt, hierarchisch strukturiert und, wenn aufgrund schneller Signale notwendig, mit Zeitwerten gebündelt übertragen werden.

4 Fallstudien und Ergebnisse

Das vorgestellte Konzept wird an drei Fallstudien getestet. Während sich die erste Fallstudie auf die Optimierung des Simulationsmodells einer Montageanlage konzentriert, wird dieses Simulationsmodell in einer zweiten Fallstudie mit einer Fabriksimulation gekoppelt. In der dritten Fallstudie wird die Kopplung unterschiedlicher Simulationstools, mit dem Ziel ein Energiemanagementsystem im Vorfeld in Betrieb zu nehmen, vorgestellt.

4.1 Optimierung des Simulationsmodells einer Montageanlage

Das für diese Fallstudie verwendete reale Montagemodul beinhaltet vier Förderbänder, eine pneumatische und zwei elektrische Handhabungseinheiten. Es ist virtuell mit dem Simulationstool ISG-virtuos abgebildet. Das Ziel besteht darin, das energetische und zeitlich Verhalten der simulierten Anlagenkomponenten in Bezug auf Ein- und Ausschaltvorgänge zu optimieren. Der Workflow wird in Abbildung 5 dargestellt.

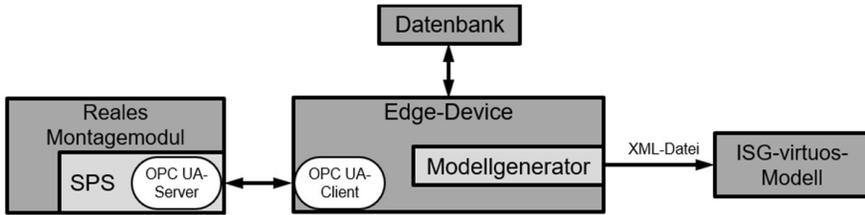


Abbildung 5: Kommunikationsübersicht zur Optimierung einer Montagesimulation

Das Montagemodul wird von einer SPS gesteuert, welche über einen integrierten OPC UA-Server alle nötigen Daten und Interaktionsmöglichkeiten bereitstellt. Die Strukturierung des Namensraumes geschieht bei einer hierarchischen und funktionalen Programmierung des SPS-Programmes (Zizler et al. 2018) automatisch. Ein Edge-Device liest die nötigen Informationen über das Modul aus und kann anhand vorgegebener Aus- und Einschaltregeln das Modul in die relevanten Zustände versetzen. Während dieses Vorgangs werden die Zustände der Komponenten sowie die Leistungen aufgezeichnet. Danach wird analytisch ausgewertet, welche Änderungen von Zuständen mit welchen Leistungsänderungen korrelieren und welche Zeiten hierfür jeweils benötigt werden. Ein ebenfalls im Edge-Device implementierter Modellgenerator erstellt anhand der Ergebnisse eine XML-Datei. Die Software ISG-virtuos importiert die XML-Datei und erstellt daraus ein neues und optimiertes Simulationsmodell.

Abbildung 6 zeigt im Vergleich den realen Leistungsverlauf, den simulierten Verlauf basierend auf vorhandenen Datenblatt- und Datenbankwerten und den optimierten Verlauf beim schrittweisen Aus- und Einschalten der Anlage.

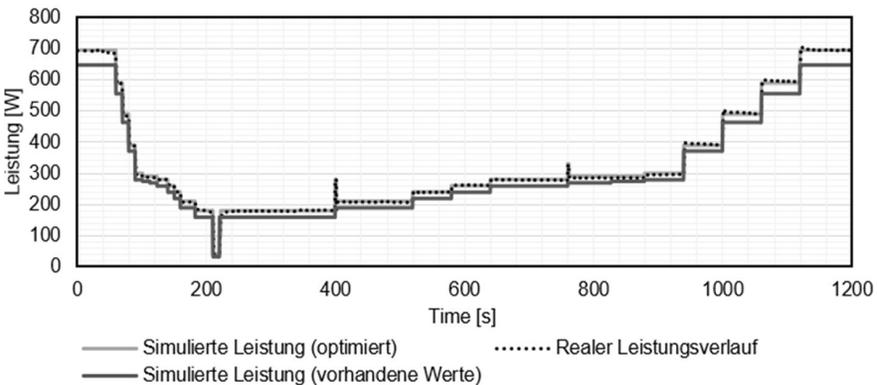


Abbildung 6: Realer und simulierter Leistungsverlauf

Die in den Datenblättern der Komponenten aufgeführten und in Datenbanken von früheren Messungen gespeicherten Leistungsdaten, sind für einige Simulationsstudien zu ungenau. Eine automatisierte Optimierung des Simulationsmodells führt, wie hier dargestellt, aufwandsarm zu einer ausreichend genauen Simulation.

4.2 Kopplung der Montageanlagensimulation mit einer Fabriksimulation

In der zweiten Fallstudie wird das in der ersten Fallstudie simulierte Montage-Modul mit einer Fabriksimulation gekoppelt (Abb. 7). Für die Fabriksimulation wird das Simulationswerkzeug Plant Simulation eingesetzt. Da beide Simulationstools nur über einen OPC UA-Client verfügen, wird für die Implementierung der OPC UA-Server jeweils ein Softwareadapter erstellt.

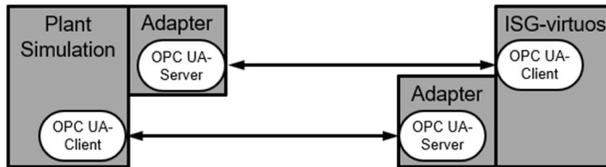


Abbildung 7: Kommunikationsübersicht zur Kopplung von Plant Simulation und ISG-virtuos

Somit kann die Software Plant Simulation die Information bereitstellen, dass ein neues Werkstück in das Montage-Modul einfährt und welche Charge produziert wird. Die Software ISG-virtuos kann diese Daten auslesen, auf das einfahrende Werkstück reagieren und einen detaillierten Montageprozess simulieren. Dabei werden Informationen, wie z.B. genaue Energiebedarfe und zeitliches Verhalten, über den OPC UA-Server bereitgestellt. Plant Simulation kann mittels OPC UA-Client auslesen, wann das Werkstück fertig bearbeitet wurde und es zur nächsten Bearbeitungsstation, welche eventuell ebenfalls detailliert über eine Co-Simulation angebunden ist, weitertransportieren. Die Simulationen laufen dabei auf getrennten Rechnern ab, wodurch mehr Rechenleistung zur Verfügung steht.

4.3 Kopplung mit der Simulation energetischer Erzeuger

Das Ziel der dritten Fallstudie ist es, zusätzlich zum Prozessablauf das Energiemanagement im Rahmen der virtuellen Inbetriebnahme im Vorfeld in Betrieb zu nehmen. Hierzu werden drei Simulationssysteme miteinander gekoppelt (Abb. 8).

Die Produktionsinfrastruktur wird in Plant Simulation symbolisch aufgebaut und bildet damit den Backbone, der mittels OPC UA-Client die Daten der anderen Simulationssysteme ausliest, zusammen mit den eigens simulierten Planungs- und Logistikdaten auswertet und mittels OPC UA-Server bereitstellt. Die physikalische Simulation wird in NX (Mechatronics Concept Designer) durchgeführt und von der simulierten SPS PLCSim Advanced gesteuert. Die Sensor- und Anlagendaten der Simulation werden über einen OPC UA-Server bereitgestellt. Die Komponenten der (Leistungs-)Elektronik (z.B. von einer PV-Anlage oder einem Akkumulator) werden in MATLAB Simulink zeitkontinuierlich simuliert. Die Ergebnisse können ebenfalls mittels OPC UA-Server bereitgestellt werden.

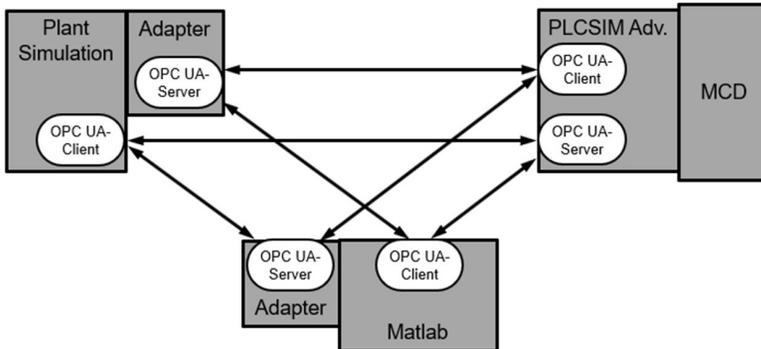


Abbildung 8: Kopplung von drei Simulationssystemen

Die PLCSIM Adv. liest über einen OPC UA-Client die Daten von Plant Simulation und MATLAB Simulink aus und kann, in Abstimmung mit dem Backbone, beispielsweise den Prozessablauf der Anlage anpassen, wenn der Akkumulator den minimalen Ladezustand erreicht hat und durch die PV-Anlage keine zusätzliche Energie bereitgestellt werden kann. Des Weiteren kann auch die Simulation in MATLAB Simulink auf die Daten der beiden anderen Simulationssysteme zugreifen, wodurch z.B. die Ladezyklen des Akkumulators optimiert werden können.

5 Zusammenfassung und Ausblick

Mit Hilfe simulationsbasierter Werkzeuge können Entwickler im Engineeringprozess vorab entscheidungsrelevante Untersuchungen durchführen. In diesem Beitrag wurden Möglichkeiten vorgestellt, die Interoperabilität zwischen realen und simulierten Produktionssystemen zu erhöhen. Durch einheitliche Informationsmodelle realer und simulierter Anlagen ist das automatisierte Auslesen benötigter Daten zur Optimierung von Simulationsmodellen möglich. Für die Co-Simulation werden den Informationsmodellen Variablenknoten zum Auslösen von Ereignissen vorangestellt. Folglich können Simulationsmodelle mit unterschiedlichen Detaillierungsgraden und Eigenschaften flexibel und ressourcenverteilend gekoppelt werden.

Damit die Interoperabilität künftig für Anwender aus unterschiedlichen Bereichen gewährleistet werden kann, ist es notwendig, Informationsmodelle für die verschiedensten Anlagentypen in gemeinsamen Gremien und für alle zugänglich zu standardisieren.

Danksagung

Die Autoren danken dem Bayerischen Staatsministerium für Wirtschaft, Landesentwicklung und Energie für die Förderung des Forschungsprojekts ISAC@OTH-AW.

Literatur

- Blochwitz, T. et al.: Functional Mockup Interface 2.0: The Standard for Tool independent Exchange of Simulation Models. In: Proceedings of the 9th International Modelica Conference, Linköping University Electronic Press, 2012, S. 173–184.
- fmi-standard: Tools. <https://fmi-standard.org/tools/>, letzter Zugriff am 10.05.2021.
- Hensel, S.; Graube, M.; Urbas, L.; Heinzerling, T.; Oppelt, M.: Co-Simulation with OPC UA. In: 2016 IEEE 14th International Conference on Industrial Informatics (INDIN), Poitiers, 2016, S. 20-25.
- IEEE Standards Association: IEEE 1516-2010 - IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA)-- Framework and Rules, 2010.
- Jaensch, F.; Csiszar, A.; Kienzlen, A.; Verl, A.: Reinforcement Learning of Material Flow Control Logic Using Hardware-in-the-Loop Simulation. In: 2018 First International Conference on Artificial Intelligence for Industries (AI4I), Laguna Hills, 2018, S. 77–80.
- Kohl, J.; Spreng, S.; Franke, J.: Discrete Event Simulation of Individual Energy Consumption for Product-varieties. In: ElMaraghy, H. (Hrsg.): Proceedings of the 47th CIRP Conference on Manufacturing Systems (CMS), Amsterdam, Elsevier B.V., 2014, S. 517–522.
- Nyhuis, P.; Fronia, P.; Pachow-Frauenhofer, J.; Wulf, S.: Wandlungsfähige Produktionssysteme - Ergebnisse der BMBF-Vorstudie „Wandlungsfähige Produktionssysteme“. In: wt Werkstattstechnik online (Band 4), 2009, S. 205-210.
- OPC Foundation: Norm IEC/TR 62541-1:18.02.2010. OPC Unified Architecture - Part 1: Overview and Concepts, 2010.
- Schmidl, E.; Fischer, E.; Steindl, J.; Wenk, M.; Franke, F.: Reinforcement learning for energy reduction of conveying and handling systems. In: Procedia CIRP, Volume 97, 2021, S. 290-295.
- Seow, Y.; Rahimifard, S.: A framework for modelling energy consumption within manufacturing systems. In: CIRP Journal of Manufacturing Science and Technology (3), 2011, S. 258–264.
- Sobie, C.; Freitas, C.; Nicolai, M.: Simulation-driven machine learning: Bearing fault classification. In: Mechanical Systems and Signal Processing Volume 99, 2018, S. 403-419.
- Straßburger, S.: On the Role of Simulation and Simulation Standards in Industry 4.0. In: 2019 Simulation Innovation Workshop (SIW), Orlando, USA, 11.-15.02.2019.
- Tur, B.; Fischer, E.; Franke, J.; Bergs, C.; Schulze, C.: Modelling of Energy Storage Devices and Converters for Energy Flow Simulation in Plant Simulation. In: Putz, M.; Schlegel, A. (Hrsg.): ASIM 2019. Simulation in Produktion und Logistik. Auerbach /Vogtl.: Wissenschaftliche Scripten, 2019, S. 97-110.
- Wenk, M.: Potentiale der Virtuellen Inbetriebnahme für die industrielle Anwendung und die Lehre. In: Tagungsband AALE 2014, Regensburg, 08.-09.05.2014, S. 175-184.
- Zizler, E.; Wenk, M.; Martens, B.: Self-Learning Energy Management System on the Process Control Level. In: Applied Mechanics and Materials, Vol. 882, 2018, S. 3-9.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Ein unternehmensübergreifender Standard für Logistiksimulation mit grafischer Prozessmodellierung

A Cross-Company Standard for Simulation of Transport Logistics with Graphic Process Modelling

Kristina Sokoll, AUDI AG, Ingolstadt (Germany), kristina.sokoll@audi.de

Philipp von Braunschweig, Volkswagen AG, Wolfsburg (Germany),
philipp.von.braunschweig@volkswagen.de

Linda Gustafsson-Ende, BMW Group, München (Germany),
linda.gustafsson-ende@bmw.de

Frank Hilmer, SimPlan AG, Hanau (Germany), frank.hilmer@simplan.de

Abstract: The ongoing transformation of the automotive industry and the increasing digitalization and automation require complex planning methods in even shorter planning cycles. Simulation is an established technique, but still very laborious for some applications such as logistics. The “Transport Logistics” module is a cross-company approach for standardization of logistic simulations within an existing object library. The module groups logistic information in structure as well as process information and provides solutions for mastering typical challenges when simulating a logistic system. Especially the new approach of graphic process modelling allows the implementation of transportation processes in an effective and comprehensible manner. This leads to a reduction of modelling times and a high level of flexibility and reusability in future projects. The module received positive feedback from first project experiences, where its suitability and the improvement in modelling of both structure and transportation processes could be confirmed.

1 Motivation und Zielsetzung

Die Automobilindustrie unterliegt einem stetigen Wandel, der nicht nur die Anforderungen an Produkte und Strategien betrifft, sondern auch Einfluss auf die Planung und Optimierung von Fertigungs- und Logistiksystemen hat. Neue Technologien und ein zunehmender Digitalisierungs- und Automatisierungsgrad beeinflussen sowohl neue Fahrzeugmodelle als auch die dafür auszulegenden Produktionsanlagen und logistischen Transportnetzwerke. Immer komplexere Steuerungen, intelligente und vernetzte Systeme und kürzere Planungszyklen haben

die Ablaufsimulation in vielen Produktionsplanungsprojekten zu einem festen Bestandteil zur Absicherung und Optimierung werden lassen.

Im Bereich Logistik bietet Simulation die Möglichkeit, Versorgungsprozesse und erforderliche Ressourcen, wie Fläche, Material oder Transportmedien zu analysieren und verschiedene Szenarien zu untersuchen, um eine optimierte Lösung im Hinblick auf Versorgungssicherheit sowie Invest- und Betriebskosten zu finden. Im Folgenden wird dafür der Begriff der Transportlogistiksimulation verwendet, der sich hier grundsätzlich auf alle üblichen Ausprägungen in der Automobilindustrie bezieht. Einsatzgebiete sind beispielsweise Werkssimulation, Belieferungssimulation, Supply-Chain-Simulation und Verkehrsflusssimulation, wobei erfahrungsgemäß der Großteil der Planungsprojekte im Bereich der Belieferungssimulation einzuordnen ist (vgl. Sprock und Hilmer 2020).

Die Herausforderung besteht darin, dass für die Logistikplanung oftmals maßstabsgetreue und hochdetaillierte Simulationsmodelle erforderlich sind, um die zur Untersuchung stehenden Fragen beantworten zu können (Müller-Sommer 2013). Die Entwicklung und Anpassung solcher Modelle ist nicht nur aufgrund des hohen Datenbedarfes sehr zeitaufwändig, sondern vor allem aufgrund der Abbildung der häufig komplexen Steuerungslogiken. Eine interne Umfrage im Volkswagen Konzern hat ergeben, dass – unabhängig von Anwendungsgebiet und erzieltm Projektergebnis – fast ein Drittel der befragten Auftraggeber die langen Entwicklungszeiten kritisch sehen, da eine simulative Unterstützung der Planung unter Umständen nicht innerhalb des Projektzeitrahmens möglich ist. Laut einer Umfrage im VDA entfallen bei Belieferungssimulationen ca. 38% des Gesamtaufwandes auf den Bereich Modellentwicklung (Müller-Sommer 2010), was – neben der Schaffung von geeigneten Datenstrukturen und -schnittstellen – zu einem großen Teil der aufwändigen Implementierung der Transportstrategien zuzuschreiben ist.

Vor diesem Hintergrund bietet sich die Chance, dass insbesondere durch eine Vereinfachung der Prozessmodellierung eine deutliche Effizienzsteigerung bewirkt werden kann. Damit kann nicht nur die Entwicklungszeit verkürzt, sondern auch die Handhabbarkeit und Flexibilität im Projektverlauf erhöht werden. Daher sollte eine neuartige grafische Prozessmodellierung im Zentrum der gemeinsamen Entwicklung stehen. Darüber hinaus sollten die Funktionen innerhalb eines möglichst großen Anwenderkreises abgestimmt werden, um für die bekannten Anwendungsfälle aus der Planungspraxis einsetzbar zu sein. So kann sich eine Modellierungstechnik schneller zu einem akzeptierten Standard entwickeln, was eine hohe Effizienz bei der Simulation bewirkt und die Wiederverwendbarkeit verbessert. Außerdem erleichtern Standards die Weiterverwendung von Simulationsmodellen durch einen anderen Entwickler oder die Einbindung externer Dienstleister, was in Anbetracht der häufig mehrjährigen Planungsphasen in der Logistik von großem Interesse ist.

Innerhalb existierender Simulationstools gibt es bereits Standards für verschiedene Anwendungsgebiete. In der Automobilindustrie ist für die Software Plant Simulation die umfangreiche Bibliothek „Automotive Library“ (ehem. „VDA Automotive Bausteinkasten“, vorgestellt durch Mayer und Pöge (2010)) zu nennen, die in den an der Entwicklung beteiligten Unternehmen (AUDI AG, BMW Group, Daimler AG, Volkswagen AG, ZF Friedrichshafen AG) als Modellierungsstandard zum Teil bereits seit 15 Jahren zum Einsatz kommt und stetig weiterentwickelt wird. Die Bibliothek

Es ist zu sehen, dass fast ein Jahr Arbeit in Spezifikation und Ausarbeitung eines geeigneten Konzeptes geflossen ist. Diese umfassende vorgelagerte Entwurfsphase war nicht nur aufgrund der Vielzahl logistischer Prozesse und Prozesseigenschaften erforderlich, die zu berücksichtigen waren, sondern auch aufgrund der neuartigen grafischen Prozessmodellierung.

Die Implementierungsphase von 2018 bis 2019 wurde nach den Prinzipien agiler Softwareentwicklung gestaltet. In kleineren Zeitabschnitten wurden sinnvolle geschnittene Entwicklungspakete umgesetzt, getestet, optimiert und kontinuierlich erforderliche Anpassungen an der Spezifikation diskutiert. In größeren Review Workshops wurden rollierend nicht nur Funktionalität, sondern auch Verständlichkeit und Usability der einzelnen Komponenten und deren Zusammenspiel anhand von realen Prozessbeispielen untersucht. Die Entwicklung wurde dabei innerhalb eines stetig wachsenden Simulationsmodells vorgenommen. Dadurch wurde sichergestellt, dass jeder neu entwickelte Baustein und jede neue Funktionalität gleich in mehreren Varianten und mit verschiedenen Parameterkombinationen getestet werden konnte. Zum Ende der Implementierungsphase enthielt dieses Modell eine Vielzahl an Modellierungsbeispielen für die häufigsten Anwendungsfälle. Es wurde daher entschieden, es den Anwendern des Moduls zur Verfügung zu stellen, damit es neben der Bibliotheksdokumentation als Modellierungsreferenz genutzt werden kann.

Trotzdem führt die neuartige Prozessmodellierung, zusammen mit der Vielzahl neuer Objekte im Modul, sowie deren Kombinations- und Parametrierungsmöglichkeiten, zu einem hohen Aufwand für die Einarbeitung im Vergleich zu anderen Modulen der Objektbibliothek „Automotive Library“. Aus diesem Grund wurde bereits zu Beginn der Entwicklung die zeitnahe Bereitstellung einer Schulung zum Modul entschieden, um möglichst schnell interessierte Anwender auf standardisiertem Weg zu befähigen, Logistiksimulationsprojekte durchführen zu können. Die Entwicklung der Qualifizierungsmaßnahme startete nach Abschluss der Implementierung im Jahr 2020.

Inhaltlich sieht das Schulungskonzept vor, dass die Teilnehmer ein vollständiges Logistiksystem bestehend aus Wegenetz sowie ein- und mehrstufigen Transportprozessen Schritt für Schritt aufbauen. Dabei werden Transportumfänge (Behälter verschiedener Teilenummern) von Staplern, Zugmaschinen und einem fahrerlosen Transportsystem zwischen Fertigungslinie und Vollgut- bzw. Leergutpuffer ausgetauscht. Die Bestückung des Vollgutpuffers erfolgt über einen nachgelagerten LKW-Versorgungsprozess. Diese Transportprozesse sind, auch wenn für die Schulung in einigen Punkten vereinfacht, bereits auf viele Anwendungsfälle in der realen Planung übertragbar. Sie bieten so eine gute Basis für den Einstieg in reale Simulationsprojekte, da die häufigsten Arbeitsschritte aus den Bereichen Struktur- und Prozessmodellierung sowie die statistische Analyse im Rahmen der Schulung ein- oder mehrfach mit verschiedenen Parametrierungen abgehandelt werden.

3 Aufbau und Funktion des neuen Moduls

In der Logistiksimulation gibt es oft komplexe Wechselwirkungen zwischen den baulichen Gegebenheiten, den abzubildenden Prozessabläufen und den Eigenschaften der transportierten Güter. Erfahrungen aus früheren Simulationsprojekten haben gezeigt, dass eine standardisierte, klar strukturierte Datenhaltung und funktionale Abgrenzung sehr wichtig sind, um Modelle effizient erstellen zu können und die

Weiterverwendung oder Weitergabe vorhandener (Teil-)Modelle zu erleichtern. Die Informationen und Funktionsbausteine werden daher konsequent in zwei Gruppen geteilt:

- Struktur: Layout (Wegenetz, Bedarfs-/Lagerorte, Fahrzeugstellplätze etc.), Transportmittel, Regeln für den Fahrzeugverkehr, Verwaltung der Lagerbestände
- Prozess: Prozessablauf (Abfolge von Transport- und Arbeitsschritten), die dafür benötigten Ressourcen, spezifische Eigenschaften der zu transportierenden Güter

Die Komponenten sind über Schnittstellen miteinander verbunden und greifen auf eine zentrale Datenbasis zu. Der zweite wichtige Ansatz, welche den höchsten Nutzen verspricht, ist die grafische Modellierung und Steuerung der Transportprozesse. Diese kann nicht nur die Entwicklungsphase verkürzen, sondern auch das Verständnis und die Übersicht über die modellierten Abläufe erleichtern. Im Folgenden werden Details aus den Bereichen Struktur- und Prozessmodellierung sowie Möglichkeiten für die statistische Auswertung des Moduls näher erläutert.

3.1 Strukturmodellierung

Die Strukturmodellierung erfordert für gewöhnlich die maßstabsgetreue Abbildung des Wegenetzes, auf dem sich die Transportressourcen bewegen. Das Modul „Transport Logistics“ beinhaltet verschiedene Werkzeuge, die die Erstellung selbst großer Wegenetze sowie deren Parametrierung und Verwaltung erleichtern.

In der Regel liegt das Hallenlayout als CAD-Datei vor. Nicht immer ist daraus ein Wegenetz einfach ableitbar. Hinzu kommt, dass es für die Objektbibliothek „Automotive Library“ und die im Einsatz befindlichen CAD-Tools keine standardisierte Schnittstelle gibt, um das Wegenetz automatisch in das Simulationstool zu übertragen, weshalb es in der Regel im Simulationstool neu modelliert werden muss. Dies kann vor allem an Kreuzungen und den Ein- und Ausfahrten zu den Be- und Entladepunkten sehr aufwändig sein.

Daher wurde zur Unterstützung das „Layout Tool“ entwickelt, welches Kreuzungen automatisch in das Wegenetz einfügt. Dies ist immer dann hilfreich, wenn die Straßen annähernd in einem Rechteckraster angeordnet sind, wie es in Fabrikhallen oft der Fall ist. Der Vorteil ist, dass durchgehende Straßen schnell ohne Berücksichtigung von Kreuzungen in einem Stück gezeichnet werden können. Das „Layout Tool“ unterbricht dann automatisch die Wege an den Kreuzungspunkten und setzt dort die Kreuzungsbausteine des Moduls „Transport Logistics“ ein. Diese Bausteine erstellen automatisch die Verbindungswege innerhalb der Kreuzung. Dafür stehen verschiedene vordefinierte Kreuzungslayouts zur Auswahl.

Abbildung 2 zeigt eine einfache Kreuzung, für deren Modellierung der Anwender lediglich die vier Streckenobjekte ins Modell einsetzen muss. Mithilfe des „Layout Tools“ können daraus verschiedene Kreuzungsvarianten erstellt werden. Die Kreuzungsobjekte beinhalten nicht nur die Wege innerhalb der Kreuzung, sondern stellen ebenfalls Steuerungen für gängige Vorfahrtsregeln zur Verfügung. Sowohl das Kreuzungslayout, als auch die Vorfahrtsregeln können individuell angepasst werden. Selbst komplexe Kreuzungslayouts lassen sich so mit wenig Aufwand modellieren. Der rechte Teil von Abbildung 2 zeigt einen solchen Anwendungsfall, bei dem unterschiedliche Fahrzeugtypen verschiedene Fahrspuren belegen und die Kreuzung mit unterschiedlichen Prioritäten passieren dürfen.

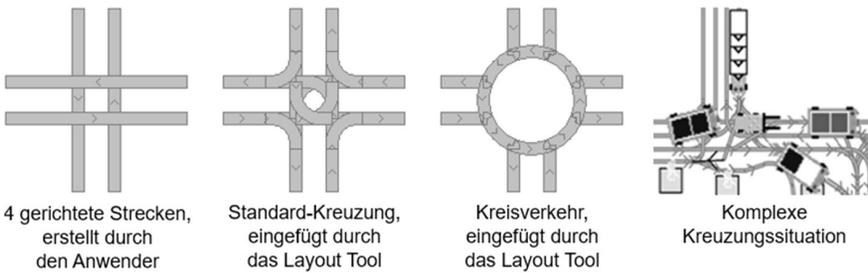


Abbildung 2: Möglichkeiten des „Layout Tools“ für verschiedene Kreuzungen

Auch das Erstellen der „Logistikstationen“, welche z. B. Be- und Entladepunkte oder Fahrzeugstellplätze abbilden, ist sehr einfach. Nachdem der Baustein in das Layout eingesetzt wurde, erstellt dieser die Verbindungen zur nächstgelegenen Straße automatisch. Es ist einstellbar, an welchen Seiten Verbindungen erstellt werden sollen. Die folgende Abbildung zeigt einen Ausschnitt aus einem Wegenetz mit verknüpften Stationen, die verschiedene Typen von Logistikstationen für Lagerung, Beladung, Batterieladen oder Parken repräsentieren. Unterschiedliche Fahrzeuge bewegen sich zwischen diesen Stationen auf den Fahrstrecken.

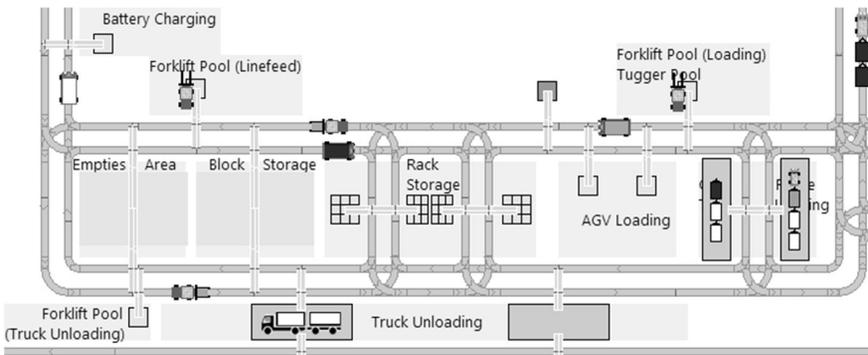


Abbildung 3: Beispiel für ein Wegenetz mit verschiedenen Logistikstationen

Zusammen mit weiteren vielfältigen Möglichkeiten, um den Fahrzeugverkehr zu beeinflussen, wie z. B. Streckensperrungen, Geschwindigkeitsrestriktionen, Überhol- oder Vorbeifahrverboten, werden dem Anwender Lösungen für die häufigsten Anforderungen bei der Modellierung von Wegenetzen zur Verfügung gestellt. Die Bausteine müssen lediglich projektspezifisch parametrisiert oder in Einzelfällen über vorhandene Programmierschnittstellen modifiziert werden.

3.2 Prozessmodellierung

Das Kernstück des Moduls ist die grafische Modellierung und Steuerung der Transportprozesse. Zu diesem Zweck stellt die Bibliothek Objekte zur Verfügung, die eine visuell leicht nachvollziehbare und flexibel anpassbare Abbildung von logistischen Abläufen in einem steuernden Prozessmodell ermöglichen. Diese Prozessbausteine gibt es für die typischen Aufgaben bei der Transportabwicklung aus

den Bereichen Auftragsverarbeitung, Prozesssteuerung/-fluss, Fahrzeugbewegung, Handling, Bestandsverwaltung und Ressourcenzuordnung. Die Symbole für diese Bereiche sind farbkodiert, um die Übersichtlichkeit und Lesbarkeit zu unterstützen.

Jedes dieser Objekte ist ein komplexes, hierarchisch aufgebautes Element mit Parametern sowie Steuerungs- und Schnittstellenmethoden. Es verfügt über einen Dialog, über welchen der Anwender alle vordefinierten Einstellungen vornehmen oder die Programmierschnittstellen öffnen kann, in welchen über Quellcode das Verhalten des Bausteins projektspezifisch angepasst werden kann. Zum Aufbau eines lauffähigen Prozessmodells gilt es, die Prozessbausteine sinnvoll zu einem Ablaufdiagramm zu verketten und zu parametrieren. Interne Logiken überprüfen sowohl die Gültigkeit der Verbindungen als auch mögliche Fehler oder Lücken in der Parametrierung und tragen damit auch zur Vereinfachung der Validierung bei.

Während der Simulation wird der Prozessablauf von beweglichen Objekten – Tour-BE genannt – gesteuert, die sich durch das Netzwerk der Prozessbausteine bewegen. Nach Eintritt des Tour-BEs in einen Prozessbaustein wird der Prozessschritt, den dieser repräsentiert, ausgeführt (z. B. eine Fahrt zu einem bestimmten Ziel, Beladen/Entladen, Warten auf Ressourcen). Der Prozessbaustein interagiert dabei mit den Objekten im Strukturnetzwerk. Nach Abschluss des Prozessschritts springt das Tour-BE zum nächsten Prozessbaustein. Abbildung 4 zeigt ein einfaches Beispiel für eine getaktete Fahrt entlang einer definierten Route, die durch eine Folge von Logistikstationen definiert ist. Das Tour-BE befindet sich auf dem „Drive“-Baustein, was bedeutet, dass zu dieser Zeit das Transportfahrzeug im Wegenetz zum nächsten Ziel fährt. Dort angekommen ist die Fahrt beendet und das Tour-BE geht in den „Decision“-Baustein über.

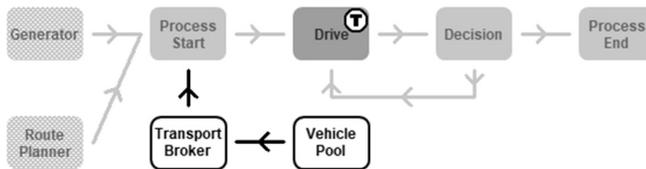


Abbildung 4: Einfacher Fahrprozess – das Tour-BE befindet sich auf „Drive“

Ablauf des gesamten dargestellten Prozesses:

- Der „Generator“ erzeugt in bestimmten Zeitabständen Tour-BEs, die jeweils eine bestimmte Fahrtroute repräsentieren.
- Am „Process Start“-Baustein erhält das Tour-BE vom „Route Planner“ die Information, welche Stationen angefahren werden müssen. Über den „Transporter Broker“ wird ein Transportfahrzeug reserviert, sobald eines verfügbar ist. Die Eigenschaften und verfügbare Anzahl der Transportfahrzeuge werden im „Vehicle Pool“ definiert.
- Der „Drive“-Baustein modelliert die Fahrt zur nächsten Station der Fahrtroute.
- Der „Decision“-Baustein setzt das Tour-BE wieder auf den „Drive“-Baustein, solange laut Route noch weitere Stationen angefahren werden müssen. Andernfalls springt das Tour-BE zum Baustein „Process End“, wo die Tour beendet wird und statistische Kenngrößen der Tour ermittelt werden.

Der hier dargestellte Prozess ist ohne jeglichen Programmieraufwand funktionsfähig. Aber auch komplexere Abläufe z. B. mit Fahrtrouten, Zugmaschinen mit Anhängern, Batterieladekonzepte oder mehrstufige Transporte können vom Anwender ohne größere zusätzliche Implementierung erstellt werden, weil die Bausteine bereits viele verschiedene Möglichkeiten für Transportvarianten zur Verfügung stellen. Die Prozessmodelle können beliebig strukturiert, verzweigt sowie dupliziert oder abgeleitet werden, was sich positiv auf die Möglichkeiten zur Modularisierung auswirkt. Dadurch können die Prozesse einfach von einem Simulationsmodell in ein anderes importiert werden, was zu einer weiteren Vereinfachung und Beschleunigung in der Modellentwicklungsphase und -validierung führt.

3.3 Statistische Auswertung

Die statistische Auswertung erfolgt auf Basis von Ereignisprotokollen, die während des Simulationslaufs aufgenommen werden. Erfasst werden unter anderem.

- Vorbeifahrt von Fahrzeugen an Erfassungspunkten im Wegenetz
- Statusänderungen im Prozess
- Eintritt und Austritt von Tour-BEs in Prozessbausteine
- Abschluss von Touren
- Versorgungsabrisse

Bei jedem Ereignis werden spezifische, für die Auswertung relevante Daten mit abgespeichert. Diese können vom Anwender projektspezifisch erweitert werden. Aus den Ereignisprotokollen können verschiedene Kenngrößen ermittelt werden, wie z. B. Auslastungen, Füllstände und Durchlaufzeiten. Für die Analyse und Visualisierung dieser umfangreichen Datensammlung liefert das Modul „Transport Logistics“ vordefinierte Statistikausteine, welche die häufigsten logistischen Kenngrößen abdecken, die für die Bewertung eines Logistiksystems erforderlich sind. Grundsätzlich ist es aber auch möglich, die Rohdaten projektspezifisch zu verarbeiten oder zu exportieren und in alternativen Softwaretools aufzubereiten und zu analysieren.

4 Fazit und Erfahrungen aus der Praxis

Ziel des Entwicklungsprojektes war es, einen Standard für die Transportlogistiksimulation als Bestandteil der unternehmensübergreifend genutzten „Automotive Library“ zu entwickeln, der die Simulation von Logistiksystemen effizienter und flexibler gestalten sollte. Dies wurde mit der Fertigstellung und dem Roll-Out des entstandenen Moduls „Transport Logistics“ erreicht. Es steht momentan den an der Entwicklung beteiligten Unternehmen sowie deren Projektpartnern zur Verfügung. In die Modulentwicklung sind Knowhow und Funktionalitäten verschiedener existierender Lösungen eingeflossen, um den aktuellen Anforderungen gerecht zu werden. Daher wird eine zügige Etablierung als akzeptierter Standard erwartet. Die bisherigen Erfahrungen aus ersten Planungsprojekten fallen positiv aus, da die Strukturen und Transportprozesse dort mit den zur Verfügung stehenden Möglichkeiten des Moduls abgebildet werden konnten.

Im Hinblick auf die erreichbare Effizienz zeichnet sich ab, dass insbesondere durch die gute Weiterverwendbarkeit der Prozessmodelle eine Steigerung zu erwarten ist, auch wenn sich diese nicht im ersten Projekt erreichen lässt. So sind im Modul z. B.

fast 40 Objekte allein für die Prozessmodellierung, zehn verschiedenen Typen Logistikstationen und über 20 vordefinierte Statistikbausteine enthalten. Die Modulnutzung ist daher erst nach einer gewissen Einarbeitung möglich, mit deren Fortschreiten eine deutliche Effizienzsteigerung erreichbar ist. Hierfür bietet jedoch die angebotene Schulung laut Teilnehmerfeedback eine adäquate Basis, da die häufigsten Objekte dort ein- oder mehrfach Verwendung finden. Die zur Verfügung stehende Moduldokumentation und das umfangreiche Demo-Modell ergänzen das Informationsangebot sinnvoll.

Eine steigende Nutzerzahl wird in absehbarer Zeit zu neuen Ideen und Optimierungen im Hinblick auf Usability und Funktionalität führen. Die entstehenden Anforderungen fließen in den Weiterentwicklungsprozess der „Automotive Library“ ein. Bereits durch ein erstes Pilotprojekt, welches in der späten Implementierungsphase in einem der Unternehmen gestartet wurde, konnten offene Punkte im Konzept aufgedeckt und gleich entsprechende Anpassungen der Spezifikation vorgenommen werden. Darüber hinaus zeigte das Pilotprojekt aus dem Bereich der Belieferungssimulation für eine Fahrzeugmontage, dass auch rollierende Änderungen am Planungsgegenstand durch die zur Verfügung gestellten Lösungen schnell nachvollzogen werden konnten. So wurde das Wegenetz nicht nur initial mithilfe des „Layout Tools“ erstellt, sondern es war auch möglich, die Änderungen der sich weiterentwickelnden Planungsstände auf einfache Art und Weise in das Modell zu übernehmen. Vereinfachend und fehlervermeidend wirkt sich dabei aus, dass aufgrund der zentralen Datenhaltung die Planungsdaten nur an einer Stelle im Modell aktualisiert werden mussten.

Die grafische Prozessmodellierung hat nicht nur eine nachvollziehbare Abbildung der Transportprozesse ermöglicht, sondern zeigte auch, dass sie neue Möglichkeiten für die Zusammenarbeit mit den Planungsabteilungen bietet. Denn die Prozessmodelle sind auch von Logistikplanern ohne Simulationsknowhow nachvollziehbar und konnten im gemeinsamen Austausch validiert und verifiziert werden. Die ersten Erfahrungen und Feedbacks zeigen, dass es sinnvoll sein kann, eine unternehmensinterne Prozessbibliothek für gängige Transportvarianten aufzubauen, um in Projekten auf bereits definierte Standardprozesse zurückgreifen zu können, die lediglich projektspezifisch angepasst werden müssen. Damit könnte den Anwendern eine noch breitere Ausgangsbasis für ihre Simulationsprojekte zur Verfügung gestellt und die Modellentwicklungszeiten damit weiter reduziert werden.

Literatur

- Mayer, G.; Pöge, C.: Auf dem Weg zum Standard – Von der Idee zur Umsetzung des VDA Automotive Bausteinastens. In: Zülch, G; Stock, P. (Hrsg.): Integrationsaspekte der Simulation: Technik, Organisation und Personal. Karlsruhe: KIT Scientific Publishing 2010, S. 29-36.
- Müller-Sommer, H.: Wirtschaftliche Generierung von Belieferungssimulationen unter Verwendung rechnergestützter Plausibilisierungsmethoden für die Bewertung von Eingangsdaten. Ilmenau: Universitätsverlag Ilmenau 2013.
- Müller-Sommer, H.: Aufwände der Belieferungssimulation – Umfrage im VDA. https://www.db-thueringen.de/receive/dbt_mods_00015869 (2010).
- Sprock, T.; Hilmer, F.: Bereitstellungssimulation. In: Mayer, G.; Pöge, C.; Spieckermann, S.; Wenzel, S. (Hrsg.): Ablaufsimulation in der Automobilindustrie. Berlin, Heidelberg: Springer 2020, S. 189-204.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

A Simulation Study on Electric Last Mile Delivery with Mobile Smart Cargo Boxes

Eine Simulationsstudie zu elektrischen Liefermodellen für die letzte Meile mit mobilen Smart Cargo Boxen

Fabian Lorig, Emil Johansson, Paul Davidsson, Jan A. Persson, Internet of Things and People Research Center, Department of Computer Science and Media Technology, Malmö University, Malmö (Sweden), fabian.lorig@mau.se, emil.johansson@mau.se, paul.davidsson@mau.se, jan.a.persson@mau.se

Abstract: The increasing popularity of e-commerce requires efficient solutions for the provision of last mile logistics. There are different approaches for delivering parcels, e.g., home delivery, service points, or parcel lockers, which have different advantages and disadvantages for customers and logistics providers in terms of flexibility, accessibility, and operating costs. We have studied a novel transportation solution where electric vehicles dynamically set up smart cargo boxes, from which customers can fetch their delivery at any time of the day. This provides customers with a more flexible access to their packages and allows the service provider to deliver the parcels more efficiently. In this article, we present the results of a feasibility study conducted in Västra Hamnen, Malmö (Sweden). The developed simulation model shows that smart boxes not only are a viable approach for efficient last mile deliveries, but also result in considerably smaller travel distances compared to conventional package delivery.

1 Introduction

Even before the beginning of the Covid-19 pandemic, business to consumer (B2C) e-commerce has experienced a steady growth. However, due to interventions for containing the pandemic such as movement restrictions and lockdowns as well as recommendations against visiting physical stores or even their closure, the importance and popularity of online shopping increased further (Elrhim & Elsayed 2020). A major challenge that arises from the growing trend towards online shopping is the effective realization of B2C last mile delivery, i.e., the delivery of the parcel from a regional depot to the customer (Mangiaracina et al. 2019).

There exists a great number of logistics service providers that pursue different last mile delivery approaches. Common last-mile solutions for B2C are to deliver the parcel to (i) the home address, (ii) service points where they are picked up by the

customers, or (iii) stationary parcel lockers that are located at supermarkets or other frequented places, where customers can pick up their parcels. As outlined by Allen et al. (2007), these delivery alternatives provide customers but also service providers with different advantages and disadvantages regarding flexibility, accessibility, and effort of the pick-up process. Customers, for instance, experience shorter retrieval times and are not limited to opening hours when using parcel lockers compared to service point deliveries. However, compared to attended home deliveries, parcel lockers and service points do not require the customer to be present upon delivery and, thus, increase the customers' flexibility in terms of the delivery time window. Yet, pick-up points are limited in their opening hours and require each customer to travel there, which they might consider to be inconvenient.

From a logistics provider's perspective, these delivery options vary in efficiency. Compared to pick-up service points, home delivery results in increased delivery costs, a higher number of failed deliveries, and a significantly greater driving distance for delivery vehicles, which also might affect traffic congestion and exhaust emissions. Moreover, the delivery option with parcel lockers is more time-consuming for the driver as he or she needs to fill the lockers with new deliveries. For both parcel lockers and service point deliveries, it can be assumed that the cumulated customer travel distance is significantly higher in relation to the travel distance of the service provider. It is challenging to take all these, potentially conflicting, requirements into account and to find a balance between the requirements of the customer and logistics providers.

To address this issue, the Swedish start-up DiPP-R (www.dipp-r.com) develops a transportation solution to improve the efficiency of last mile delivery of e-commerce parcels. The idea is that electric vehicles dynamically set up smart cargo boxes at different locations in the city. These boxes can hold between 50 and 100 packages and are dynamically placed at, for instance, parking areas. This enables the customers to pick-up their deliveries at any time during the day. After all parcels were picked up by the customers or after a certain time, the boxes are collected by the vehicles, refilled at the depot, and placed at other locations. The aim is to reduce the handling of parcels outside the depot and to decrease the distance recipients must travel to fetch their parcels. This results in increased convenience and shorter total travel distance, ultimately reducing traffic congestion and environmental impact.

This article presents the results of a feasibility study that was conducted by Malmö University, the city of Malmö, and DiPP-R. As part of this study, a simulation model was developed to investigate the effects of this new delivery concept and how it can be realized. The simulation model allows for analysing the efficiency of different service designs in Västra Hamnen, a district in northern Malmö (Sweden). The simulation model also enables the comparison of this new approach with traditional delivery concepts, such as home delivery and central service points.

In summary, the simulation model can be used to answer research questions such as:

- How does the new concept perform compared to existing delivery services in terms of, for instance, travelled distance, accessibility for customers, environmental and traffic impact?
- How shall smart cargo boxes be configured and how many compartments are required per box?

- How many setup locations are required to efficiently serve a particular area and where should they be located?
- Which pick-up and drop-off strategy is most efficient in terms of travel distance?
- How do variations in demand affect the service quality?

The paper is structured as follows: In Section 2, we provide a description of related work on last mile logistics and its simulation. We then describe the simulation model we developed, followed by a specification of the experiments and the results. After a discussion of the results, we provide conclusions and discuss future work.

2 State of the Art

The use of simulation for analysing and comparing logistics processes is well established and studied (Manuj 2009). In transportation logistics, simulation is used to analyse, for instance, how transport tasks can be allocated to vehicles (Davidsson et al. 2005). According to Olsson et al. (2019), modelling and simulation is the leading methodology used in the emerging research area of last mile logistics. It is applied to investigate, e.g., the effects of different means of delivery such as electric vehicles and cargo-bikes but also the feasibility of crowdsourcing. This is, as simulation allows for creating digital copies (*digital twins*) of real-world systems, that can be used to efficiently investigate the system's behaviour under different circumstances, without influencing or jeopardizing the real-world system.

Grando & Gosso (2005) refer to the issue of identifying the optimal delivery solution as “*Last Mile Logistics Dilemma*” and present a reference model for comparing home delivery with pick-up points. To overcome this dilemma, Perboli et al. (2018) propose a multimodal simulation optimization framework for urban freight transportation of e-commerce deliveries, which allows for analysing different delivery modes in realistic scenarios. Besides such frameworks, there exists other studies for specific scenarios in last mile logistics. This includes, for instance, the use of robots for autonomous last mile deliveries, e.g., (Poeting et al. 2019), or for crowdsourced delivery, e.g., (Guo et al. 2019), where local non-professional couriers deliver the parcels to the customers' homes.

The use of cargo boxes has been mostly studied for scenarios with stationary boxes that are equipped to the customer's house or set up at fixed publicly accessible locations. To optimize the last mile in electronic grocery shopping, Punakivi et al. (2001) simulate the use of delivery and reception boxes for unattended delivery of groceries. Yetis & Karakose (2018) propose the use of smart cargo cabinets that are located within buildings and fed by unmanned aerial vehicles (drones). For a Polish city, a study has been conducted by Iwan et al. (2016). The results show that a reduction of the environmental impact of last mile delivery can only be achieved by alternative delivery concepts such as parcel lockers. A similar study has also been conducted in the Netherlands, which investigated the potential of cost reductions when shifting from home delivery to parcel lockers (Van Duin et al. 2020).

To our knowledge, there exists no simulation studies on smart cargo boxes that are dynamically placed at different locations in the city with the aim of optimizing delivery processes for both customers and logistic service providers.

3 Modelling Last Mile Delivery Options

We implemented an agent-based model (ABM) to investigate the effects of different last mile delivery options. The delivery vehicle, customers, deliveries, smart cargo boxes, potential locations of the boxes, and the depot are implemented as agents. Each day, a number of deliveries arrives to the depot each of which is designated for a specific customer in the simulated area. According to the customers' home addresses, the deliveries will be allocated to boxes such that the customers' travel distance for picking up their parcels is minimized. This includes the clustering of the deliveries for the allocation to the boxes as well as the identification of optimal setup locations for each box. A vehicle will then transport the boxes, one at a time, to their designated location. In case there is already a box standing at this location whose minimum setup time (e.g., after 24 hours) has been reached, it will be replaced, and the previous box is returned to the depot. Packages remaining in the returned box will then be allocated to new boxes in the same manner as newly arrived packages. The vehicle visits the box locations in an order prioritizing empty boxes and those that have exceeded their minimal setup time. Boxes that are placed at a location may not be completely filled with packages, however, the vehicle will never deliver empty boxes and will skip locations to which no packages are to be delivered. During hours with high volume of traffic (e.g., 6 a.m. – 9 a.m. and 3 p.m. – 6 p.m.) the vehicle will not leave the depot to reduce traffic congestion.

Each customer has a home address, from where he or she will pick up the parcel. Once a box with a parcel arrives at a pick-up location, there will be a random delay representing that the customers are occupied with other activities and that they pick up their deliveries at a later point in time. If the box with the package has not been returned to the depot by then, the customer walks to the location of the box, takes its package and walks back home. Otherwise, the customer will be informed when the delivery can be collected from another box. In case the distance to the box is greater than a given threshold, the recipient will choose to take the car instead of walking.

Modelling of customer demand is challenging and requires data on where customers live as well as on their habits. For this study, each customer and delivery need to be assigned to a specific building to adequately simulate driving and walking distances to distribute and pick up parcels. Hence, address data is needed on where people live. This data is usually not openly available such that other data sources must be used to generate realistic artificial data on customer demand. OpenStreetMap (OSM; www.openstreetmap.org) data, for instance, can be used to geographically distribute customers in a realistic way. From OSM, positions of buildings can be extracted as well as their size and utilization. This allows us to identify the potential home addresses of customers and to estimate the likely number of residents. We do this by distributing the known number of inhabitants of the simulated area to the buildings we identified in OSM. Here, we use the floor area of the houses to estimate the number of residents by calculating the average floor area per resident. Due to a lack of data, the modelled population is homogenous in terms of their behaviour and habits, e.g., the threshold when they will use a car to fetch their delivery.

The model allows for comparing the new delivery concept to two traditional package delivery systems: deliveries to service points and home deliveries. For service point deliveries, different pick-up locations are defined, where the service points are located. Each time a customer fetches a package from a box, the walking

distance from their home to one of the delivery locations and back is simulated as well. For home delivery, every time new packages are delivered to the depot, a route is iteratively planned such that packages are delivered to the recipient closest to the last one. This does not return the minimum distance required to deliver all packages but overestimates the delivery distance. Yet, we do not consider extra driving distances potentially caused by time windows for home delivery. Moreover, we assume that only one vehicle is in charge of all home deliveries.

For analysing different scenarios, the model provides the following parameters:

- The **number of packages delivered to the depot** each day. Each arriving package has an individual defined as its recipient.
- The **minimum set-up time of boxes** to stay at a location before it can be picked up or replaced by the vehicle.
- The **package capacity of the boxes**. Packages that do not fit in a box will be delivered with the next box to the same location or a suitable location close by.
- The **rate at which customers collect their packages** as the lambda parameter of the exponential function for determining waiting times of customers.
- The **maximum walking distance of recipients** before taking the car for fetching a delivery. We assume all individuals have access to a car.
- The **ratio of individuals fetching their packages combined with other activities**. If individuals fetch their package together with other activities, e.g., buying groceries, only the additional distance needed to fetch the package is considered. The customer will move from its home to the grocery store, to the box, and back home. In case the grocery store contains a delivery point, the additional distance is zero.

The following outputs are provided for each run of the simulation model:

- The **total distance the delivery vehicle has travelled**.
- The **total time the vehicle is being active**.
- The **total distance of customers** to pick up their packages and to return home.
- The **distance of customers travelled by car** in case the distance to the box is above the car threshold. Driving distance can be longer than walking distance.
- The **number of deliveries that have been picked up** by the recipients.
- The **number of packages that have not been picked up by the customers** and thus were returned to the depot for a new delivery with another box.
- For *service point deliveries*:
 - The **total distance of customers** to fetch their packages and return home.
 - The **total distance of customers travelled by car**.
- For *home delivery*:
 - The **approximated total distance** travelled by the home delivery vehicle.
 - The **approximated total time** all deliveries will take.

4 Case Study: Smart Cargo Boxes in Västra Hamnen

The model was implemented using the AnyLogic simulation framework (www.anylogic.com). For this feasibility study, we have chosen the neighbourhood

of Västra Hamnen in Malmö (Sweden) as the setting for our experiments. The potential locations of the boxes as well as for the depot can be chosen manually.

To generate more realistic results, we have chosen the location of the depot to be close to the depots of other existing logistics providers and identified suitable locations, e.g., parking areas, for setting up the boxes. For the simulation of both service point deliveries and grocery stores, two existing service point locations were chosen. AnyLogic includes built-in geographic information system (GIS) support with real-world road networks, which is used to create routings for vehicles and individuals. The user interface of the simulation is shown in Figure 1.

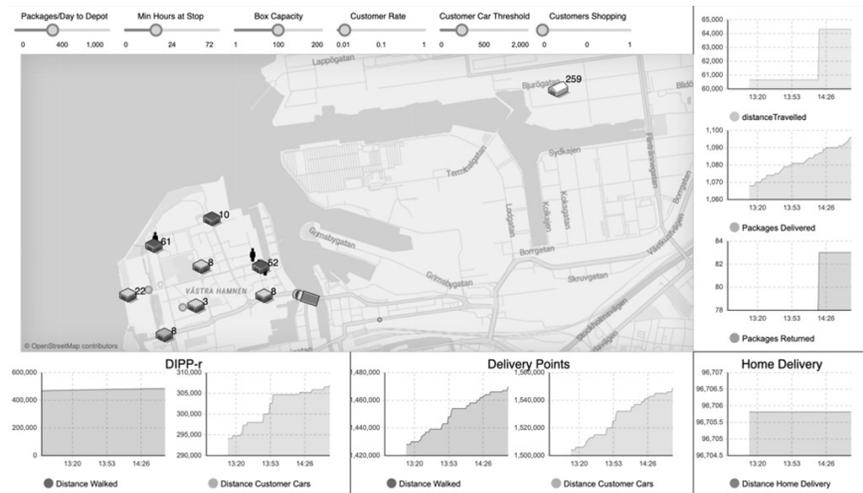


Figure 1: The user interface of the simulation model in AnyLogic.

For the generation of customer address data and customer demand, we have used OSM data of Västra Hamnen in Malmö. According to the data, there are 298 buildings in this area. Buildings with a floor area over 2000 m² were assumed to be industrial buildings and not considered as residence of customers. Buildings between 200 and 2000 m² are assumed to be apartment buildings and below 200 m² as single-family house. In total, Västra Hamnen has 9 739 inhabitants, which were distributed to the existing buildings according to their floor area, resulting in 9 155 customers living in apartment buildings and 584 living in single-family houses. Each of the 9 739 customers was assigned an address according to this distribution.

5 Results of the Simulation Study

For the study, the simulation model was run using different combinations of input parameter values. The simulation starts at 8 a.m. and we simulate an entire week. The presented results were generated using 8 box locations, 9 739 customers, and a single distribution vehicle. With respect to the comparability of the results, all simulations used a fixed random seed.

Figure 2 shows the distance travelled by vehicles for different delivery options and scenarios, i.e., thresholds when customers use their car to get their packages as well as packages per day. When customers chose not to use their car for picking up parcels if the distance is less than 1 km, the mobile smart boxes system results in considerably shorter driving distances compared to the delivery point system. Also, the distance is similar to the distance the home delivery vehicle has to drive, assuming it has a capacity of 100 parcels.

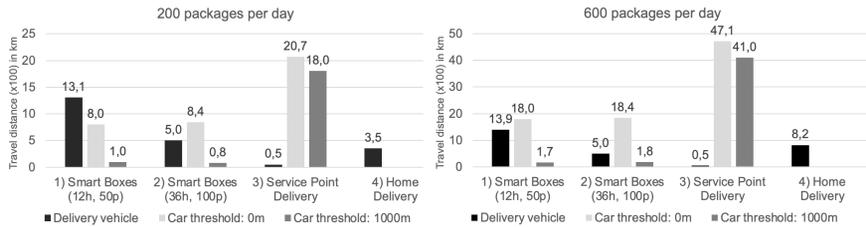


Figure 2: The distance (in km) travelled by the delivery vehicle and by customers using cars at thresholds of 0km and 1km for four scenarios: 1) Boxes, with a capacity of 50 packages waiting for 12 hours at stops; 2) Boxes, with a capacity of 100 packages waiting for 36 hours at stops; 3) Service point delivery; 4) Home delivery. The depot received 200 resp. 600 packages per day.

For the effectiveness of the service, it is not only relevant how many boxes are used but also where they are located. The placement of boxes and its effects on the distance customers must walk can also be explored using the model. For instance, in the two set-ups shown in Figure 3, the cumulated walking distance differs by 3.4%. Hence, the model can be used by decision makers to identify most suitable locations.



Figure 3: Total customers' walking distance for two different placements of boxes.

The model also allows for varying the number of boxes and to investigate the effect this has on the service provision. We simulated the parallel set-up of 4, 6, 8, and 10 boxes with the locations of the boxes being determined using k-means clustering (see Figure 4). The results show a decreasing customers' travel distance and an increasing distance driven by the delivery vehicle, when the number of parallelly used cargo boxes increases. As shown in Table 1, increasing the number of boxes from 4 to 6 results in a 13.9% decrease of the customers' walking distance and 23.2% decrease of the distance driven by car (in total -18.8%) whereas the travel distance of the delivery vehicle increases by 48.8%. When increasing the number of boxes from 6 to 8, the total decrease in customers' travel distance is only 7.5% whereas the distance of the delivery vehicle almost doubles (+91.7%).

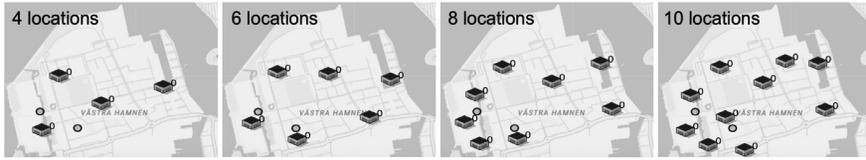


Figure 4: Different number and placement of set-up locations for smart boxes.

Table 1: Traveling distance of customers and delivery vehicle for different number of stops assuming that customers will walk in case the distance is less than 1km.

Number of cargo boxes	Distance walked by customers (km)	Distance driven by customers (km)	Distance delivery vehicle (km)
4	1717	1974	180
6	1478	1516	268
8	1453	1319	345
10	1346	1083	451

It can be assumed that some customers will combine the collection of their delivery with other activities such as grocery shopping, as service points often are located at grocery stores. Figure 5 shows how the percent of individuals fetching their package combined with another activity affects the additional distance travelled by private car for both mobile smart boxes and delivery points. More people combining fetching their package with grocery shopping leads to smaller additional distances travelled. More interestingly, the smart boxes system is shown to lead to smaller distances travelled than the delivery point system for almost all scenarios.

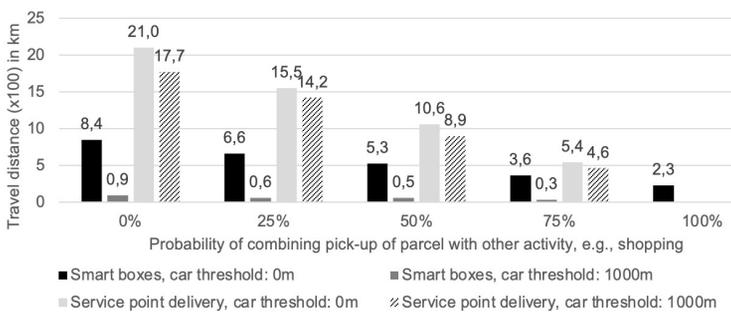


Figure 5: The distance (in km) travelled by customers by private car for different probability of combining the fetching of parcels with other trips, e.g., shopping.

An estimate of the cumulative time it takes to distribute the mobile smart boxes and to make all home deliveries is shown in Figure 6. For home delivery, each delivery is assumed to take one minute per address, the vehicle’s movement speed is 15 km/h, and the capacity is 100 parcels. The figure shows that the time the smart box delivery vehicle is active correlates with the frequency at which boxes are

delivered and returned to the depot. Yet, it is largely unaffected by the number of packages being delivered. This is not true for the home delivery vehicle.

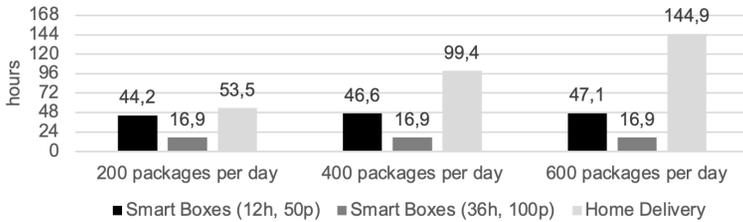


Figure 6: The cumulative time the delivery takes for smart boxes with a capacity of 50 packages being replaced after 12 hours, smart boxes with a capacity of 100 packages being replaced after 36 hours, and home delivery.

6 Conclusions

In this article, we have presented an agent-based simulation model for comparing a delivery solution with mobile smart cargo boxes to existing systems for last-mile delivery. The simulation explores the effects of different smart box service designs and results show that smart boxes are not only feasible as a delivery solution, but significantly decrease the distance customers must travel to fetch their packages and the total distance driven by vehicles compared to service point deliveries. For a car threshold of 1 km, the total vehicle distance is similar to the one of home delivery.

Yet, existing delivery systems have limitations that have not been included in this study. For instance, home delivery might require the recipient to be at home and service points usually have opening hours. With smart boxes solution, however, customers can fetch their packages whenever they desire during the day, allowing for increased flexibility. Moreover, the service provider can set up and collect boxes all day through, which increases the utilization of the vehicles.

There is a trade-off concerning the time boxes stay out before being returned to the depot. A shorter setup time reduces the time packages stay at the depot before being distributed. Recipients, however, have a smaller time window for fetching their packages. This, as well as the fact that not all individuals fetch their packages right away, increases the load at the depot and requires the use of more boxes. Also, reducing the time packages are available for pickup is less convenient to customers.

Examples of simplifications made in the model are the homogeneity of individuals and their habits, the assumption of a static threshold for fetching a parcel by car, and the exclusion of workplaces and other venues than grocery stores and service points. There is also no consideration of exhaust emissions of vehicles, which might be relevant for cities with low-emission zones. Another assumption is that only one delivery vehicle is used for all home deliveries. An extension of the model requires, e.g., data on the capacity of home delivery vehicles and the time to deliver packages.

Besides the design of the service, local regulations and policies might affect the feasibility and viability of deliveries using mobile smart cargo boxes. This includes, for instance, parking regulations that might limit potential locations for setting up boxes and how long they can stand at a location. Moreover, it is uncertain how

different configurations of the service, e.g., the minimum setup time, affect customer acceptance and satisfaction. Yet, the proposed simulation model can be used to investigate different scenarios and to identify potential challenges and opportunities.

With respect to future trends, it is planned to use electric vehicles for the distribution of the boxes. To this end, the effect the battery capacity of the vehicles has on the service needs to be investigated as well as the approaches for charging the vehicles. It can also be assumed that many customers will combine the pick-up of their parcels with, for instance, their work trip. This might affect the optimal placement of the boxes as the location closest to the home might not be most convenient.

Acknowledgement

This work was partially funded by VINNOVA through the project *Smarta Offentliga Miljöer* (Smart Public Environments), a part of the strategic innovation program for the Internet of Things (IoT Sverige), and partially by the Knowledge Foundation through the Internet of Things and People Research Profile.

References

- Allen, J.; Thorne G.; Browne M.: BESTUFS – Good Practice Guide on Urban Freight Transport (2007). URL: http://www.bestufs.net/gp_guide.html
- Davidsson, P.; Henesey, L.; Ramstedt, L.; Törnquist, J.; Wernstedt, F.: An analysis of agent-based approaches to transport logistics. *Transportation Research part C: Emerging technologies*, 13 (2005) 4, pp. 255-271.
- Guo, X.; Jaramillo, Y. J. L.; Bloemhof-Ruwaard, J.; Claassen, G. D. H.: On integrating crowdsourced delivery in last-mile logistics: A simulation study to quantify its feasibility. *Journal of Cleaner Production*, 241 (2019), 118365.
- Iwan, S.; Kijewska, K.; Lemke, J.: Analysis of parcel lockers' efficiency as the last mile delivery solution – the results of the research in Poland. *Transportation Research Procedia*, 12 (2016), pp. 644-655.
- Manuj, I.; Mentzer, J. T. M.; Bowers, M. R.: Improving the rigor of discrete-event simulation in logistics and supply chain research. *International Journal of Physical Distribution & Logistics Management*, 39 (2009) 3, pp. 172-201.
- Olsson, J.; Hellström, D.; Pålsson, H.: Framework of last mile logistics research: A systematic review of the literature. *Sustainability*, 11 (2019) 24, 7131.
- Perboli, G.; Rosano, M.; Saint-Guillain, M.; Rizzo, P.: Simulation-optimisation framework for City Logistics: An application on multimodal last-mile delivery. *IET Intelligent Transport Systems*, 12 (2018) 4, pp. 262-269.
- Poeting, M.; Schaudt, S.; Clausen, U.: Simulation of an optimized last-mile parcel delivery network involving delivery robots. In: Clausen U.; Langkau S.; Kreuz F. (Eds.): *Interdisciplinary Conference on Production, Logistics and Traffic*, Springer, Cham, 2019, pp. 1-19.
- Punakivi, M.; Yrjölä, H.; Holmström, J.: Solving the last mile issue: reception box or delivery box? *International Journal of Physical Distribution & Logistics Management*, 31 (2001) 6, pp. 427-439.
- Van Duin, J. H. R.; Wiegman, B. W.; van Arem, B.; van Amstel, Y.: From home delivery to parcel lockers: A case study in Amsterdam. *Transportation Research Procedia*, 46 (2020), pp. 37-44.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Entwicklung eines anwendungsorientierten Bausteinkastens zur Simulation kombinierter Transportmodelle mittels autonomer Fahrzeuge

Development of an Application-Based Object Library for the Simulation of Combined Transport Systems Using Shared Autonomous Vehicles

Johannes Staritz, Marvin Auf der Landwehr, Maik Trott, Christoph von Viebahn,
Hochschule Hannover, Hannover (Germany), Johannes.Staritz@hs-hannover.de,
marvin.auf-der-landwehr@hs-hannover.de, maik.trott@hs-hannover.de, christoph-
von.viebahn@hs-hannover.de

Abstract: This paper presents an approach to develop region-specific simulation models for quantifying and evaluating the traffic-related, environmental, and economic implications of combined passenger and freight transportation via shared autonomous vehicles (SAV). Based on a broad literature review, conceptual peculiarities, interdependencies and characteristics were derived and transferred into a conferrable, agent-based object library. Finally, to ensure its usefulness and credibility, an initial evaluation of the object library was carried out by developing and visually validating a simulation-prototype for a rural area in Hanover, Germany.

1 Einleitung

Mobilität und Logistik stoßen angesichts gesamtgesellschaftlicher Herausforderungen wie dem Klimawandel, dem demografischen Wandel oder der zunehmenden Urbanisierung sowohl in urbanen als auch in ruralen Gebieten an ihre Grenzen (Moerner 2019; Rodrigue et al. 2020). Angesichts der mit diesen Entwicklungen einhergehenden Implikationen für Umwelt, Gesellschaft und Wirtschaft, gilt es, neueste technische Errungenschaften intelligent zu nutzen und innovative Ansätze für Mobilitäts- und Transportlösungen zu erschließen. Zur Lösung mobilitätsbezogener Herausforderungen entstehen seit geraumer Zeit verschiedene Konzeptansätze, die sich sogenannter geteilter autonomer Fahrzeuge (Shared Autonomous Vehicle, SAV) bedienen (Narayanan et al. 2020). Aufgrund ihrer einzigartigen Betriebscharakteristika versprechen autonome Fahrzeuge hohe Effizienzgewinne, die sich sowohl in wirtschaftlichen (z. B. Profit) als auch gesellschaftlichen (z. B. Verkehrsbelastung) Zielgrößen widerspiegeln und entsprechend nicht nur in mobilitätsbezogenen Lösungsansätzen Betrachtung finden, sondern auch im Bereich der Gütertransporte (Monios und Bergqvist 2019; Yu et al. 2020). Trotz der evidenten,

individuellen Potentiale wurden kombinierte Transportmodelle mittels geteilter autonomer Fahrzeuge bislang jedoch selten untersucht (Beirigo et al. 2018; Meyer und Shaheen 2017). Solche Modelle kombinieren Mobilität und Logistik und stellen durch den simultanen Transport von Gütern und Personen einen vielversprechenden und potenziell von der jeweiligen Raumtopologie unabhängigen Ansatz zur Lösung akuter Probleme dar. Durch diverse Spezifika der betroffenen Akteure, Handlungsfelder und Betriebsstrukturen wird in einem solchem System jedoch ein hohes Maß an Wechselwirkungen, Zielkonflikten und konzeptionellem Planungsbedarf verursacht. Aufgrund dieser Komplexität auf Konzeptions- und Umsetzungsebene eignen sich daher insbesondere dynamische Computersimulationen zur Abbildung und Analyse solcher Konzepte (Auf der Landwehr et al. 2020). Simulationen werden häufig bei logistischen und mobilitätsbezogenen Fragestellungen und Systemen eingesetzt, da sie die Möglichkeit bieten, real nicht existente Zusammenhänge, Interdependenzen mit analytisch nicht abbildbarer Komplexität, mehrere Gestaltungsvarianten eines Systems sowie das langfristige Systemverhalten zu analysieren (Wenzel 2018). Hierdurch können sowohl potenzielle, konzeptinduzierte Fehlerquellen frühzeitig identifiziert als auch Investitionsbedarfe bspw. in Flotte und Infrastruktur, adäquat abgeschätzt werden.

2 Problemstellung

Mobilitätskonzepte, die SAV einsetzen, verbinden meist die Komponenten Nachfrage, Routing, Flotte, Fahrzeugzuordnung, Fahrzeugneuverteilung und elektrische Flottenbeladung (Narayanan et al. 2020). Konzeptionelle Überlegungen für einen kombinierten Transport aus Paketen (Logistikaufträgen) und Personen (Mobilitätsaufträgen) legen nahe, die Komponenten Distributionszentren bzw. Depots sowie Halte- bzw. Auslieferungspunkte zu ergänzen. In Abhängigkeit vom jeweiligen Transportmodell und dessen Untersuchungsrahmen treten die genannten Komponenten in variierenden Ausprägungen auf, wodurch sich die zugrundeliegenden Annahmen der jeweiligen Konzepte teilweise stark unterscheiden (Narayanan et al. 2020). Da sich diese Annahmen für gewöhnlich auf die jeweiligen, im beispielhaften Anwendungsgebiet vorliegenden Rahmenbedingungen stützen bzw. sich auf Erhebungsdaten dieser Gebiete beziehen, sind die einzelnen (Simulations-) Szenarien jedoch nicht mit davon abweichenden Bedingungen wie verändertem Nachfrageniveau, unterschiedlichen räumlichen Strukturen oder einem veränderten Kundenverhalten wie bspw. höherer oder niedrigerer Akzeptanz von Wartezeit kompatibel. Aufgrund fehlender Referenzmodelle mangelt es bei Konzeption und Simulation außerdem häufig an geeigneten Orientierungspunkten (Hyland und Mahmassani 2017). Bei simulationsgestützten Ansätzen zeigt sich in der Praxis, dass konzeptionell komplexe Konstrukte, zu denen auch kombinierte Transportmodelle zählen, oft nur ineffizient oder in Teilssegmenten in den Modellen abgebildet werden (Clausen et al. 2019).

3 Lösungsansatz

Angesichts mangelnder Referenzmodelle und um die evaluative Übertragbarkeit zwischen verschiedenen Raumstrukturen zu gewährleisten, besteht das Ziel dieser Arbeit darin, einen anwendungsorientierten Baustein zur Simulation kombinierter Transportmodelle mittels SAV zu konzipieren und dessen

Anwendbarkeit an einem beispielhaften Simulationsprototypen zu demonstrieren. Aufgrund des breiten Methodenspektrums und des Bedarfs zur Implementierung mikroskopischer, nur mittelbar mit dem Verkehr zusammenhängender Zielgrößen (z. B. Profitabilität, Kosten) erfolgt die Umsetzung der Bausteine in der Software AnyLogic (v. 8.7.2) und nicht in einem spezifischen Verkehrssimulationstool.

Konkret ergeben sich eine Reihe spezifischer Anforderungen für ein entsprechendes Simulationsmodell. So sollten die einzelnen Elemente sowohl einfach in Bedienbarkeit und Handhabung sein, als auch schnell erstellt und flexibel an veränderte Rahmenbedingungen und Datenlagen angepasst werden können. Die Nutzung der bereitgestellten, standardisierten Bausteine sollte außerdem Einstiegshürden senken. Zuletzt müssen die Simulationsergebnisse sowohl aussagefähig als auch vergleichbar sein (Wenzel 2018). Ein vielversprechender Lösungsansatz für die geschilderte Problemstellung ist die Entwicklung von Simulationsbausteinen, deren Eingangsparameter variiert werden können, sodass Abweichungen im Rahmen des kombinierten Transportmodells einfach darstellbar sind. Nach Verbraeck et al. (2002) lassen sich Simulationsbausteine als eigenständige, interoperable, wiederverwendbare und austauschbare Einheit definieren, die ihre interne Struktur kapseln und nützliche Dienste oder Funktionen für ihre Umgebung über genau definierte Schnittstellen bereitstellen. Um eine zielgerichtete, zweckdienliche Entwicklung von Simulationsbausteinen zu gewährleisten, muss zunächst der inhaltliche Rahmen des zugrundeliegenden Einsatzgebietes eruiert werden (Verbraeck und Valentin 2008). Die getroffenen Annahmen und Ausprägungen bilden die Grundlage für die spätere Simulationsanwendung und werden deshalb nachfolgend im Detail beschrieben. Dies dient zum einen dazu, ein Verständnis der Funktionsweise des kombinierten Transportkonzepts zu gewährleisten und zum anderen, um zu verdeutlichen welche Eingangsparameter für die Bausteine von Nöten bzw. an welchen Stellen diese zu implementieren sind.

Beim Fahrzeugrouting beziehen sich die relevanten konzeptionellen Überlegungen zum einen auf die Zeit, die ein Kunde bereit ist an der Abholhaltestelle auf ein SAV zu warten. Zum anderen besteht im Rahmen der Fahrzeugführung die Möglichkeit sensible Zonen (z. B. Naturschutzgebiete, Baustellen), welche sowohl in ländlichen als auch in urbanen Räumen existieren, vor dem Hintergrund ökonomischer, ökologischer oder sozialer Zielaspekte zu umfahren (Narayanan et al. 2020). In Bezug auf die Kundenwartezeiten legen qualitative Erhebungen nahe (z. B. Martinez und Viegas 2017), dass der Kunde bereit ist, an der Abholhaltestelle etwa fünf Minuten auf ein SAV zu warten. Die herangezogenen Wartezeiten sind variabel und können dynamisch in den Simulationsbausteinen angepasst werden. Hierbei ist jedoch immer die Fahrtzeit zwischen Depot und der am weitest entfernten Haltestelle zu beachten, die über einen bidirektionalen A* Punkt-zu-Punkt Distanzalgorithmus automatisch als Mindestanforderung an die maximale Wartezeit bestimmt wird.

Das hier vorgestellte Konzept sieht zentrale Depots vor, weshalb die Flotte zu Beginn der Simulation dort initialisiert wird (Ma et al. 2017). Die Depots selbst werden als Black-Box mit uneingeschränkter Verlade- und Lagerkapazität für Sendungselemente (z. B. Pakete) modelliert. Die Anzahl (Integer) der zentralen Depots sowie deren geografische Standorte (Koordinaten) können vom Anwender frei gewählt werden. Während des operativen Betriebes wird von einer Zielhaltestelle direkt die Starthaltestelle des nächsten Auftrages angefahren. Zweimal täglich und außerhalb der Hauptverkehrszeiten muss für Reinigung, Reparatur, Wartung und Beladung mit

Paketen für Logistikaufträge in das Depot zurückgekehrt werden (Hyland und Mahmassani 2017). Fahrzeuge, die sich in einem vom Anwender spezifizierten Radius vom Depot befinden, können optional innerhalb dieses Radius immer über das Depot geleitet werden, um weitere Logistiksendungen aufzunehmen und im weiteren Fahrverlauf zuzustellen. Die Flottengröße, also die im System des Szenarios befindliche Fahrzeuganzahl kann dynamisch über das Serviceversprechen, das gegenüber den Kunden gegeben wird (z. B. Wartezeiten), bestimmt werden oder vom Anwender für die Durchführung verschiedener Was-wäre-wenn-Analysen und Optimierungsansätze parametrisch gewählt werden.

Damit kombinierter Transport gewährleistet wird, erfolgt der Transport von Mobilitäts- und Logistikaufträgen simultan im selben SAV (Beirigo et al. 2018). Sowohl die Passagierkapazität (Mobilitätsaufträge) als auch Lagerkapazität (Logistikaufträge) der Fahrzeuge hängt primär von der Fahrzeugkonzeption und der Beschaffenheit der zu transportierenden Güter ab (Narayanan et al. 2020). Um ein entsprechend vielfältiges Einsatzspektrum zu gewährleisten, können dieses Parameter innerhalb der Bausteine frei parametrisiert werden. Die Zuordnung von Mobilitäts- und Logistikaufträgen zu einem SAV lässt sich bei Modellen, die unter Verwendung des Bausteinkastens implementiert werden, auf drei verschiedene Weisen initiieren (vgl., Narayanan et al. 2020). So kann die Neuordnung je nach Untersuchungskontext und Anwenderpräferenz gemäß eines k-Nearest-Neighbour Algorithmus' erfolgen (vgl., Dudani 1976), bei dem die Fahrzeuge basierend auf einer Distanzmatrix (wird bei Initialisierung der Simulation berechnet) immer dem nächstgelegenen Mobilitätsauftrag zugeordnet werden. Die zweite Auswahlmöglichkeit impliziert die optimierte Zuordnung auf Basis eines Reinforcement Learning Algorithmus' (vgl., Zhang et al. 2020), bei dem die Fahrzeugagenten im Routingnetzwerk vor jedem Simulationsdurchlauf auf Basis der prognostizierten Nachfrage auf die Minimierung von Strecken und Wartezeiten trainiert werden. Die dritte Möglichkeit bietet ein lokales meta-heuristisches Suchverfahren (vgl., Groër et al. 2010), bei dem je nach Anwenderspezifikation aus einer Reihe von generierten Zuordnungslösungen die mit der (1) kürzesten Fahrtzeit oder (2) kürzesten Fahrtstrecke für den entsprechenden Auftrag gewählt wird. Angesichts stark variierender Effizienz verschiedener Routinglösungen unter veränderten Kontextbedingungen (z. B. Hyland und Mahmassani 2017) und im Sinne einer guten Übertragbarkeit des Bausteinkastens, werden dem Anwender hier drei unterschiedliche Routingansätze zur Auswahl gestellt. Neu- oder Umverteilung der SAV im Untersuchungsgebiet wird konzeptionell dann notwendig, wenn sich aufgrund unterschiedlicher Nachfrageschwerpunkte die Fahrzeuge in bestimmten Gebieten sammeln und es dadurch unmöglich wird, Mobilitätsaufträge innerhalb der vorgegebenen Servicezeiten zu bedienen (Brendel et al. 2017). Die Neuverteilung der SAV erhöht im Allgemeinen zwar die insgesamt im System gefahrene Strecke, verbessert jedoch auch die Reaktionszeiten auf Mobilitätsaufträge und somit die Wirtschaftlichkeit des Konzepts (Fagnant und Kockelman 2014). Der hier vorgestellte Bausteinkasten sieht eine Umverteilung der SAV dann vor, wenn sich die Fahrzeuge entweder an einer Haltestelle befinden, an der nach den Nachfragedaten innerhalb eines Zeitraums von 60 Minuten keine weiteren Aufträge zu erwarten sind oder wenn es im Untersuchungsgebiet Haltestellen gibt, die aufgrund der ungleichmäßigen Fahrzeugverteilung nicht innerhalb der angestrebten Servicezeiten bedient werden können. Die Umverteilung ist für das Konzept auch dahingehend von hoher Relevanz

als das sie auch dafür genutzt wird SAV für die Aufnahme von Paketen zum zentralen Depot zu führen.

Zur elektrischen Beladung der Fahrzeuge legen bestehende Untersuchungen nahe (z. B. Brendel et al. 2017; Narayanan et al. 2020), eine Systematik zu nutzen, bei der die Aufladung direkt an einer Haltestelle oder einem Depot stattfindet, sobald die Wartezeit des Fahrzeuges dort einen Zeitraum von fünf Minuten übersteigt. Die Ladeinfrastruktur wird in den Bausteinen in beiden Fällen ebenfalls als Black-Box modelliert. Der Ladezyklus beginnt nach Ablauf der fünf Minuten und endet sobald ein neuer Auftrag eingeht. Fällt der Ladestand eines SAV nach Beendigung eines Mobilitätsauftrages unter einem vom Simulationnutzer festgesetzten Prozentsatz (Min. 10 %, Max. 80 %), wird das Fahrzeug als nicht mehr verfügbar angesehen und kehrt zum nächstgelegenen Ladepunkt zurück. Ausgehend von bisherigen Forschungsarbeiten beträgt die Reichweite der elektrifizierten SAV 175 Kilometer mit einer durchschnittlichen Wiederaufladezeit von 30 Minuten (Chen et al. 2016; Martinez und Viegas 2017). Der letzte Bestandteil des Transportmodells sind die Halte- bzw. Auslieferungspunkte für Personen und Logistiksendungen. Diese können je nach Anwendungskontext frei vom Anwender verteilt werden (Moerner 2019). Ergänzend wird davon ausgegangen, dass zusätzlich zur Infrastruktur für die elektrische Beladung der Flotte auch geeignete Vorrichtungen für Entladung und Zwischenlagerung der Pakete sowie für einen komfortablen Zu- und Ausstieg der Passagiere vorhanden sind. Diese Haltestelle können vom Anwender selbst über die entsprechenden Koordinaten spezifiziert oder automatisch vom Simulationssystem festgelegt werden. In letzterem Fall muss der Anwender lediglich die gewünschte Anzahl an Haltestellen vorgeben, sodass diese bei der Simulationsinitialisierung über das Streckennetz des Untersuchungsgebietes gleichverteilt wird. Die parametrisierte Mobilitätsnachfrage wird über die Entfernung der Haltestellen zu den individuellen Nachfragelokationen (distanzbasierter A* Punkt-zu-Punkt Algorithmus) in individuelle Nachfrageerwartungswerte für jede Haltestelle überführt und dient als Grundlage für die Routing- und Umverteilungsprozesse.

4 Umsetzung und Demonstration

Zur Umsetzung des bausteinbasierten Lösungsansatzes wurde in Anlehnung an die Systematik des BPMN 2.0-Modells (Allweyer 2016) zunächst ein Gesamtprozess (s. Abb. 1) konzipiert, der die zuvor genannten Konzeptionsbestandteile modelliert. Als Kernelemente dieses Gesamtprozesses wurden drei interdependente Subprozesse identifiziert, die als Bausteine dargestellt und in dem Bausteinkasten zusammengefasst werden. Zu den wichtigsten Eigenschaften eines Bausteins zählt, dass er übertragbar bzw. wiederverwendbar sein muss (Verbraeck und Valentin 2008). Daher wurden die Subprozesse so konzipiert, dass sie sowohl unter gleichen Rahmenbedingungen als auch unter veränderten Annahmen auf andere kombinierte Transportmodelle übertragen werden können.

Wie in Abb. 1 dargestellt beginnt der Gesamtprozess mit der Initialisierung der SAV-Flotte am Depot. Anschließend werden Mobilitäts- und Logistikaufträge generiert, die auf der Nachfrage der einzelnen Haltestellen basieren. Dieser Nachfrage wiederum liegen nutzerspezifizierte Mobilitätsdaten zugrunde, aus denen Erwartungswerte gebildet und über eine Liste in das System eingespeist werden. Nächster Prozessschritt ist die Fahrzeugdisposition sowie Zuteilung der Aufträge auf die SAV-

Flotte durch die Haltestellen. Handelt es sich bei einem zugeteilten Auftrag um einen Mobilitätsauftrag, wird das SAV zur Ausgangshaltestelle dieses Auftrags geroutet und der Auftrag bearbeitet. Das Fahrzeug nimmt also an der Starthaltestelle des Auftrags Passagiere auf und bringt diese zur gewünschten Zielhaltestelle. Liegen in dem Quadranten, in der sich das SAV zu diesem Zeitpunkt befindet, keine weiteren Aufträge vor, erfolgt eine Umverteilung der Fahrzeuge in Quadranten mit höherer Nachfrage. Routing und Auftragsbearbeitung finden dann von Neuem statt.

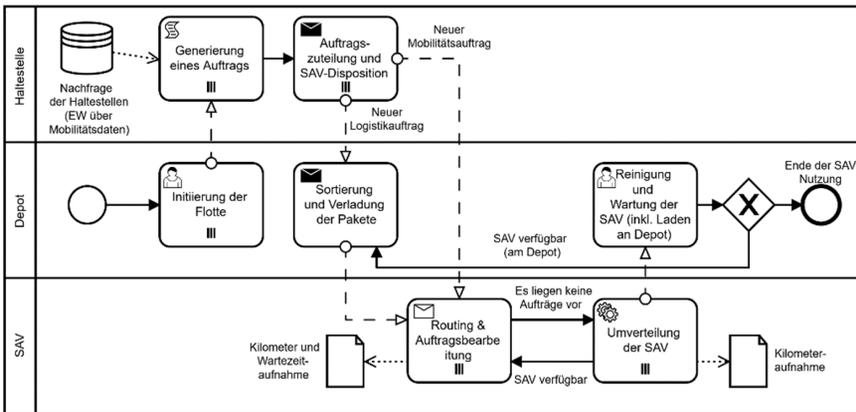


Abbildung 1: Darstellung des Gesamtprozesses in Anlehnung an BPMN 2.0

Zweimal am Tag wird ein SAV nicht umverteilt (zeitgesteuerte Bedingung mit den Parametern Uhrzeit, Auftragslage, Zeit seit letzter Wartung und verbliebenes Sendungsvolumen), sondern stattdessen für Reinigung, Wartung, elektrische Beladung und Sendungsaufnahme zum Depot zurückgeführt. Hier endet die SAV-Nutzung oder es erfolgt die Sortierung und Verladung der Pakete mit anschließender Auslieferung im Rahmen der Auftragsbearbeitung. Handelt es sich bei der anfänglichen Auftragszuteilung um einen Auftrag vom Typ "Logistik", setzt die Sortierung und Verladung der Pakete schon dann ein und die Fracht wird anschließend direkt vom Depot zum Bestimmungsort verbracht. Der Gesamtprozess verfügt über variable Eingangsparameter. Diese gelten über die Grenzen eines Bausteins hinaus und können vom Anwender angepasst werden. Dazu zählen die geografische Spezifikation des Untersuchungsgebietes mittels eines Open Streetmap Netzwerkes, die Anzahl und Größe der Quadranten, in die dieses Gebiet unterteilt ist, die Zahl und ggf. Lokation der Depots und Haltestellen sowie die Kapazität und Geschwindigkeit der SAV. Ebenfalls bausteinübergreifend wirken sich die logistik- und mobilitätsbezogene Nachfrage und die maximalen Kundenwartezeiten aus. Wie der Abb. 1 zu entnehmen ist, beinhaltet der Bausteinkasten sieben Bausteine, von denen nachfolgend beispielhaft die Konzeption und Implementierung des Bausteins *Routing und Auftragsbearbeitung* erläutert wird. Der Subprozess *Routing und Auftragsbearbeitung* beginnt mit der Bedingung, dass das SAV verfügbar ist. Anschließend erfolgt die Prüfung, ob sich das SAV am zentralen Depot oder an einer der Haltestellen befindet. Ist ersteres der Fall, werden Pakete eingeladen, deren Zielhaltestellen alle im selben Quadranten liegen. Die Logistikaufträge werden dabei durch den Erwartungswert (kummulierte Logistknachfrage der Haushalte in

Reichweite einer Haltestelle auf Basis der Daten eines Kourrier-Express-Paket-Dienstleisters) der einzelnen Haltestellen in eine Liste überführt. Anschließend wartet das SAV auf eingehende Mobilitätsaufträge. Befindet sich das Fahrzeug hingegen an einer der Haltestellen, wartet es direkt auf den Eingang von Mobilitätsaufträgen. Sind keine Mobilitätsaufträge vorhanden, wird geprüft, ob der aktuelle Zeitpunkt innerhalb eines Zeitfensters mit hohem Mobilitätsauftragsvolumen liegt, wobei der Schwellenwert für den Erwartungswert zur Abgrenzung von Zeitfenstern mit hohem Volumen über einen Parameter vom Nutzer bestimmt werden kann. Liegt ein Zeitfenster mit Mobilitätsbewertungswerten über dem Schwellenwert vor, wird keine Verteilung von Logistikaufträgen vorgenommen, sondern stattdessen weiter auf Mobilitätsaufträge gewartet. Liegt der aktuelle Zeitpunkt jedoch außerhalb eines solchen Zeitfensters, wird geprüft, ob Logistikaufträge vorliegen. Ist dem nicht so, endet der Subprozess. Sind hingegen Logistikaufträge vorhanden, wird das Routingnetzwerk aktualisiert und auf vom Anwender definierte sensible Zonen hin untersucht. Während des darauffolgenden Transports der Pakete zur Zielhaltestelle werden diese sensiblen Zonen dann aufgrund einer dynamischen Routinganpassung umfahren. Ist das Paket ausgeliefert, wartet das SAV erneut auf eingehende Mobilitätsaufträge. Sind während das SAV an Haltestelle oder Depot wartet Mobilitätsaufträge vorhanden bzw. gehen diese ein, wird ebenfalls eine Aktualisierung des Routingnetzwerks hinsichtlich sensibler Zonen vorgenommen. Bei der anschließenden Fahrt zur Starthaltestelle des Mobilitätsauftrags werden diese Zonen dann umgangen. An der nächsten Haltestelle angekommen kommt es dann zu Zu- bzw. Ausstieg der Passagiere. Hat das SAV zusätzlich zu Passagieren auch Sendungen geladen, die für die Haltestelle bestimmt sind an der sich das SAV zu diesem Zeitpunkt befindet, werden auch diese ausgeladen. Nachdem dies geschehen ist, oder für den Fall, dass sich keine Sendungen an Bord befinden, die die aktuelle Haltestelle zum Bestimmungsort haben, wird in den umliegenden Quadranten nach weiteren Mobilitätsaufträgen gesucht, die den gleichen Quadranten wie der soeben zugestiegene Mobilitätsauftrag zum Ziel haben. Sind entsprechende Aufträge vorhanden, wird erneut das Routingnetzwerk aktualisiert und die beschriebenen Prozessschritte wiederholen sich. Liegen in näherer Umgebung hingegen keine Mobilitätsaufträge vor, wird geprüft, ob an Bord des SAV ein weiterer Mobilitätsauftrag befindlich ist. Falls dem so ist, fährt das SAV zur Zielhaltestelle dieses Anschlussauftrags. Ist dort kein Mobilitätsauftrag vorhanden, erfolgt eine Prüfung, ob der aktuelle Zeitpunkt unmittelbar nach einem Zeitfenster mit hohem Mobilitätsaufkommen liegt und ob das Depot bisher maximal einmal pro Tag aufgesucht wurde. Treffen beide Bedingungen zu, ist das SAV nun nicht mehr verfügbar und fährt zum Depot, wo der Subprozess endet. War das SAV bisher jedoch bereits zweimal am selben Tag am Depot oder liegt der aktuelle Zeitpunkt unmittelbar vor oder in einem Hochfrequenzzeitfenster, endet der Subprozess ohne die Fahrt zum Depot und der Subprozess *Umverteilung der SAV* beginnt. Abb. 2 zeigt das Implementierungsbeispiel für den Baustein *Routing und Auftragsbearbeitung* in der Simulationssoftware AnyLogic. Die Bausteine wurden mittels einer agentenbasierten Modellierungsmethodik (ABM) im Rahmen von Verhaltensdiagrammen umgesetzt und weisen an den entsprechenden Schnittstellen wechselseitige Interdependenzen zu anderen relevanten Bausteinen auf (z. B. wird über den in Abb. 2 erkennbaren Zustand „starteUmverteilung“ per Nachrichtenübermittlung das Verhaltensdiagramm des Umverteilungsbausteins initiiert). Die Bausteine wurden in Form Agentenklassen als wiederverwendbare Objektbibliotheken in AnyLogic gespeichert.

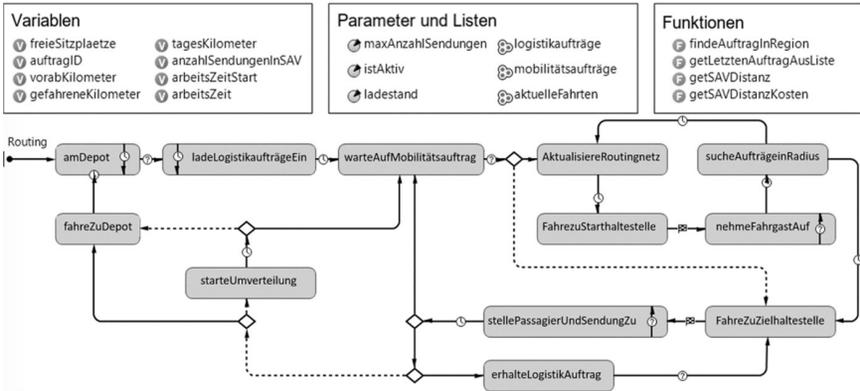


Abbildung 2: Beispiel eines Bausteins für den Auftragsprozess in AnyLogic

Zur Demonstration und Evaluation des Bausteinkastens wurden die Bausteine der vorgestellten Logik des Gesamtprozesses folgend für die Entwicklung eines Simulationsmodells für kombinierte Verkehre im Raum Hannover entwickelt. Einen exemplarischen Ausschnitt aus dem entwickelten Simulationsmodell zeigt Abb. 3.



Abbildung 3: Beispielmodell für kombinierten Verkehr im Raum Hannover

Der angewendeten Modellierungslogik der Simulationsbausteine folgend wurde die Simulation als agentenbasiertes Modell umgesetzt. Auf diese Weisen können die wechselseitigen Abhängigkeiten der einzelnen Systementitäten (z. B. SAV, Passagiere und Sendungen) abgebildet und imitiert werden, wodurch letztendlich die globale Systemdynamik mithilfe von Netzwerkeffekten reproduziert wird, die sich aus den modellierten Agenteninteraktionen ergeben. Dieser Ansatz ist besonders geeignet, um Verhaltenseinflüsse und Komponenten mit geografischem Bezug widerzuspiegeln (Heppenstall et al. 2012). Der Hauptvorteil von ABM im Rahmen unserer Studie ist seine immanente Fähigkeit, emergent auftkommende Verhaltensweisen zu erfassen, die von den Eigenschaften der einzelnen Agenten entkoppelt sind und vom Modellierer noch nicht in ihrer Gänze durchdrungen wurden. Dies sichert die Übertragbarkeit der Bausteine.

5 Fazit und Ausblick

Die hier vorgestellten Bausteine sowie deren beispielhafte Implementierung in AnyLogic dienen als ein Ansatzpunkt für die simulationsgestützte Darstellung kombinierter, autonomer Transportsysteme mittels geteilter autonomer Fahrzeuge. Es wurde sowohl ein konzeptioneller Rahmen als auch ein Vorschlag für die Simulation solcher Modelle entwickelt. Hiervon profitieren potenziell nicht nur Forschungs- und Entwicklungsvorhaben, sondern auch Anwendung und Implementierung in der planerischen Praxis. Aufbauend auf den vorliegenden Ergebnissen sind weitere Untersuchungen im Bereich kombinierter Verkehre vorgesehen, die die konzeptionelle und simulationsgestützte Betrachtung des Transportsystems erweitern und bisher nur oberflächlich betrachtete Aspekte vertiefen und kontextabhängig quantifizieren.

Danksagung

Die Autoren bedanken sich herzlich beim Bundesministerium für Verkehr und Infrastruktur (BMVI) für die finanzielle Unterstützung und beim Projektträger TÜV Rheinland für die Betreuung des Projektes „Kombinom“ unter dem Förderkennzeichen FKZ 19F1072B.

Literatur

- Allweyer, T.: BPMN 2.0: Introduction to the Standard for Business Process Modelling. Norderstedt: Books on Demand 2016.
- Auf der Landwehr, M.; Trott, M.; Viebahn, C. von: Waste of Time and Money? Constructing an Applicability Framework for Organizational Use of Simulation Studies and Digital Twins. In: George, J.F.; Paul, S.; De', R.; Karahanna, E.; Sarker, S.; Oestreicher-Singer, G. (Hrsg.): Proceedings of the 41st International Conference on Information Systems (ICIS), Virtual Conference, 13.-16. Dezember 2020.
- Beirigo, B.; Schulte, F.; Negenborn, R.: Integrating People and Freight Transportation Using Shared Autonomous Vehicles with Compartments. IFAC-PapersOnLine 51 (2018) 9, S. 392–397.
- Brendel, A.B.; Nastjuk, I.; Lichtenberg, S.; Kolbe, L.M.: Adapting Carsharing Vehicle Relocation Strategies for Shared Autonomous Electric Vehicle Services. In: Kim, Y.J.; Agarwal, R.; Lee, J.K. (Hrsg.): Proceedings of the 38th International Conference on Information Systems (ICIS), Seoul (South Korea), 10.-13. Dezember 2017.
- Chen, T.D.; Kockelman, K.M.; Hanna, J.P.: Operations of a shared, autonomous, electric vehicle fleet: Implications of vehicle & charging infrastructure decisions. Transportation Research Part A: Policy and Practice 94 (2016), S. 243-254.
- Clausen, U.; Brueggenolte, M.; Kirberg, M.; Besenfelder, C.; Poeting, M.; Gueller, M.: Agent-based simulation in logistics and supply chain research. In: Clausen, U.; Langkau, S.; Kreuz, F. (Hrsg.): Interdisciplinary Conference on Production, Logistics and Traffic: Literature review and analysis: Springer 2019, S. 45–59.
- Dudani, S. A.: The distance-weighted k-nearest-neighbor rule. IEEE Transactions on Systems, Man, and Cybernetics 4 (1976), S. 325-327.

- Fagnant, D.J.; Kockelman, K.M.: The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios. *Transportation Research Part C: Emerging Technologies* 40 (2014), S. 1–13.
- Gröer, C.; Golden, B.; Wasil, E.: A library of local search heuristics for the vehicle routing problem. *Mathematical Programming Computation* 2 (2010), S. 79–101.
- Heppenstall, A.; Crooks, A.; See, L. M.; Batty, M.: *Agent-based models of geographical systems*. Dordrecht: Springer Netherlands; 2012. 760 p.
- Hyland, M. F.; Mahmassani, H.S.: Taxonomy of shared autonomous vehicle fleet management problems to inform future transportation mobility. *Transportation Research Record: Journal of the Transportation Research Board* 2653 (2017), S. 26–34.
- Ma, J.; Li, X.; Zhou, F.; Hao, W.: Designing optimal autonomous vehicle sharing and reservation systems: A linear programming approach. *Transportation Research Part C: Emerging Technologies* 84 (2017), S. 124–141.
- Martinez, L. M.; Viegas, J. M.: Assessing the impacts of deploying a shared self-driving urban mobility system: An agent-based model applied to the city of Lisbon, Portugal. *International Journal of Transportation Science and Technology* 6 (2017) 1, S. 13–27.
- Meyer, G.; Shaheen, S.: *Disrupting Mobility: Impacts of Sharing Economy and Innovative Transportation on Cities*: Springer International Publishing 2017.
- Moerner, M. von: Demand-oriented mobility solutions for rural areas using autonomous vehicles. In: Coppola, P.; Esztergár-Kiss, D. (Hrsg.): *Autonomous Vehicles and Future Mobility*. Amsterdam, Netherlands, Oxford, United Kingdom, Cambridge, United States: Elsevier 2019, S. 43–56.
- Monios, J.; Bergqvist, R.: The transport geography of electric and autonomous vehicles in road freight networks. *Journal of Transport Geography* 80 (2019).
- Narayanan, S.; Chaniotakis, E.; Antoniou, C.: Shared autonomous vehicle services: A comprehensive review. *Transportation Research Part C: Emerging Technologies* 111 (2020), S. 255–293.
- Rodrigue, J.-P.; Comtois, C.; Slack, B.: *The geography of transport systems*. London, New York: Routledge 2020.
- Verbraeck, A.; Saanen, Y.; Stojanovic, Z.; Valentin, E.; van der Meer, K.; Meijer, A.; Shishkov, B., 2002: What are building blocks? In: Verbraeck, A.; Ajantha Dahanayake (Hrsg.): *Building Blocks For Effective Telematics Application Development And Evaluation*. Delft University of Technology, S. 8–21.
- Verbraeck, A.; Valentin, E.C.: Design guidelines for simulation building blocks. In: 2008 Winter Simulation Conference, Miami, FL, USA, 2008, S. 923–932.
- Wenzel, S.: Simulation logistischer Systeme. In: Tempelmeier, H. (Hrsg.): *Modellierung logistischer Systeme*. Berlin Heidelberg: Springer Vieweg 2018, S. 1–34.
- Yu, S.; Puchinger, J.; Sun, S.: Two-echelon urban deliveries using autonomous vehicles. *Transportation Research Part E: Logistics and Transportation Review* 141 (2020), <https://doi.org/10.1016/j.tre.2020.102018>.
- Zhang, K.; He, F.; Zhang, Z.; Lin, X.; Li, M.: Multi-vehicle routing problems with soft time windows: A multi-agent reinforcement learning approach. *Transportation Research Part C: Emerging Technologies* 121 (2020), <https://doi.org/10.1016/j.tre.2020.102861>.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Direktumschlag an der Kaikante – Eine Machbarkeitsstudie für hafeninterne Containertransporte

Direct Container Handling at Waterside - A Feasibility Study for Inter-Terminal Transports

Nicole Nellen, Ann-Kathrin Lange, Carlos Jahn, TU Hamburg, Hamburg
(Germany), nicole.nellen@tuhh.de, ann-kathrin.lange@tuhh.de,
carlos.jahn@tuhh.de

Abstract: The increasing ship size growth and land-based capacity restrictions at container terminals increase the need for optimization approaches in the process design of container ports. Several strategies can be found in the literature, ranging from the optimization of terminal processes to the introduction of truck appointment systems and the direct handling of transshipment containers. In this paper, a possibility for process redesign of waterside container handling to external trucks is analysed. For this purpose, the impact of direct waterside container handling on the terminal is analysed using a discrete event-based simulation model. The study shows that the approach for new process design has potential, although implementation is not easy.

1 Containerterminals und hafeninterne Transporte

Seehäfen bilden in der maritimen Lieferkette die Schnittstelle zwischen verschiedenen Verkehrsträgern. Im Jahr 2019 umfasste das Volumen des globalen containerisierten Handels 152 Millionen TEU (Twenty-foot Equivalent Units). Dies entspricht einer Verdreifachung verglichen mit 1997 (UNCTAD, 2020). Darüber hinaus steigt die Komplexität maritimer Transportketten durch das Größenwachstum der Containerschiffe. Hatten Containerschiffe um die Jahrtausendwende eine Kapazität von 8.000 TEU, fassten die größten Schiffe in 2020 knapp 24.000 TEU (Allianz, 2020). Somit laufen immer häufiger Containerschiffe mit größeren Volumina die Terminals an, während die Summe aller Schifffanläufe reduziert wird. Folglich entstehen höhere Lastspitzen im Containerumschlag einschließlich der landseitigen (Lkw-) Abfertigung (Ramírez-Nafarrate et al., 2017).

Um dieser Entwicklung begegnen zu können, ist eine effiziente Gestaltung der Prozesse von Containerterminals notwendig. Hierbei ist zu beachten, dass jeder Container im Terminal bei den herkömmlichen Prozessen mindestens dreimal, meist

aber häufiger, durch internes Equipment zwischen verschiedenen Verkehrsträgern und Abstellflächen umgeschlagen werden muss. Dies resultiert in einem hohen Koordinationsaufwand für die Terminals und setzt gleichzeitig die Verfügbarkeit von Handlingequipment voraus. Weiterhin haben insbesondere stadtnahe Terminals stark eingeschränkte Erweiterungsflächen, was Maßnahmen zur Reduzierung der Menge an Containern im Yard bei einer Beibehaltung der Umschlagleistung besonders attraktiv macht. Eine Möglichkeit der Reduzierung des Handlingaufwands und des Flächenbedarfs am Terminal ist der direkte Umschlag von Containern an der Kaikante auf externe Lkw. Hierbei entfallen sowohl der terminalinterne Transport von der Kaikante zum Yard mit dem zugehörigen Umschlag als auch die sich daran anschließenden Ein- und Auslagervorgänge. Da der direkte Umschlag einer großen Menge an Containern an der Kaikante umfangreiche Prozess- und Layoutanpassungen für die Terminals bedeuten würde, ist es angebracht, zunächst nur die Auswirkungen auf einen Teil der Transporte des Terminals zu untersuchen und dort das Potenzial zur Effizienzsteigerung zu bewerten. Dafür eignen sich insbesondere die hafeneninternen Transporte zwischen Terminals, Leercontainerdepots, Packstationen und dem Zoll, die von einer begrenzten Menge an Fuhrunternehmen und zugehörigen Fahrern durchgeführt werden und im Verhältnis zur Transportstrecke eine sehr hohe Anzahl an Umschlägen im Hafen auf sich vereinen. Aufgrund der hohen Kosten und Risiken eignet sich dies aber nicht für die sofortige Erprobung im Echtbetrieb und auch der Einsatz von mathematischer Optimierung ist aufgrund hoher Stochastizität nicht sinnvoll. Viel mehr bietet sich die Untersuchung mittels Simulation an. Ziel dieser Studie ist es dementsprechend, mithilfe eines diskreten, ereignisorientierten Simulationsmodells die Auswirkungen eines direkten Containerumschlags an der Kaikante auf die hafeneninternen Containertransporte zu untersuchen.

Dafür werden in Kapitel 2 zunächst aktuelle Forschungsfelder zu Optimierungsansätzen auf Containerterminals und im Bereich der hafeneninternen Transporte dargestellt. In Kapitel 3 folgt ein Überblick über den Aufbau des Simulationsmodells. Weiterhin werden notwendige Annahmen zur Modellierung erörtert. Eine Darstellung und Auswertung der durchgeführten Experimente erfolgt in Kapitel 4. Der Beitrag schließt in Kapitel 5 mit einem Fazit und Ausblick ab.

2 Stand der Forschung

In der Literatur lassen sich zahlreiche Ansätze zur Prozessoptimierung in Containerhäfen finden. Der Fokus dieser Studien liegt zumeist auf einer isolierten Betrachtung der Knoten oder Kanten im Transportnetzwerk. Die wichtigsten Knoten bei hafeneninternen Containertransporten sind die Containerterminals, welche grob in drei Bereiche aufgeteilt werden können. Dies sind die Abfertigung der Schiffe an der Kaikante, auch wasserseitige Abfertigung genannt, die Einlagerungen der Container im Yard und die landseitige Abfertigung, bei der die Container von und auf Lkw und Bahnen umgeschlagen werden. Für die Umschlagprozesse kann unterschiedliches Equipment eingesetzt werden. Die Zusammenstellung des Equipments auf einem Terminal wird als Betriebssystem bezeichnet. Bei größeren Containerterminals werden fast immer sogenannte Containerbrücken für die wasserseitige Abfertigung verwendet. Dies sind schienengebundene Portalkrane mit einem Ausleger, der die gesamte Breite des abzufertigenden Schiffs überspannt. Die Container werden anschließend mit Fahrzeugen des Horizontaltransports zum Containeryard gebracht.

Für den Umschlag im Yard ist eine weit verbreitete Lösung der Einsatz von gummibereiften Portalkranen, zumeist bezeichnet als Rubber Tyred Gantry Cranes (RTG). Diese nehmen die Container von den Lkw, Bahnen oder dem Horizontaltransport ab, lagern sie ein und vice versa. Häufig verwendete Alternativen dazu sind Rail Mounted Gantry Cranes (RMG) oder Straddle Carrier. Die wissenschaftlichen Untersuchungen bei Terminals reichen von langfristigen Entscheidungen bzgl. der Gestaltung des Terminal-Layouts oder des Betriebssystems bis hin zu operativen Entscheidungsproblemen. Kizilay und Eliyi (2020) und Schwientek et al. (2017) geben u. a. einen aktuellen Überblick über Entscheidungsprobleme auf Containerterminals. Bei der Betrachtung von straßenseitigen Transporten ist in den letzten Jahren insbesondere die Gestaltung und Einbindung von sogenannten Truck Appointment Systemen (TAS), welche zur Steuerung von Lkw-Ankünften an Terminals eingeführt werden, in das Interesse der Wissenschaft gerückt. Eine Übersicht zu den aktuellen Publikationen im Bereich von TAS ist u. a. in Huiyun et al. (2018) und Lange et al. (2017) dargestellt. Durch die Nutzung von TAS wird die Planbarkeit der Lkw-Ankünfte am Terminal deutlich erhöht, da die Lkw zumeist nur ein Zeitfenster von einer Stunde haben, in dem sie am Terminal ankommen dürfen. Abgesehen von TAS konzentrieren sich einige Autoren auf den Vergleich verschiedener Transportsysteme bei Durchführung von Containertransporten im Hafen. Beispielfhaft können hier Duinkerken et al. (2007), Gharehgozli et al. (2017) und Nellen et al. (2020) genannt werden. Da Container innerhalb eines Hafens nicht nur an dem Terminal, den sie wasserseitig erreichen, benötigt werden, müssen diese häufig zwischen den logistischen Knoten eines Hafens transportiert werden. Diese Knoten können neben anderen Terminals insbesondere auch Leercontainerdepots, Packstationen und der Zoll sein. Beispielsweise stellen Heilig und Voß (2017) einen aktuellen Stand der Forschung zu hafeninternen Containertransporten dar und identifizieren darüber hinaus zukünftige Forschungsfragen.

Weiterhin existieren Ansätze für den direkten Umschlag von Containern zwischen Containerschiffen (direktes Transshipment). Zeng et al. (2017) untersuchten das direkte Transshipment mittels Heuristiken im Hinblick auf die Betriebskosten von Yard Trucks und RTG. Die Autoren kamen zu dem Ergebnis, dass durch einen direkten Containerumschlag zwischen Schiffen die Betriebskosten gesenkt und gleichzeitig die Umschlageffizienz verbessert werden können. Auch Liang et al. (2012) haben den teilweise direkten Umschlag zwischen Schiffen betrachtet. Der Fokus der Studie lag auf dem Problem der Liegeplatzzuweisung unter Berücksichtigung von direkten Containerumschlägen zwischen Großcontainer- und Feederschiffen. Zur Untersuchung der Problemstellung wurde ein hybrider mehrstufiger genetischer Algorithmus angewendet.

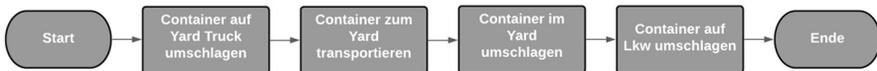
Den Autoren ist kein Ansatz eines direkten Umschlags von Containern an der Kaikante bekannt, die hafenintern weitertransportiert werden. Deshalb wird dies für Importcontainer, also die Container, die per Seeschiff am Terminal ankommen, in der in Kapitel 3 beschriebenen Simulationsstudie grundlegend untersucht und in Kapitel 4 ausgewertet. Um einen Direktumschlag zu ermöglichen, ist eine sehr pünktliche Ankunft der Lkw am Terminal notwendig, damit die Abfertigung des Schiffes nicht verzögert wird. Da dies bisher nach Kenntnis der Autoren nirgendwo umgesetzt ist, können die Untersuchungen zu TAS als Grundlage verwendet werden, um geeignete Annahmen für das Simulationsmodell zu generieren.

3 Aufbau der Simulationsstudie

Um Potenziale durch die Minimierung von Containerumschlägen bei hafeninternen Transporten an einem Terminal zu untersuchen, wird ein ereignisorientiertes diskretes Simulationsmodell eines Containerhafens in Tecnomatix Plant Simulation (Version 15.1) erstellt. Dabei bildet das Modell den Hafen vereinfacht mit einem Terminal und vier weiteren Akteuren ab. Die Prozesse am Terminal sind detailreich modelliert und werden durch das Layout sowie das Betriebssystem (RTG mit Yard Trucks) beeinflusst. Es werden die wasserseitige Abfertigung eines Containerschiffs, Ein- und Auslagervorgänge im Yard, der Horizontaltransport zwischen den Funktionsbereichen sowie die landseitige Abfertigung externer Lkw dargestellt.

Um den Einfluss der Ablaufgestaltung beim Umschlag von hafenintern weiter zu transportierenden Containern auf den Terminal zu bestimmen, wird bei der Prozessgestaltung zwischen zwei unterschiedlichen Umschlagprozessen am Terminal differenziert. Herkömmlicher Weise werden alle wasserseitig am Terminal eintreffenden Container auf Equipment für den horizontalen Transport umgeschlagen und in den Containeryard transportiert. Hier findet ein weiterer Umschlag zur Zwischenlagerung des Containers statt. Dies wiederholt sich, wenn dieser ausgelagert und auf das Fahrzeug für den Weitertransport im Hafen oder in das Hinterland geladen wird (siehe Abbildung 1). Für den Direktumschlag an der Kaikante muss, wie in Abbildung 1 zu erkennen, zwischen Containern unterschieden werden, die im Hafengebiet bleiben oder die auf nachfolgende Schiffe umgeschlagen (Transshipmentcontainer) bzw. ins Hinterland transportiert werden. Container für den hafeninternen Transport sollen in diesem Prozess direkt an der Kaikante auf Lkw umgeschlagen werden und den Terminal verlassen. Somit entfallen mindestens zwei Umschläge. Alle anderen Container werden weiterhin herkömmlich umgeschlagen.

Herkömmliche Prozessgestaltung



Direktumschlag an der Kaikante

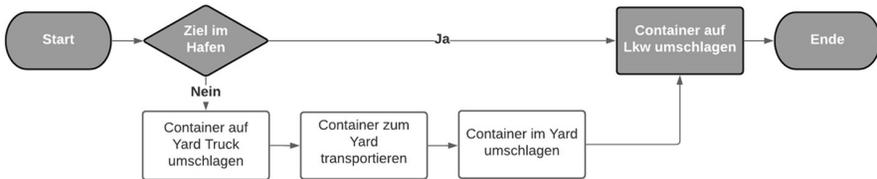


Abbildung 1: Ablaufgestaltung des wasserseitigen Containerumschlags

Da der Fokus der Studie auf den Auswirkungen der Prozessumgestaltung für den Terminal liegt, werden die Prozesse der weiteren Akteure, verglichen mit dem Terminal, abstrahiert modelliert. Der abgebildete Hafen orientiert sich hinsichtlich der Entfernungen zwischen den Akteuren und dem Umschlagvolumen am Hamburger Hafen. Dabei werden die Inputparameter entsprechend der Modellgröße skaliert. Fahrwege zwischen den Akteuren sind in der Simulation vereinfacht mithilfe einer Matrix der Wegzeiten abgebildet.

Zur Modellerstellung wurden folgende Annahmen getroffen:

- Der Terminal wird wasserseitig durch die Abfertigung eines Großschiffs sowie landseitig durch eintreffende Lkw belastet.
- Belastungsschwankungen durch wechselnde sowie parallele Ankünfte unterschiedlich großer Containerschiffe an der Kaikante werden nicht abgebildet.
- Es wird lediglich die Entladung der Container vom Schiff dargestellt.
- Gesetzlich vorgeschriebene Lenkzeiten der Lkw- Fahrer werden außer Acht gelassen.
- Das Modell unterscheidet nicht zwischen Containertypen. Es werden nur 40 ft Standardcontainer abgefertigt. Daraus ergibt sich, dass sowohl Yard Trucks als auch Lkw im Hafen nur einen Container zurzeit transportieren können.

Zu Beginn der Simulation werden Aufträge generiert, die im betrachteten Zeithorizont zwischen den Akteuren im Hafen zu transportieren sind. Hierbei wird jedem Auftrag zufällig eine Quelle und eine Senke zugewiesen. Weiterhin besitzt jeder Transportauftrag einen Zeitpunkt, bis zu dem er im Idealfall abgefertigt werden soll. Transportaufträge werden als mit Verspätung abgefertigt gewertet, wenn der Zeitpunkt der vorgesehenen Abfertigung mehr als 30 Minuten überschritten wird. Dabei erfolgt die Verteilung der Transportaufträge über den Tag entsprechend der Tagesganglinie eines großen Hamburger Terminals. So fallen in die frühen Morgen- und Abendstunden weniger Transportaufträge, während sich die im Hafen zu transportierenden Container um die Mittagszeit häufen.

Die Lkw im Hafen folgen vereinfacht keiner festen Tourenplanung. Nach dem erfolgreichen Abschluss eines Transports wird über eine zentrale Auftragsliste dem Lkw ein neuer Transportauftrag zugewiesen. Hierbei werden zum einen die Dringlichkeit der Abfertigung verfügbarer Transportaufträge und zum anderen die aktuelle Position des Lkw und die sich daraus ergebende Fahrzeit zum Start des möglichen neuen Transportauftrags beachtet. Bei zeitkritischen Transportaufträgen wird der Parameter des geplanten Abfertigungszeitpunkts höher priorisiert. In Experimenten, in denen ein Direktumschlag der Container auf Lkw an der Kaikante vorgesehen ist, werden an der Kaikante abzuholende Container gegenüber Transportaufträgen zwischen den anderen Akteuren im Hafen stark priorisiert, sodass an der Kaikante nach Möglichkeit keine Wartezeiten entstehen.

Am Containerterminal werden zu Beginn der Simulation die Container an den jeweiligen Containerbrücken dem Modell zugefügt. Das Simulationsmodell ist auf ein am Terminal anlegendes Containerschiff mit einer Länge von knapp 400 Metern sowie einer Stellplatzkapazität von rund 17.000 TEU ausgelegt. Davon sind am Terminal zwischen 8.000 und 10.000 TEU abzuladen, welche sich an Deck des Containerschiffs auf 18 Bays verteilen. Die mittlere wasserseitige Umschlag-

geschwindigkeit ergibt sich aus der Annahme, dass je Containerbrücke in der Stunde 24-30 Container umgeschlagen werden können. Stochastische Einflüsse auf die wasserseitige Umschlaggeschwindigkeit werden über eine Dreiecksverteilung abgebildet. Der Horizontaltransport zwischen Kaikante und Containeryard erfolgt durch Yard Trucks. Dabei sind jeder Containerbrücke vier Yard Trucks zugeordnet. Yard Trucks können im Modell mit einer Höchstgeschwindigkeit von 8,4 m/s über Fahrspuren auf dem Terminalgelände fahren. Für die RTG im Containeryard wird eine Produktivität von 15 Umschlägen in der Stunde angenommen. Daraus ergibt sich eine erwartete Abfertigungsdauer von 240 Sekunden pro Container. Abweichungen und Unregelmäßigkeiten im Prozess werden ebenfalls durch eine Dreiecksverteilung berücksichtigt. Diesbezüglich wird ein Minimalwert von 180 Sekunden und ein Maximalwert von 420 Sekunden bestimmt. Die Abfertigungszeit an allen weiteren Akteuren beträgt 30 bis 60 Minuten. Diese Annahmen basieren auf Experteninterviews im Hamburger Hafen.

Mithilfe von Simulationsexperimenten soll u. a. untersucht werden, welchen Einfluss die Umgestaltung der Prozesse am Terminal auf die Produktivität der Containerbrücken sowie den Yard und die Yard Trucks hat. Neben der Ablaufgestaltung bei der Abfertigung von Containern für den hafeninternen Transport wird der Anteil der Container im direkten Umschlag, der Lagerfüllstand im Yard sowie die Wahrscheinlichkeit, dass für einen Importcontainer bei seiner Abholung am Yard Block durch Lkw Umstapelvorgänge notwendig sind, variiert. Hierbei wird die Produktivität der Containerbrücken durch deren Umschläge pro Stunde ausgedrückt. Weiterhin werden die pro Lkw durchgeführten Transportaufträge sowie die Verspätungen der Lkw im Hafen ausgewertet. Zum Schluss wird überprüft, wie viele der geplanten Direktumschläge tatsächlich stattgefunden haben.

4 Auswertung der Ergebnisse

Die in diesem Kapitel dargestellten Ergebnisse basieren auf den zuvor vorgestellten Annahmen und dienen als Basis für eine geplante, hierauf aufbauende umfangreiche Simulationsstudie. Tabelle 1 veranschaulicht die durchgeführten Experimente mit den jeweils variierten Parametern.

Tabelle 1: Experimentplan mit zu variierten Parametern

Exp.	Direktums.	Direktums.[%]	Füllgrad Yard	Umstapelws. Yard	Σ Lkw
1-5	nein	–	50 % - 90 %	gering	75
6-10	nein	–	50 % - 90 %	hoch	75
11-15	ja	1 % - 5 %	60 %	gering	75
16-20	Ja	1 % - 5 %	80 %	gering	75
21	Ja	10 %	60 %	gering	75
22-27	ja	10 %	80 %	gering	75-100

Die Experimente eins bis zehn dienen als Vergleichsszenarien, in denen die Container am Terminal wasserseitig auf herkömmlich Weise umgeschlagen werden (siehe Abbildung 1 für die Prozessgestaltung). In diesen Experimenten wird der Füllgrad der Yard Blöcke um jeweils zehn Prozent erhöht. Weiterhin sind die abzuholenden Container in den Experimenten eins bis fünf vorsortiert, wodurch die Umstapelwahrscheinlichkeit für landseitig abgeholte Container eher gering ist. Dagegen ist die Umstapelwahrscheinlichkeit in den Experimenten sechs bis zehn deutlich erhöht. In den Experimenten elf bis 27 werden Container für den Hafen wasserseitig direkt auf externe Lkw umgeschlagen (siehe Abbildung 1 für die Prozessgestaltung). Dabei wird während der Experimente elf bis 20 der Anteil des Direktumschlags an der Kaikante jeweils um ein Prozent erhöht. Darüber hinaus wird der Lagerfüllstand variiert. In den Experimenten 20 bis 27 wird angenommen, dass zehn Prozent aller Container wasserseitig direkt durch externe Lkw abzuholen sind. Dafür wird die Anzahl an Lkw erhöht, die für den Transport von Containern im Hafen verfügbar sind.

Tabelle 2 stellt die Auswertung der Experimente eins bis zehn dar. Es ist zu erkennen, dass die durchschnittliche Umschlaggeschwindigkeit der Containerbrücken deutlich unter der erwarteten Annahme von 24 - 30 Container pro Stunde liegt. Jedoch zeigen die Experimente auch, dass in den Randzeiten die Anzahl an Containern, die pro Stunde vom Schiff geladen werden, deutlich über dem Durchschnitt liegt. Dies lässt sich damit erklären, dass das Containerschiff in den frühen Morgenstunden den Terminal anläuft. In diese Zeit fallen nur sehr wenige hafeninterne Containertransporte. Somit kommen nur gelegentlich Lkw am Terminal an, welche Container bringen oder abholen. Folglich konkurrieren die Yard Trucks nicht so stark mit den Lkw um die RTG im Yard und können schneller abgefertigt werden. Weiterhin lässt sich eine Abnahme der Umschlaggeschwindigkeit mit zunehmendem Füllgrad der Yard Blöcke erkennen. Ein höherer Füllgrad des Yards sorgt dafür, dass vermehrt Umstapelvorgänge notwendig sind, wenn ein Container landseitig abgeholt wird. Somit verzögert sich die Umschlaggeschwindigkeit im Yard, was wiederum Einfluss auf die Produktivität der Containerbrücken hat. Außerdem sind eine Abnahme der durchgeführten Containertransporte sowie eine Zunahme der durchschnittlichen Verspätung der Lkw bei der Durchführung der Transporte zu beobachten.

Tabelle 2: Auswertung der Experimente mit herkömmlicher Ablaufgestaltung

Exp.	Ø moves/h CB	Max. moves/h CB	Σ Tour/Lkw	Ø Verspätung Lkw
1	19,8	28	14,1	01:03
2	19,8	27	13,0	01:38
3	19,7	26	12,2	01:19
4	19,5	26	12,3	01:19
5	19,8	28	12,5	02:23
6	19,7	26	14,1	02:17
7	19,1	26	12,6	02:55
8	19,0	25	13,1	03:22
9	19,2	26	12,7	02:46
10	19,0	26	12,1	03:07

Testläufe vor dem Start der Simulationsstudie haben ergeben, dass für die Durchführung von Direktumschlägen auf externe Lkw eine Anpassung des Modells notwendig ist. Um eine vergleichbare Performance der Containerbrücken zwischen den beiden Alternativen der Ablaufgestaltung zu erreichen, ist es erforderlich den Direktumschlag nur dann stattfinden zu lassen, wenn auch ein externer Lkw für die Übernahme an der Kaikante verfügbar ist. Hierfür wird in dem Simulationsmodell ein direkt umzuschlagender Container, welcher länger als drei Minuten auf seine Übergabe an einen externen Lkw wartet, an den nächsten verfügbaren Yard Truck übergeben und in den Yard transportiert.

Die Auswertung der Experimente zeigt, dass die durchschnittliche Umschlaggeschwindigkeit an den Containerbrücken abnimmt, sobald der Anteil an Direktumschlägen auf externe Lkw steigt. Dafür nimmt die maximale Umschlaggeschwindigkeit in den Randstunden verglichen mit den Experimenten eins bis zehn zu, da Wechselwirkungen zwischen der Performance der Containerbrücke und den RTG teilweise wegfallen. Ebenso lässt sich erkennen, dass die Summe der Touren je Lkw im Vergleich zu der herkömmlichen Ablaufgestaltung zunimmt. Dies lässt sich durch den Tagesgang der Auftragsplanung begründen. Morgens sind wenige hafeninterne Transporte geplant, die Lkw stehen aber dennoch zur Verfügung. Diese ungenutzten Kapazitäten werden für Direktumschläge an der Kaikante genutzt. Schlussendlich nimmt die Verspätung der Lkw über den Tag zu, da Lkw Aufträge mit Direktumschlag höher priorisieren und den Terminal zum Teil bis zu 30 Minuten vor dem tatsächlichen Umschlag anfahren.

Tabelle 3: Auswertung der Experimente mit wassers. Direktumschlag auf Lkw

Exp.	Ø moves/h CB	Max. moves/h CB	Σ Tour/Lkw	Ø Verspätung Lkw
11	18,8	29	16,3	02:31
12	18,4	29	16,47	03:13
13	18,3	28	15,1	03:27
14	19,0	29	14,9	02:57
15	19,3	30	14,3	02:11
16	18,5	30	14,1	04:03
17	17,6	29	14,4	01:28
18	18,1	29	14,2	04:03
19	19,3	28	14,3	06:20
20	18,3	28	14,7	06:12
21	17,6	29	17,5	09:04
22	18,1	29	17,3	08:57
23	18,8	28	15,8	01:18
24	18,1	29	14,6	02:26
25	17,9	28	13,9	01:52
26	18,2	31	13,8	01:38
27	19,6	29	10,4	00:34

Abbildung 2 veranschaulicht den Anteil der tatsächlich durchgeführten wasserseitigen Direktumschläge im Verhältnis zu den geplanten. Es ist zu erkennen, dass nicht einmal jeder dritte geplante Direktumschlag tatsächlich durchgeführt werden konnte.

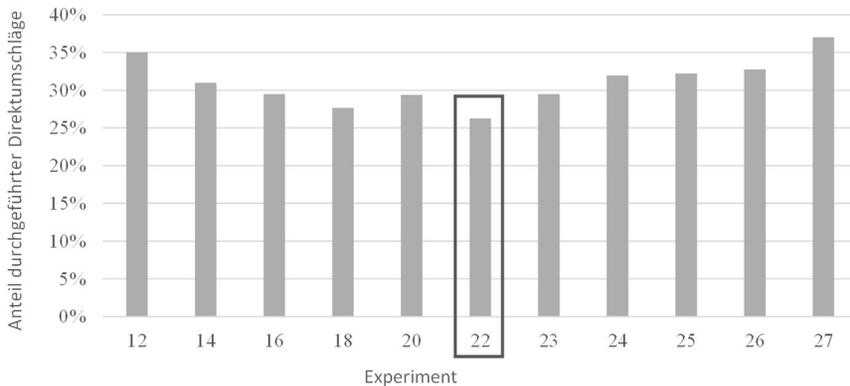


Abbildung 2: *Tatsächlich durchgeführte Direktumschläge*

Durch die Erhöhung des Anteils geplanter Direktumschläge sinkt die Auftragsbefüllungsquote. Bei einem geplanten Anteil von zehn Prozent Direktumschlag können davon lediglich 26 Prozent tatsächlich durchgeführt werden (siehe Experiment 22 in Abbildung 2). Eine höhere Anzahl verfügbarer Lkw im Hafen kann die Auftragsbefüllung zwar steigern, diese liegt jedoch immer noch unter 40 Prozent.

5 Fazit und Ausblick

Diese Studie stellt einen ersten Versuch dar, den Einfluss des Direktumschlags von hafenintern weiter zu transportierenden Containern auf den Terminal mittels Simulation zu untersuchen. Hierfür wird ein neuer Prozess des wasserseitigen Containerumschlags vorgestellt. Es wird deutlich, dass die wasserseitige Umschlaggeschwindigkeit sehr stark von der Verfügbarkeit und Performance des nachgelagerten Equipments abhängig ist. Während bei hafeninternen Transporten Verzögerungen im Prozess toleriert und durch eine geschickte Tourenplanung der Fuhrunternehmen abgefangen werden, wirken sich kleinste Verzögerungen im Direktumschlag stark auf Prozesse am Terminal aus. Dies macht den wasserseitigen Direktumschlag extrem aufwendig. Gleichzeitig bleibt eine hohe Planungsunsicherheit für den Terminal. Dennoch lassen die durchgeführten Experimente Potenziale erkennen. So konnte gezeigt werden, dass sich durch eine Neugestaltung der Terminalabläufe die maximale wasserseitige Umschlaggeschwindigkeit erhöhen lässt.

Ausgehend von dieser Machbarkeitsstudie lassen sich Potenziale für aufbauende Forschungsarbeiten erkennen. Hierzu zählt die Integration von weiteren Hafenteilnehmern nach dem Vorbild unterschiedlicher Containerhäfen. Weiterhin wird in dieser ersten Studie ein einziges Terminal mit RTG Betriebssystem untersucht. Für die Weiterentwicklung des Modells sollten unterschiedliche Betriebssysteme Betrachtung finden. Darüber hinaus wird in der durchgeführten Simulationsstudie lediglich die Entladung eines Containerschiffs nachgebildet. Eine Erweiterung des

Modells um die sich abwechselnden Be- und Entladeprozesse unterschiedlich großer Containerschiffe bildet natürliche Belastungsschwankungen des Systems nach und erhöht so die Realitätsnähe der Studie. Hierbei ist zusätzlich der Betrachtungshorizont zu erhöhen. Abschließend ist zu erwähnen, dass in der Realität zum einen zwischen 20 ft und 40 ft Containern unterschieden wird. Zum anderen können Lkw zwei 20 ft Container gleichzeitig transportieren. Diese Implementierungen sind ebenfalls für zukünftige Untersuchungen geplant.

Literatur

- Allianz: SAFETY AND SHIPPING REVIEW 2020. An annual review of trends and developments in shipping losses and safety 2020.
- Duinkerken, M.B.; Dekker, R.; Kurstjens, S.T.; Ottjes, J.A.; Dellaert, N.P.: Comparing transportation systems for inter-terminal transport at the Maasvlakte container terminals. In: Kim, K.Hwan; Günther, H.-O. (Hrsg.): Container Terminals and Cargo Systems. Springer Berlin Heidelberg 2007, S. 37–61.
- Gharehgozli, A.H.; Koster, R. de; Jansen, R.: Collaborative solutions for inter terminal transport. *Int. J. Prod. Res.* 55 (2017) 21, S. 6527–6546.
- Heilig, L.; Voß, S.: Inter-terminal transportation: an annotated bibliography and research agenda. *Flex. Serv. Manuf. J.* 29 (2017) 1, S. 35–63.
- Huiyun, Y.; Xin, L.; Lixuan, X.; Xiangjun, L.; Zhihong, J.; Zhan, B.: Truck appointment at container terminals: Status and perspectives. In: 2018 Chinese Control And Decision Conference, 2018, S. 1954–1960.
- Kizilay, D.; Eliiyi, D.T.: A comprehensive review of quay crane scheduling, yard operations and integrations thereof in container terminals. *Flex. Serv. Manuf. J.* 264 (2020) 1, S. 138.
- Lange, A.-K.; Schwientek, A.; Jahn, C.: Reducing truck congestion at ports – classification and trends. In: Jahn, C.; Kersten, W.; Ringle, C.M. (Hrsg.): Digitalization in maritime and sustainable logistics, 2017, S. 37–58.
- Liang, C.; Hwang, H.; Gen, M.: A berth allocation planning problem with direct transshipment consideration. *J. Intell. Manuf.* 23 (2012) 6, S. 2207–2214.
- Nellen, N.; Grafelmann, M.; Ziegenbein, J.; Lange, A.-K.; Kreutzfeldt, J.; Jahn, C.: Literature Classification on Container Transport Systems for Inter-terminal Transportation. In: Freitag, M.; Haasis, H.-D.; Kotzab, H.; Pannek, J. (Hrsg.): DYNAMICS IN LOGISTICS, 2020, S. 52–61.
- Ramírez-Nafarrate, A.; González-Ramírez, R.G.; Smith, N.R.; Guerra-Olivares, R.; Voß, S.: Impact on yard efficiency of a truck appointment system for a port terminal. *Annals of Operations Research* 258 (2017) 2, S. 195–216.
- Schwientek, A.; Lange, A.-K.; Jahn, C.: Literature classification on dispatching of container terminal vehicles. In: Jahn, C.; Kersten, W.; Ringle, C.M. (Hrsg.): Digitalization in maritime and sustainable logistics, 2017, S. 3–36.
- UNCTAD: Review of Maritime Transport 2020: UNITED NATIONS 2020.
- Zeng, Q.; Feng, Y.; Chen, Z.: Optimizing berth allocation and storage space in direct transshipment operations at container terminals. *Marit. Econ. Logist.* 19 (2017) 3, S. 474–503.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Simulation von kombinierter Stau- und Routenplanung in Multi-Terminal Häfen

Simulation based analysis of the combined container stowage and ship routing problem in multi-terminal ports

Janna Franzkeit, Hamburg University of Technology, Hamburg (Germany),
janna.franzkeit@tuhh.de

Thomas Voß, Leuphana University Lüneburg, Lüneburg (Germany),
thomas.voss@leuphana.de

Hannah Pache, Hamburg University of Technology, Hamburg (Germany),
hannah.pache@tuhh.de

Jens Heger, Leuphana University Lüneburg, Lüneburg (Germany),
jens.heger@leuphana.de

Carlos Jahn, Hamburg University of Technology, Hamburg (Germany),
carlos.jahn@tuhh.de

Abstract: The contribution presents a combination of analytical approach and simulation to solve the problem of combined container stowage and ship routing. Based on a real-world use case, the approach is demonstrated, and multiple scenarios are presented to emphasize the different pitfalls and challenges. The presented newly developed mathematical model is used to compute the optimal stowage plan while simultaneously optimizing terminal rotation. The optimal solution is tested for robustness using simulation. The contribution shows that the stowage plan and ship routing are interdependent in a multi-terminal port and a multi-criteria consideration of the departure time, overstows and robustness can be beneficial.

1 Einleitung

Der Transport von Containern ist der am schnellsten wachsenden Sektor in der Handelsschifffahrt und der Containerumschlag in Häfen nimmt rasant zu. Im Jahr 2019 wurden weltweit 811,2 Millionen TEU in Häfen umgeschlagen (UNCTAD, 2021). Gleichzeitig hat die internationale maritime Transportkette von Containern, bedingt durch das Wachstum der Schiffsgrößen und der steigenden Anzahl umgeschlagener Container, in den letzten zwanzig Jahren immer weiter an Komplexität gewonnen. So transportieren große Containerschiffe (ab ca. 8.000 TEU) Container zwischen international bedeutenden Containerhäfen (Hubs). Von dort aus

werden die Container mittels Feederschiffen (bis ca. 3.000 TEU) an kleinere Häfen (Spokes) weiter verteilt. In Abbildung 1 ist ein entsprechendes Hafennetzwerk exemplarisch dargestellt.

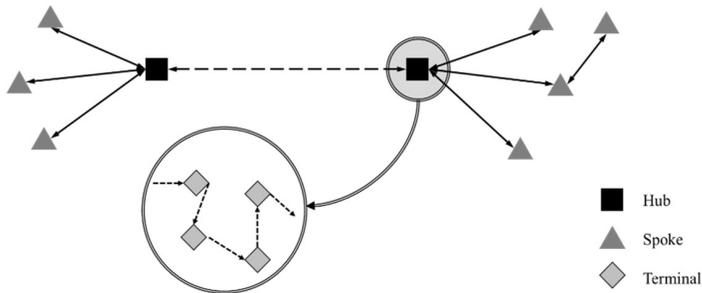


Abbildung 1: Transport zwischen Hubs und Spokes sowie das Anlaufen mehrerer Terminals innerhalb eines Hubs

International bedeutende Containerhäfen (z. B. Singapur, Shanghai, Rotterdam oder Hamburg) haben üblicherweise mehrere Containerterminals und werden daher auch als Multi-Terminal Häfen bezeichnet. Innerhalb eines Multi-Terminal Hafens laufen Feederschiffe, im Gegensatz zu großen Containerschiffen, mehrere Terminals an, um Container zu be- und entladen. Die Reihenfolge in der ein Feederschiff die einzelnen Terminals innerhalb des Hafens anläuft, kann sich kurzfristig durch Liegeplatz- und Containerverfügbarkeiten ändern. Dies hat sowohl Einfluss auf die Stauplanung der Container an Bord des Schiffes als auch auf die Reisedauer. Hieraus ergibt sich das Problem der Kombinierten Stau- und Routenplanung für Schiffe in Multi-Terminal Häfen.

Zur Lösung dieses Problems werden in diesem Beitrag eine mathematische Modellierung und eine Simulation zur Evaluation von Stau- und Routenplänen entwickelt. Ziel der mathematischen Optimierung ist es, die Aufenthaltszeit von Feederschiffe in Multi-Terminal Häfen zu reduzieren und Umstapelvorgänge der Container zu vermeiden. Das Simulationsmodell wird zur Evaluation des dynamischen Verhaltens der Lösungsalternativen unter Unsicherheit genutzt. Die besondere Herausforderung dieses Beitrages liegt in der Dynamik des Szenarios. Er leistet einen praxisnahen Beitrag und zeigt Möglichkeiten zur kombinierten Betrachtung von Stau- und Routenplanung in Multi-Terminal Häfen auf.

2 Stand der Forschung

Im Folgenden werden die einzelnen Aspekte der kombinierten Stau- und Routenplanung sowie die Herausforderungen beschrieben.

Das Schiffsroutenproblem, nah verwandt mit dem Traveling Salesman Problem (TSP), wird in diesem Kontext als das Finden der schnellsten Tour zwischen den Terminals, die jeweils genau einmal angelaufen werden, definiert. In den meisten Fällen ist dies nur unter Berücksichtigung von Anlaufzeitfenster und Abfertigungszeiten möglich (Desrosiers et al., 1995; Christiansen et al., 2013). Besonders im Feederverkehr ist dieses Problem relevant, da Feederschiffe in großen

Häfen, sogenannten Multi-Terminal Häfen (Hellsten et al., 2020), üblicherweise mehrere Terminals anlaufen. Dieses Problem lässt sich grundsätzlich in zwei Kategorien einteilen: statische und dynamische Routenplanung. Bei der statischen Planung wird davon ausgegangen, dass die relevanten Informationen über die Tour konstant sind. Die zu Beginn der Berechnung benötigten Informationen wie Lage, Servicezeiten und Beladung sind konstant und zugänglich. Dabei sind Anwendungsbeispiele sowohl für Schubverbände (Maraš et al., 2013) wie auch für Containerschiffe (Vaferi et al., 2018) in der Literatur bekannt. Diese Annahme ist in realen Anwendungsszenarien unzureichend und im Normalfall sind diese Informationen mit Unsicherheit behaftet. So kommt es vor, dass die Tour dynamisch als Reaktion auf eine Änderung, basierend auf der vorherigen Lösung, umgeplant werden muss. In der Literatur zeigt sich, dass die genaue Betrachtung von Unsicherheit und das angepasste Planen zu einer Verbesserung der Leistung führen kann (Rodrigues et al., 2019).

Ein weiteres Problem ist die Positionierung der zu transportierenden Container auf dem Schiff. Die Positionen (Slots) sind in Bezug auf Längsposition (Bay), Breitenposition (Row) und Höhe (Tier) nummeriert und anhand dieser drei Werte eindeutig definiert. Da die Container übereinandergestapelt werden, folgen sie dem First-In-Last-Out Prinzip. Wenn an einem Hafen ein Container gelöscht werden soll, auf dem ein Container platziert ist, der für einen späteren Hafen bestimmt ist, so muss dieser zunächst entfernt werden. Diese Situation wird als Überstauung und die zusätzlichen Containerbewegungen als Umstapelvorgänge bezeichnet. Diese zusätzlichen Containerbewegungen kosten Zeit und Geld. Daher ist es ein Hauptanliegen der Stauplanung diese zu vermeiden.

Die beiden beschriebenen Probleme beeinflussen sich beim Anlauf mehrerer Häfen bzw. Terminals gegenseitig. Somit ergibt sich das Problem der kombinierten Stau- und Routenplanung. Die zwei Bestandteile des hier untersuchten Problems werden als das Containerstauproblem und das Schiffsroutenproblem (Container Stowage and Ship Routing Problem (CSSRP)) bezeichnet (Martins, P.T.a, Lobo, V.a, Moura, A, 2012; Moura et al., 2013). Das Schiffsroutenproblem ist NP-schwer (Vaferi et al., 2018) und das Problem der optimalen Stauplanung sogar NP-vollständig (Avriel et al., 1998). Es wurde gezeigt, dass das kombinierte Stau- und Routenplanungsproblem über mehrere Häfen mit exakten Methoden gerade noch lösbar ist (Li et al., 2018; Li et al., 2020).

Die Betrachtung von stochastischen Schwankungen in diesem Zusammenhang wurde bisher noch nicht analysiert. Insbesondere im Zusammenhang mit Multi-Terminal Häfen ist dies auf Grund der kurzen Fahrzeiten zwischen den Terminal von Relevanz.

3 Szenario

Das in diesem Paper gewählte Szenario orientiert sich an der Anordnung der Containerterminals im Hamburger Hafen. Der Hamburger Hafen ist Deutschlands größter Containerseehafen mit einem jährlichen Umschlag von 8,5 Mio. TEU in 2020 (Hafen Hamburg Marketing e.V., 2021) und beinhaltet vier Containerterminals. Diese Terminals liegen nicht direkt aneinander, sodass ein Feederschiff mehrere Anlegemanöver benötigt, um alle Terminals anzulaufen. Ähnliche Anordnung von Containerterminals sind auch in anderen historisch gewachsenen Multi-Terminal Häfen üblich (Nellen et al., 2020).

In diesem Beitrag wird ein Feederschiff betrachtet, das innerhalb eines Hubs an vier Terminals Container für vier zukünftig zu besuchenden Häfen (Spokes) auflädt. In der Ausgangssituation des Problems ist die Containerzuordnung zu diesem Schiff sowie die zukünftige Anlaufreihenfolge der Spokes bereits festgelegt. Zu Beginn jeder Instanz sind an den Terminals mehrere Zeitfenster definiert, innerhalb derer das Schiff beladen werden kann. Für jedes Terminal ist die Anzahl an aufzuladenden Containern für die folgenden Spokes festgelegt. Hierbei werden zu Zwecken der Vereinfachung nur 40-Fuß Container für eine Bay mit konstantem Gewicht und Höhe betrachtet. Die Bay umfasst je acht Rows und Tiers und kann somit mit insgesamt 64 Container beladen werden. Die Aufenthaltsdauer am jeweiligen Terminal ist proportional zu der Anzahl an aufzuladenden Container. Die Routenplanung beeinflusst die Reihenfolge, in der die Container aufgeladen werden. Dies beeinflusst die Stauplanung und insbesondere die möglichen Überstauungen, da die Container gestapelt werden und somit dem First-In-Last-Out Prinzip folgen. Das Zielkriterium des kombinierten Ansatzes zur Stau- und Routenplanung ist ein möglichst früher Zeitpunkt zum Verlassen des Hubs mit dem bestmöglichen Stauplan, d.h. mit minimaler Anzahl an Überstauungen. Somit werden Liegezeiten und Umstapelvorgänge reduziert.

Als multikriterielle Optimierung unter Unsicherheit kann abgewogen werden, ob mögliche Zeiteinbußen durch alternative Routen durch den Hub zu einer besseren Stauplanung führen können.

3.1 Mathematische Formulierung des Problems

Das Problem wird in einem ersten Schritt mit mathematischer Optimierung ohne Einbezug von Unsicherheiten gelöst. In Tabelle 1 sind die Konstanten des Modells aufgelistet, die je nach Szenario unterschiedlich konfiguriert werden. Tabelle 2 listet die Entscheidungsvariablen des Modells auf.

Tabelle 1: Konstanten innerhalb eines Szenarios

Konstanten	Beschreibung (Anzahl)
T	Terminals (4)
W	Zeitfenster (1 oder 3)
R	Row (8)
I	Tier (8)
I^*	Tier ohne oberste Position (7)
P	Spokes (4)
$task_t$	Beladungszeit am Terminal t
$open_{t,w}$	Beginn des Zeitfensters w am Terminal t
$close_{t,w}$	Ende des Zeitfensters w am Terminal t
$gatedist_t$	Fahrtzeit zwischen Gate und Terminal t
$dist_{s,t}$	Fahrtzeit zwischen Terminal s und Terminal t
$container_{t,p}$	Anzahl Container am Terminal t für Spoke p

Tabelle 2: Entscheidungsvariablen

Entscheidungsvariable	Typ	Beschreibung
xa	Integer	Ankunftszeit am Gate nach vollständiger Beladung
xu	Integer	Anzahl an Überstauungen
xb_t	Integer	Startzeit der Beladung am Terminal t
$xv_{s,t}$	Binär	1, falls Terminal s vor Terminal t angefahren wird
$xz_{t,w}$	Binär	1, falls Terminal t im Zeitfenster w angefahren wird
$xc_{r,i,t,p}$	Binär	1, falls in Row r , Tier i an Terminal t ein Container für Spoke p geladen wird

Ziel des Optimierungsmodells ist die Minimierung der Zielfunktion:

$$xa + xu \tag{1}$$

Hiermit wird die Aufenthaltsdauer im Hafen sowie die Anzahl an Überstauungen minimiert. Dies geschieht jeweils mit einer Gewichtung von 1, sodass eine Stunde Fahrzeitverlängerung und eine Überstauung als gleichwertig betrachtet werden. Das Mathematische Modell wird definiert als die Minimierung der Zielfunktion (1) unter Einhaltung der Nebenfunktionen (2) – (12):

$$xv_{s,t} + xv_{t,s} = 1 \quad \forall s, t \in T : s \neq t \tag{2}$$

$$xv_{s,t} \times (xb_s + task_s + dist_{s,t}) \leq xb_t \quad \forall s, t \in T : s \neq t \tag{3}$$

$$\sum_{w \in W} xz_{t,w} = 1 \quad \forall t \in T \tag{4}$$

$$xb_t \geq open_{t,w} \times xz_{t,w} \quad \forall t \in T, \forall w \in W \tag{5}$$

$$(xb_t + task_t) \times xz_{t,w} \leq close_{t,w} \quad \forall t \in T, \forall w \in W \tag{6}$$

$$xb_t \geq gatedist_t \quad \forall t \in T \tag{7}$$

$$xb_t + task_t + gatedist_t \leq xa \quad \forall t \in T \tag{8}$$

$$\sum_{t \in T, p \in P} xc_{r,i,t,p} = 1 \quad \forall r \in R, \forall i \in I \tag{9}$$

$$\sum_{r \in R, i \in I} xc_{r,i,t,p} = container_{t,p} \quad \forall t \in T, \forall p \in P \tag{10}$$

$$xc_{r,i,s,p} \times xc_{r,i+1,t,q} \leq xv_{s,t} \quad \begin{aligned} &\forall r \in R, \forall i \in I^*, \\ &\forall s, t \in T : s \neq t, \\ &\forall p, q \in P \end{aligned} \tag{11}$$

$$\sum_{r \in R, i \in I^*, s, t \in T, p, q \in P} xc_{r,i,s,p} \times xc_{r,i+1,t,q} \leq xu \tag{12}$$

Die Bedingungen (2) bis (8) werden für die Routenplanung und die Bedingungen (9) bis (12) für die Stauplanung benötigt. Bed. (2) sorgt dafür, dass für je zwei verschiedene Terminals genau eines zuerst angefahren wird. Auf Grund von Bed. (3) beginnt die Beladung des Schiffes an einem Terminal erst nachdem alle Beladungen am vorherigen Terminals abgeschlossen sind und das Schiff genügend Zeit hatte zum neuen Terminal zu fahren. Aus Bed. (2) und (3) folgt zusammen die Transitivität der Terminalreihenfolge. Bed. (4) fordert, dass das Schiff an jedem Terminal während genau eines Zeitfensters beladen wird. Mit Bed. (5) und (6) wird sichergestellt, dass die Beladung innerhalb dieses Zeitfensters geschieht. Bed. (7) sorgt dafür, dass die Beladung am ersten Terminal erst beginnt, nachdem das Schiff vom Gate zum Terminal gefahren ist. Bed. (8) fordert, dass das Schiff an jedem Terminal beladen wurde und Zeit hat, zum Gate zu fahren, bevor es das Gate verlassen kann. Bed. (9) sorgt dafür, dass nur ein Container pro Slot geladen wird. Mit Bed. (10) wird sichergestellt, dass alle Container geladen werden. Bed. (11) sorgt dafür, dass die Container von unten nach oben geladen werden. Bed. (12) zählt die Anzahl an Überstauungen, definiert als die Anzahl an übereinanderliegender Containerpaare, bei denen der untere Container an einem früherem Spoke gelöscht werden muss.

3.2 Simulation

Die Betrachtung von komplexen Problemstellungen kann in einigen Fällen nicht mit Hilfe mathematischer Modelle durchgeführt werden. Gründe dafür sind zum Beispiel stochastische Einflüsse und dynamische Abhängigkeiten einzelner Systemkomponenten. Besonders in den Fällen, in denen Versuche im realen System zu teuer oder mit verheerenden Folgen verbunden sind, ist dies der Fall (Suhl und Mellouli, 2013).

In diesem Beitrag wurde analog zum mathematischen Modell eine ereignisdiskrete Simulation des Hubs erstellt, um die Zusammenhänge abzubilden. Basierend auf den statischen Werten zur Berechnung des optimalen Stau- und Routenplans wird durch Verifikation sichergestellt, dass das Simulationsmodell identische Werte liefert. Das Modell hat keine Einschwingphase. Ausgehend von der berechneten Terminalrotation fährt das Schiff die Terminals nacheinander an und wird dort, mit einer gegebenen Beladungszeit, abgefertigt. Je nach Reihenfolge der Terminals entstehen reihenfolgeabhängige Fahrzeiten. Da Stau- und Routenplanung bereits abgeschlossen sind, wird die Abfahrtszeit des Schiffes beim Verlassen des Hafens dokumentiert. Die Betrachtung der Leistungsindikatoren wird über 100 Replikationen gemessen.

Da in Abhängigkeit von externen Einflüssen die Fahr- und Beladezeit variieren können, ist der Einfluss dieser zwei Faktoren, unter Anwendung der ereignisdiskreten Modellierung, auf die Leistungsindikatoren zu prüfen. Zu diesem Zweck wird der vorab berechnete, statische Fahr- und Stauplan an die Simulation übergeben und dort unter verschiedenen Stufen von Dynamik bezüglich seiner Leistung evaluiert.

Es werden drei Ausprägungen von Dynamik (± 15 Minuten, ± 30 Minuten und ± 120 Minuten) auf die Beladungs- und Fahrzeiten zur Bewertung betrachtet. Der Leistungsindikator in der Simulation, wie auch im mathematischen Modell, ist die Aufenthaltsdauer. Zur erweiterten Betrachtung wird zusätzlich die Wartezeit in Stunden dokumentiert.

In Abb. 2 ist der Zusammenhang zwischen der Optimierung mit dem mathematischen Modell und der Simulation dargestellt.



Abbildung 2: Zusammenhang Optimierung und Simulation

4 Ergebnisse

Das im Rahmen des Beitrags entwickelte mathematische Modell ist in der Lage Stau- und Routenplanung gleichzeitig zu optimieren. Erstmals wird dies analytische Betrachtung mit Simulation zur Evaluation der optimalen Stau- und Routenplanung erfolgreich kombiniert. Die Ergebnisse werden exemplarisch an fünf ausgewählten Szenarien beschrieben. Die Konfiguration der Szenarien sowie die Ergebnisse sind in Tabelle 3 dargestellt. Die Anzahl an Zeitfenstern pro Terminal sind in Klammern unter Zeitfensterkonfiguration angegeben. Die Betrachtung der optimalen Ergebnisse unter Unsicherheit gibt Aufschluss über die Robustheit der berechneten Pläne.

Tabelle 3: Konfiguration und Ergebnisse der Szenarien

	Container- konfiguration	Zeitfenster- konfiguration	Ohne Überstauung	Optimale Route	Ohne Wartezeiten
Szenario 1	A	a (1)	Ja	Ja	Ja
Szenario 2	A	b (3)	Ja	Ja	Nein
Szenario 3	B	c (3)	Nein	Ja	Ja
Szenario 4	C	d (3)	Ja	Nein	Nein
Szenario 5	D	d (3)	Ja	Nein	Nein

Szenario 1 ist das einzige Szenario, in dem nur 1 Zeitfenster pro Terminal vorgegeben wird. Dieses Szenario ist ohne Wartezeiten und mit optimaler Route in Bezug auf Fahrzeitminimierung lösbar. Szenario 2 verwendet dieselbe Containerkonfiguration wie Szenario 1, aber restriktivere Zeitfenster. Die Rotation in Szenario zwei ist in Bezug auf Fahrzeiten ebenfalls optimal, es kommt hingegen zu Wartezeiten. Beide Szenarien sind ohne Überstauung lösbar.

Unter stochastischen Schwankungen zeigt sich (siehe Abbildung 3), dass die mittlere Wartezeit bei 2 Stunden, analog zum optimalen Fahr- und Stauplan, liegt. Bei Schwankungen der Fahr- und Beladezeit um 120 Minuten können sich, in seltenen Fällen, bis zu 6 Stunden Wartezeit ergeben.

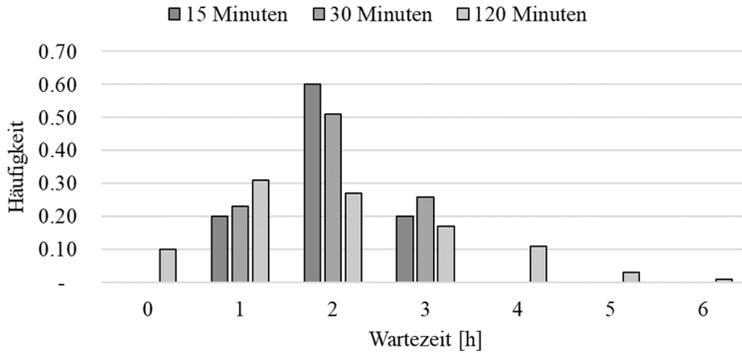


Abbildung 3: Häufigkeit der Wartezeit unter schwankenden Fahr- und Beladezeiten

Die im 3. Szenario gewählte Konfiguration führt zu minimaler Fahrzeit mit zwei Fällen von Überstauung im Stauplan (siehe Abbildung 4). Die grauen Bereiche kennzeichnen bereits verstaute Container vom vorherigen Terminal.

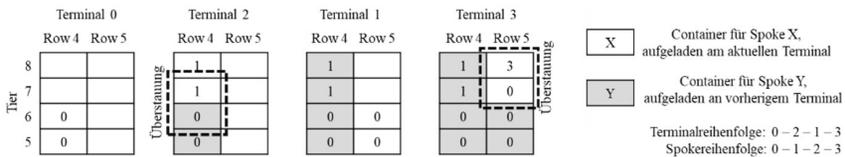


Abbildung 4: Nicht zu vermeidende Überstauung bei der gewählten Rotation

Die Szenarien 4 und 5 verwenden dieselbe Konfiguration von restriktiven Zeitfenstern und unterscheiden sich in Bezug auf die Containerkonfiguration. In Szenario 4 ist die Ankunftszeit am Gate um 6 Stunden schlechter als die optimale Rotation unter Missachtung der Zeitfenster wäre, da die Fahrzeit sich um 2 Stunden verlängert und es zu 4 Stunden Wartezeit kommt. Die Rotation aus Szenario 4 würde in Szenario 5 zu Überstauung führen, weswegen sich in diesem Szenario die Rotation verändert, was zu einer zusätzlichen Stunde Wartezeit führt. Die verschiedenen Routen der Szenarien werden in Abbildung 5 gezeigt, wobei die durchgezogene und die gestrichelte Linie die unterschiedlichen Lösungen darstellen.

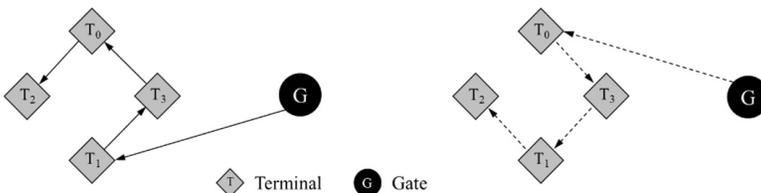


Abbildung 5: Zwei Rotationen mit unterschiedlichen Abfahrtszeiten und Beladungsplänen bei gleichbleibenden Zeitfenstern

An dieser Stelle (siehe Abbildung 6) zeigt sich, dass es, auf Grund der gegebenen Zeitfenster, bei beiden Rotationen mit starken Schwankungen der Fahr- und Beladezeiten zu einer breiten Verteilung an Wartezeiten kommt. Im Umkehrschluss kann festgestellt werden, dass eine geringfügige Schwankung, in diesem Szenario, ohne gravierende Auswirkungen auf die Aufenthaltsdauer sind. Die kombinierte Betrachtung der Stau- und Routenplanung unter Unsicherheit zeigt, dass in den meisten Fällen Schwankungen bis 30 Minuten nicht zu zusätzlichen Wartezeiten führen. Weiterhin zeigt sich in diesen Szenarien, dass bei starker Streuung der Fahr- und Beladezeiten, bis zu 4 Stunden zusätzliche Wartezeit eingeplant werden müssen. Dieses Verhalten gilt sowohl bei Routen- und Stauplänen mit und ohne Überstauung.

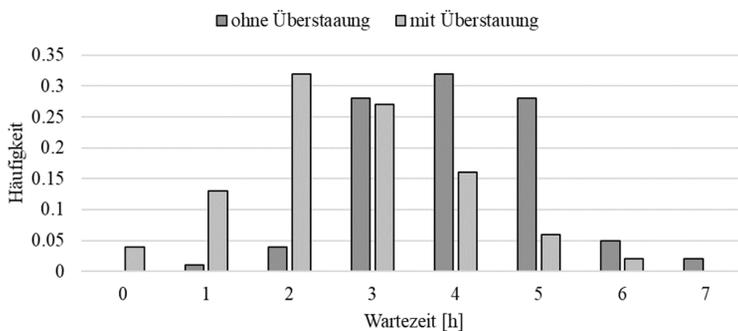


Abbildung 6: Unabhängig von der Rotation, führt eine strake Dynamik zu einer breiten Verteilung der Wartezeiten

Der Mehrwert einer solchen Betrachtung zeigt die möglichen Auswirkungen des dynamischen Charakters des Problems auf einen statischen Plan. Weiterhin kann dies als Entscheidungsunterstützung genutzt werden.

5 Zusammenfassung und Ausblick

Der Beitrag zeigt die Kombination aus analytischem Vorgehen und Simulation zur Lösung des Problems der kombinierten Stau- und Routenplanung. Das im Rahmen des Beitrags entwickelte mathematische Modell wird zur zeitgleichen Berechnung des optimalen Stauplans und der Terminalrotation genutzt. Die optimale Lösung wird durch Simulation auf ihre Robustheit geprüft. Der Beitrag zeigt, dass die Stau- und Routenplanung abhängig voneinander sind und eine multikriterielle Betrachtung der Abfahrtszeitpunkte, Überstauung und Robustheit sinnvoll ist. Der Ausblick erstreckt sich über ein breites Spektrum an Erweiterungen. Exemplarisch aufgezeigt sei die Implementierung von praxisnahen Details wie z.B. Größe und Transportanforderungen oder unsichere Parameter von Containern wie das Gewicht. Zusätzlich kann eine Steigerung der Komplexität durch die Betrachtung von mehreren Bays an Bord des Feederschiffes den Lösungsraum sinnvoll erweitern. Weiterhin steht die Betrachtung der multikriterielle Optimierung unter Einbeziehung mehrere Leistungsindikatoren als Erweiterung zur Möglichkeit. Schlussendlich kann ein verbessertes methodisches Vorgehen zur Berechnung und Evaluation der Lösung und dem ableiten von Handlungsempfehlungen mit Simulation entwickelt werden.

Literatur

- Avriel, M.; Penn, M.; Shpirer, N.; Witteboon, S.: Stowage planning for container ships to reduce the number of shifts. *Annals of Operations Research* 76 (1998), S. 55–71.
- Christiansen, M.; Fagerholt, K.; Nygreen, B.; Ronen, D.: Ship routing and scheduling in the new millennium. *European Journal of Operational Research* 228 (2013) 3, S. 467–483.
- Desrosiers, J.; Dumas, Y.; Solomon, M.M.; Soumis, F.: Time constrained routing and scheduling. In: *Network Routing: Elsevier* 1995, S. 35–139.
- Hafen Hamburg Marketing e.V., 2021: Containerumschlag. Online verfügbar unter <https://www.hafen-hamburg.de/de/statistiken/containerumschlag/>, 04.05.2021.
- Hellsten, E.O.; Sacramento, D.; Pisinger, D.: An adaptive large neighbourhood search heuristic for routing and scheduling feeder vessels in multi-terminal ports. *European Journal of Operational Research* 287 (2020) 2, S. 682–698.
- Li, J.; Zhang, Y.; Ji, S.; Zheng, L.; Xu, J.: Multi-stage hierarchical decomposition approach for stowage planning problem in inland container liner shipping. *Journal of the Operational Research Society* 71 (2020) 3, S. 381–399.
- Li, J.; Zhang, Y.; Ma, J.; Ji, S.: Multi-Port Stowage Planning for Inland Container Liner Shipping Considering Weight Uncertainties. *IEEE Access* 6 (2018), S. 66468–66480.
- Maraš, V.; Lazić, J.; Davidović, T.; Mladenović, N.: Routing of barge container ships by mixed-integer programming heuristics. *Applied Soft Computing* 13 (2013) 8, S. 3515–3528.
- Martins, P.T.a, Lobo, V.a, Moura, A.: Short sea shipping route and stowage plan optimization. In: *Maritime Engineering and Technology - Proceedings of 1st International Conference on Maritime Technology and Engineering, MARTECH 2011 2012*,
- Moura, A.; Oliveira, J.; Pimentel, C.: A Mathematical Model for the Container Stowage and Ship Routing Problem. *Journal of Mathematical Modelling and Algorithms in Operations Research* 12 (2013) 3, S. 217–231.
- Nellen, N.; Poeting, M.; Bschorer, K.; Jahn, C.; Clausen, U.: Impact of Port Layouts on Inter-Terminal-Transportation Networks. Unter Mitarbeit von TUHH Universitätsbibliothek, 2020,
- Rodrigues, F.; Agra, A.; Christiansen, M.; Hvattum, L.M.; Requejo, C.: Comparing techniques for modelling uncertainty in a maritime inventory routing problem. *European Journal of Operational Research* 277 (2019) 3, S. 831–845.
- Suhl, L.; Mellouli, T.: *Optimierungssysteme*. Berlin, Heidelberg: Springer Berlin Heidelberg 2013.
- UNCTAD: *Review of Maritime Transport 2020*. Hg. v. United Nations Conference on Trade and Development, 2021,
- Vaferi, M.; Fallah, M.S.; Tayebi, A.H.: A Metaheuristic for the Containership Feeder Routing Problem with Port Choice Process. *The Asian Journal of Shipping and Logistics* 34 (2018) 2, S. 119–128.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Erarbeitung einer Prozesssteuerungsstrategie für zwei Transportmittel mit gemeinsamen Aktionsbereich am Beispiel eines Prozesskransystems zur Kommissionierung von Schüttgut

***Development of a control strategy for two transport vehicles in a
shared space for a process crane system for handling of bulk material***

Vincent Betker, Michael Völker, Thorsten Schmidt,
TU Dresden, Dresden (Germany), vincent.betker@tu-dresden.de

Abstract: This paper gives an account of the application of the discrete-event simulation method in the planning process of twin process cranes intended for handling of bulk material. We detail the lack of research in real-world crane scheduling under uncertainty needed for our use case and propose a testing framework for twin crane systems by enhancing an existing crane model with the implementation of collision avoidance routines. We then use the model for evaluating the performance of different dispatching rules in our use case. We identify a set of rules that lead to a cycle time reduction of 16.2% compared to the zoning method commonly used for job allocation in these contexts.

1 Einführung

Eine Möglichkeit zur Steigerung des Durchsatzes von Transportsystemen ist die Erhöhung der Anzahl der Transportmittel. In Systemen mit großen Überschneidungen des Aktionsraumes der Transportmittel ist die dadurch erzielbare Durchsatzerhöhung jedoch schwer zu ermitteln. Beispiele für solche Systeme sind Krane mit gemeinsamer Kranbahn oder Hochregalgassen mit mehreren Regalbediengeräten. Analytische Methoden verlieren in diesen Fällen schnell an Aussagekraft, da die Wartezeiten vom Systemzustand abhängig sind. Erwartungswerte für die Durchsatzerhöhung durch Hinzunahme eines zweiten Transportmittels existieren, sind jedoch abhängig von den räumlichen Gegebenheiten und den anfallenden Transportaufträgen und damit nur von beschränkter Aussagekraft für abweichende Anwendungsfälle (Schulze und Könze 2008). Abhilfe kann hier die ereignisdiskrete Simulation schaffen. In einer Praxisanwendung wurde die Software *AutoMod* (Version 14.0) genutzt, um ein in Planung befindliches Kransystem mit zwei Kranen zu modellieren. Hierbei handelt es sich um eine Krananlage zur Beschickung eines Produktionsprozesses, welche im

Zuge von Maßnahmen zur Steigerung der Produktionsleistung als Engpass erkannt wurde.

Eine Voruntersuchung ergab, dass der geplante Durchsatz mit einem einzelnen Kran nicht zu erreichen ist. Das Unternehmen im Anwendungsfall ist zudem in einer dem globalen Wettbewerbsdruck ausgesetzten Branche der Prozessindustrie tätig und hat daher ein besonderes Interesse an automatisierten Abläufen. Der Schüttgutumschlag soll vor diesem Hintergrund von einer Zweikrananlage übernommen werden. Mit der Simulation wurden zwei Ziele verfolgt: zum einen die Dimensionierung der (individuell gefertigten und damit bedarfsgerecht konfigurierbaren) Krananlage; zum anderen die Entwicklung von Steuerungsregeln für einen automatisierten Betrieb des Kransystems. In diesem Beitrag wird die dafür erforderliche Zweikransteuerung vorgestellt und bewertet.

2 Vorstellung des Anwendungsfalls

2.1 Untersuchungsgegenstand

Im vorliegenden Anwendungsfall wurde die Neuplanung einer Halle mit Prozesskränen zur Versorgung eines Produktionsprozesses betrachtet. Abbildung 1 zeigt das untersuchte Kransystem und den Aufbau der Kranhalle.

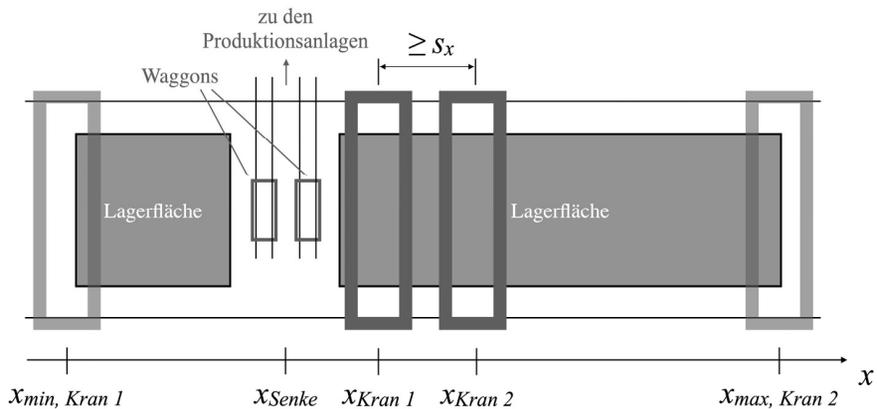


Abbildung 1: Schematische Darstellung des Untersuchungsgegenstandes – Aufsicht der Kranhalle mit Prozesskränen, Lagerflächen und Bahngleisen

Die Kranhalle dient als Zwischenlager vor den Produktionsanlagen, um für diese eine ununterbrochene Materialversorgung sicherzustellen. Das Material wird entsprechend eines Mischungsverhältnisses vorkommissioniert in Waggons bereitgestellt. Dafür werden auf den Lagerflächen verschiedene Materialarten sortenrein gelagert. Die Waggons transportieren das Material zu den Produktionsanlagen. Da die Menge an aufgenommenem Material pro Hub starken Schwankungen unterliegt, lässt sich das Mischungsverhältnis nur annähernd in Krantransportspiele umrechnen. Die Lagerflächen werden per Förderband und Lkw beschickt; die Abkippsstellen der Lkws

und Abwurfstellen der Förderbänder müssen freigehalten werden, um einen möglichst großen Materialzufluss zu sichern. Dies wird erreicht, indem die Krane bevorzugt Material aus diesen kritischen Bereichen zur Beschickung der Waggons entnehmen bzw. von dort in die übrigen Bereiche umlagern, wenn kein Waggon zur Beladung bereitsteht. Da die Beschickung des Zwischenlagers nur werktags und tagsüber erfolgen kann, müssen diese Umlagerungen dazu führen, dass sich eine ausreichende Lagerreichweite für die Phasen ohne Materialeingang bildet.

2.2 Stand der Technik

Im Bereich der Steuerung von Mehrkransystemen sind insbesondere Container-Verladeterminals Gegenstand der Forschung. Hier wurde bereits eine Vielzahl von Lösungsansätzen zur effizienten Steuerung von Mehrkransystemen entwickelt (Boysen et al. 2017). Die im Falle von Containerkränen ebenfalls auftretenden Zwischenlagerungen mit Ein-, Aus- und Umlagervorgängen ähneln dem Betrieb in der untersuchten Kranhalle.

Die Optimierung der Transportspielzeiten in einem Mehrkransystem stellt bereits bei zwei Kranen ein NP-schweres Planungsproblem dar (Lim et al. 2007). Die bekannten Ansätze zur Optimierung der Transportspielzeiten bilden das Transportsystem daher vereinfacht ab. In einem umfassenden Review konstatieren Boysen et al. (2017), dass unter den über 90 untersuchten Optimierungsansätzen des Mehrkranplanungsproblems nicht ein einziger Ansatz die Bewegungen entlang aller drei Achsen mit seinen Beschleunigungs- und Verzögerungsvorgängen abbildet. Ein kollisionsfreier Mehrkranbetrieb kann auf Basis des Optimierungsergebnisses demnach nicht garantiert werden. Im Realbetrieb muss daher eine Optimierung immer mit einer Echtzeit-Kollisionsvermeidung gepaart werden, die bestenfalls nicht nur Kollisionen verhindert, sondern proaktiv vermeidet und so den Weiterbetrieb der Anlage sicherstellt (Speer et al. 2011). Im Echtzeitbetrieb haben sich zudem (regelbasierte) Dispatchingansätze als vorteilhaft gegenüber (vorausplanenden) Schedulingansätzen erwiesen, da zur Verhinderung von *Deadlocks* die Schedulingansätze so trivial gehalten werden müssen, dass sie regelbasiert abbildbar sind bzw. den regelbasierten Dispatchingansätzen ähneln (Petering et al. 2009).

Im vorliegenden Anwendungsfall wurde im Hinblick auf die spätere Übertragbarkeit der Steuerung auf das Realsystem wie auch eine möglichst realistische Abschätzung der Systemleistung eine detaillierte ereignisdiskrete Simulation gewählt, mit der sich regelbasierte Steuerungsstrategien vergleichen lassen. So ließ sich neben der Entwicklung einer Steuerungsstrategie für den automatisierten Betrieb auch die bei Beschaffung der Krananlage vordergründig notwendige Dimensionierung des Kransystems erreichen: statt theoretische, aber vielleicht nicht umsetzbare Abläufe für die Durchsatzermittlung heranzuziehen, führt das simulierte System die Anweisungen der Steuerung streng unter Beachtung der Durchführbarkeit (und damit unter Umständen mit längeren Zykluszeiten) aus.

2.3 Formale Beschreibung

Das betrachtete Kransystem besteht aus zwei Kranen auf einer gemeinsam genutzten Kranbahn. Ein Überholen ist nicht möglich, vielmehr muss jederzeit $x_{Kran\ 1} + s_x \leq x_{Kran\ 2}$ mit dem sicheren Abstand s_x zwischen den Mitten der Kranbrücken

eingehalten werden. Die Arbeitsbereiche der Krane überschneiden sich zu großen Teilen, lediglich die Kranbahnen sind einzelnen Kranen vorbehalten (Gl. 1 und 2).

$$x_{min, Kran 1} \leq x_{Kran 1} \leq x_{max, Kran 1} = x_{max, Kran 2} - S_x \quad (1)$$

$$x_{min, Kran 1} + S_x = x_{min, Kran 2} \leq x_{Kran 2} \leq x_{max, Kran 2} \quad (2)$$

Beide Krane besitzen identische technische Parameter. Die Bewegungen der Last bzw. des Lastaufnahmemittels im Raum definieren sich über die von der Kranbrücke (Dimension x), der Katze (Dimension y) und dem Hubwerk (Dimension z) erreichbare Beschleunigung, Verzögerung und Höchstgeschwindigkeit. Die Beschleunigungs- und Verzögerungswerte werden als gleich angenommen. Alle Bewegungen können überlagert ausgeführt werden. Während der Bewegungen muss zudem sichergestellt sein, dass die Position der Last hoch genug ist, um eine Kollision mit dem Lagergut bzw. den Anlagen in der Kranhalle zu verhindern. Dabei muss eine sichere Höhe s_z über dem Hindernis eingehalten werden. Die Lastaufnahmemittel definieren sich über stochastische Aufnahmemengen und Lastaufnahmezeiten.

Das Ziel dieser Arbeit ist das Aufstellen von Regeln, mit Hilfe derer die Krane die Beschickung der Waggonen entsprechend eines vorgegebenen Mischungsverhältnisses in möglichst kurzer Zeit durchführen können.

3 Vorstellung des Lösungsansatzes

Für die Umsetzung des Simulationsmodells wurde das *Bridge Crane*-Modul der Simulationssoftware *AutoMod* derart erweitert, dass ein Zweikranbetrieb auf der vollen Länge der Kranbahn möglich ist. Zudem wurde die Steuerung der dritten Dimension (Hubwerk) ergänzt. Zwei Kollisionsschutzalgorithmen sorgen dafür, dass zwischen den Kranen wie auch zu Objekten im Raum gewisse Mindestabstände eingehalten werden. Daraus resultierende Wartezeiten werden zur Laufzeit berechnet. Die Logik zur Steuerung der Kranbrücken wird im Abschnitt 3.1 vorgestellt, jene zur Steuerung der Hubwerke in Abschnitt 3.2.

3.1 Implementierung des Zweikranbetriebs

Die Echtzeitsteuerung eines Mehrkransystems muss folgende Eigenschaften erfüllen (in Anlehnung an Petering et al. 2009):

- Die nächste Aktion eines Krans muss für jeden Systemzustand bekannt sein.
- Der Sicherheitsabstand zwischen den Kranen muss zu jeder Zeit eingehalten werden.
- Ein *Deadlock* muss in jedem Systemzustand ausgeschlossen sein.

Ein sog. *Deadlock* ist eine Situation, in dem beide Krane ein Ziel ansteuern sollen, welches nicht erreichbar ist, da der jeweils andere Kran dieses blockiert. Im Realbetrieb würde dies zu einer Notbremsung beider Krane und einem Stillstand des Systems führen. Um die Situation zu lösen, müsste einer der Krane manuell so weit zurückgefahren werden, dass der andere Kran sein Ziel erreichen kann. Diese Situation wird im Simulationsmodell verhindert, indem ein bereits wartender Kran priorisiert wird, woraufhin der andere Kran den Weg zum Ziel zeitnah freigibt. Abbildung 2 stellt den implementierten Algorithmus dar.

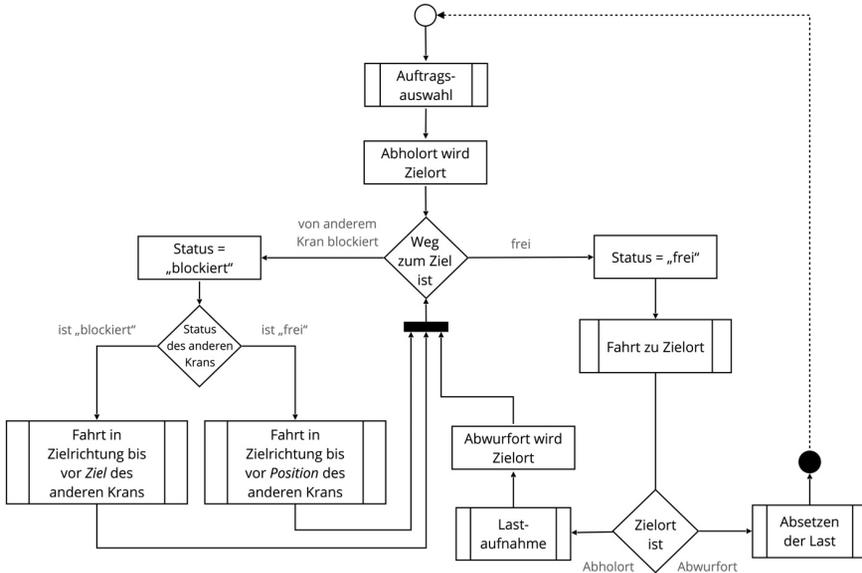


Abbildung 2: Prozessablaufdiagramm eines Krans unter Berücksichtigung des zweiten Krans

Ein zentrales Entscheidungskriterium ist dabei die Erreichbarkeit des vorgegebenen Ziels. Der Weg zum Zielort x_{Ziel} wird als blockiert angesehen, wenn $x_{Ziel, Kran 1} + s_x > x_{Ziel, Kran 2}$ gilt. Wenn der andere Kran nicht in Bewegung ist, gilt dabei dessen aktueller Ort als dessen Zielort. Der Ablauf stellt einen kollisionsfreien und unterbrechungsfreien Betrieb sicher, wenn Start und Ziel der Transportaufträge vom jeweiligen Kran gemäß Gleichung 1 bzw. 2 erreichbar sind. Die Kollisionsfreiheit wurde im Nachgang der Simulationsläufe mithilfe von während der Simulation aufgezeichneten Daten und einem daraus erstellte Weg-Zeit-Diagramm überprüft (s. Abb. 3). Der unterbrechungsfreie Betrieb wurde in einer Vielzahl von Simulationsläufen festgestellt.

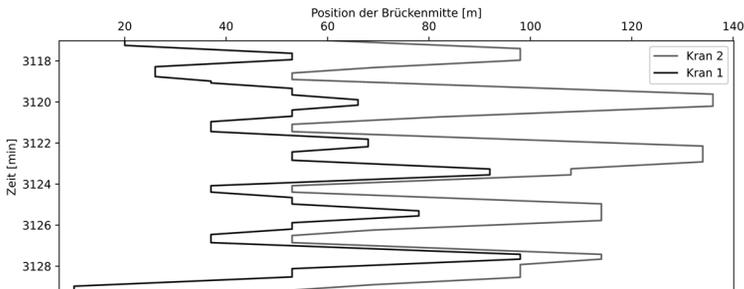


Abbildung 3: Ausschnitt eines Weg-Zeit-Diagramms eines Simulationslaufs

3.2 Implementierung der Hindernisüberquerung

Die Überquerung von auf dem Hallenboden befindlichen Hindernissen in sicherer Höhe wird durch Zeiten reiner Hubbewegung zu Beginn und zum Ende der Kranfahrt sichergestellt. In diesen Zeiten wird die Last gehoben bzw. abgesenkt, Brücke und Katze werden nicht verfahren. Die überlagerte Bewegung findet im Anschluss statt bzw. endet davor (s. Abb. 4).

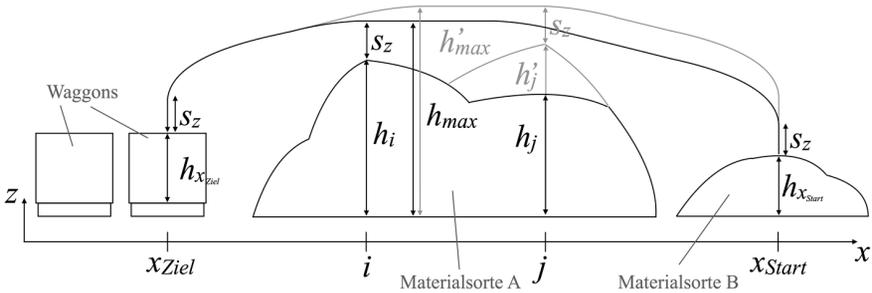


Abbildung 4: Relevante Größen zur Ermittlung der Zeiten mit reiner Hubbewegung am Beispiel zweier Lagertopologien

Zur Ermittlung der Wartezeit zu Beginn der Fahrt (t_{Start} , Gl.3) wird für jeden zu überquerenden Punkt x die Dauer der Bewegung der Brücke dorthin ($t_{x,Brücke}$, Gl.4) und die Dauer des Hebens der Last auf die dort nötige Höhe $h_x + s_z$ ($t_{x,Hubwerk}$, Gl. 5) ermittelt. Die Bewegungsdauern werden unter Berücksichtigung der anteiligen beschleunigten und gleichförmigen Bewegung ermittelt (Gl. 7 mit der Beschleunigungsstecke s_a , Gl. 6). Kann die Hubbewegung nicht schnell genug erfolgen und droht daher eine Kollision, wird die Differenz der Bewegungsdauern als Wartezeit veranschlagt. Die Berechnung erfolgt analog beim Absenken der Last, dafür wird der Startzeitpunkt der Absenkbewegung entsprechend verzögert.

$$t_{Start} = \max(0, t_{x,Brücke} - t_{x,Hubwerk} \mid x \in \{x_{Start}, x_{Start+1}, \dots, x_{Ziel}\}) \quad (3)$$

$$t_{x,Brücke} = \text{Dauer}_{\text{sektional}}(|x_{Start} - x_{Ziel}|, |x - x_{Start}|, v_{Brücke}, a_{Brücke}) \quad (4)$$

$$t_{x,Hubwerk} = \text{Dauer}_{\text{sektional}}(h_{max} - h_{Start}, \max(0, h_x + s_z - h_{Start}), v_{Hubwerk}, a_{Hubwerk}) \quad (5)$$

$$s_a = \frac{v^2}{2a} \quad (6)$$

$$\text{Dauer}_{\text{sektional}}(s_{Ges}, s_{Sek}, v, a)$$

$$= \begin{cases} 2 * \sqrt{\frac{s_{Ges}}{a}} - \sqrt{2 * \frac{s_{Ges} - s_{Sek}}{a}} & ; s_{Sek} > \frac{s_{Ges}}{2} \wedge s_{Ges} \leq 2s_a \\ \frac{s_{Sek}}{v} + \frac{v}{2a} & ; s_{Ges} - 2s_a \geq s_{Sek} > \frac{s_{Ges}}{2} \wedge s_{Ges} > 2s_a \\ \frac{s_{Ges}}{v} - \sqrt{2 * \frac{s_{Ges} - s_{Sek}}{a}} & ; s_{Ges} > s_{Gesamt} - s_a \wedge s_{Ges} > 2s_a \\ 2 * \sqrt{\frac{s_{Sek}}{a}} & ; sonst \end{cases} \quad (7)$$

4 Auswertung der Steuerungsstrategien

4.1 Untersuchte Dispatchingregeln

Die im Anwendungsfall vorgefundenen baulichen Gegebenheiten sind nicht veränderbar. Dies betrifft insbesondere die Ausdehnung der Lagerflächen und die Lage der Gleise der Waggon zur Beschickung der Produktionsanlagen. Damit ergibt sich eine asymmetrische Verteilung der Umlageraufträge; die Beschickung der Waggon findet mehrheitlich von rechts statt. Eine Aufteilung der Krane auf zwei Bedienbereiche mit Prioritätsregeln für den gemeinsamen Arbeitsbereich (hier die Bahngleise) ist vor diesem Hintergrund als Lösungsansatz weniger geeignet, da damit der mögliche Durchsatz des Kransystems nicht voll ausgeschöpft wird (Briskorn et al. 2016). Als Vergleichsgröße findet dieser in der Praxis übliche Ansatz jedoch Eingang in das Simulationsmodell. Die ebenfalls übliche Nutzung einer der beiden Krane als „Zulieferer“ für den anderen Kran wurde verworfen, da dies durch die zusätzlichen Handhabungsvorgänge eine Durchsatzverringerung erwarten lässt (Gharehgozli et al. 2017). Folgende Regeln wurden abgebildet und untersucht:

fixZone teilt den Arbeitsbereich des Kransystems an der Materialsenke auf. Die Auftragsmenge teilt sich so in zwei Gruppen: Kran 1 bearbeitet ausschließlich Aufträge mit $x_{start} < x_{senke}$, Kran 2 solche mit $x_{start} > x_{senke}$. Aus der Gruppe der möglichen Aufträge wird der Auftrag mit größtem x_{start} zuerst ausgeführt. Die Arbeitsbereiche überschneiden sich lediglich über der Senke.

dynZone teilt den Auftragsbestand in zwei gleich große Mengen. Dafür wird der Median aus allen x_{start} gebildet. Kran 1 bearbeitet jene Aufträge, deren $x_{start} \leq \tilde{x}$, während Kran 2 jene annimmt, deren $x_{start} \geq \tilde{x}$.

mostNeighbours Kran 1 nimmt den Auftrag mit kleinstem x_{start} an, während Kran 2 aus den Aufträgen mit einer Startposition gemäß Gleichung 8 denjenigen Auftrag auswählt, in dessen Umgebung sich am meisten Aufträge befinden.

$$x_{senke} < x_{start} \leq x_{max}, \text{ Kran 1} \quad (8)$$

Die Anzahl $n_{Umgebung,i}$ der einen Auftrag mit Startposition $x_{start,i}$ umgebenden Aufträge ergibt sich aus der Menge an Aufträgen, deren Startpositionen innerhalb eines durch den Sicherheitsabstand definierten Bereichs liegen:

$$\begin{aligned} n_{Umgebung,i} \\ = \{ \{ A_j \mid x_{start,i} - s_x < x_{start,j} < x_{start,i} + s_x \} \} \end{aligned} \quad (9)$$

Diese Regel der Auftragsauswahl führt Kran 2 so lange aus, wie es mindestens einen Auftrag mit $x_{start} < x_{Schrottfähre}$ und einen weiteren mit x_{start} entsprechend Gleichung 8 gibt. Ist dies nicht der Fall, wählt Kran 2 den Auftrag mit dem höchsten x_{start} .

Die Priorisierung im Fall einer drohenden Kollision erfolgt in allen Varianten gemäß der in Abschnitt 3.1 vorgestellten Zweikransteuerung. Es erhält demnach in jeder Variante derjenige Kran die Vorfahrt, welcher zuerst eine Fahrt in einen Konfliktbereich anmeldet.

4.2 Bewertung der Dispatchingregeln

Zur Bewertung der Dispatchingregeln wurden diese in mehreren Simulationsläufen in einem realitätsnahen Szenario genutzt. Der zu liefernde Materialmix wurde mit 25 t aus der kleinen Lagerfläche und 75 t aus der großen Lagerfläche angesetzt. Diese Menge wird alle 40 Minuten von den Produktionsanlagen abgerufen. Die Lastaufnahmekapazität wurde als gleichverteilt von 4 t bis 8 t definiert. Die Lastaufnahmedauer beträgt 20 s bis 40 s und unterliegt ebenfalls einer gleichförmigen Verteilung. Relevante Daten zu den räumlichen Ausmaßen sind $x_{min, Kran 1} = 0 m$, $x_{max, Kran 2} = 136 m$, $x_{Senke,1} = 39 m$ und $x_{Senke,2} = 47 m$. Für den Sicherheitsabstand zwischen den Brückenmitten gilt $s_x = 16 m$. Die technischen Daten der Krane für die relevanten Bewegungsrichtungen lauten $v_{Brücke} = 120 m/s$, $a_{Brücke} = 0,2 m/s^2$, $v_{Hubwerk} = 20 m/s$, sowie $a_{Hubwerk} = 0,25 m/s^2$. Zur Auswertung wurden die Zykluszeiten gemessen, d. h. die Dauer des Beladungsvorgangs einer Materiallieferung. Je Dispatchingregel sind fünf Simulationsläufe à 350 Tagen ausgeführt worden, pro Regelwerk sind somit etwa 70.000 Zyklen simuliert worden. Die dabei erreichten Zykluszeiten sind in Abbildung 5 zusammengefasst.

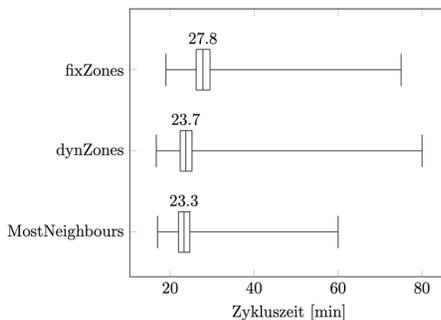


Abbildung 5: Verteilung der Zykluszeiten in Abhängigkeit der verwendeten Dispatchingregeln

Für die Auftragsverteilung nach festen Arbeitsbereichen (*fixZones*) ergibt sich im Mittel (Median) eine Zykluszeit von 27,8 Minuten. Mit der dynamischen Aufteilung der Aufträge (*dynZones*) und der damit einhergehenden stärkeren Überschneidung der Arbeitsbereiche wird im Mittel eine Zykluszeit von 23,7 Minuten erreicht (– 14,7 %). Die Anwendung der Regeln gemäß *mostNeighbours* ermöglicht im Mittel eine weitere

Reduktion der Zykluszeit auf 23,3 Minuten (− 16,2 %). Während der Simulation werden zudem die Statusänderungen der Krane protokolliert. Somit lässt sich auswerten, welchen Tätigkeiten die Krane nachgegangen sind (s. Abb. 6). Daraus geht hervor, dass Kran 1 mit *mostNeighbours* eine geringere Auslastung erfährt als mit *dynZones*, während die Tätigkeitsverteilung von Kran 2 mit diesen beiden Strategien nahezu identisch ist. Damit ergeben sich mit *mostNeighbours* weitere Reserven für eventuelle Ausfälle oder Durchsatzsteigerungen.

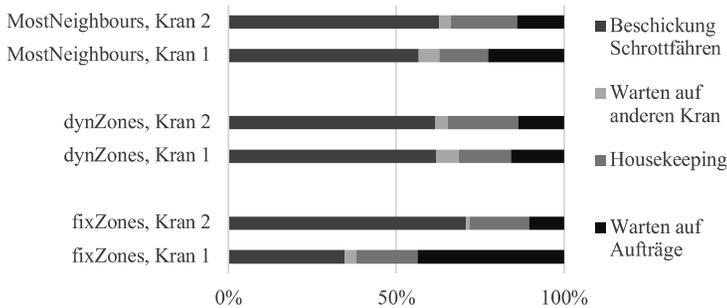


Abbildung 6
Simulation

5 Zusammenfassung und Ausblick

In diesem Beitrag wurde aufgezeigt, wie eine Zweikransteuerung mit Hindernisumfahrung in der Simulationssoftware *AutoMod* implementiert werden kann. Dadurch wird die Überprüfung von Steuerungsstrategien bzw. -algorithmen in stochastisch geprägten Anwendungsfällen möglich, da in überschaubarer Zeit eine Vielzahl von Transportspielen simuliert werden kann. Durch die Entkopplung der Steuerungs- und Ausführungsebene kann die Simulation auch als Emulation zur Bewertung von Kransteuerungen genutzt werden. Dafür wird eine Vereinfachung bzw. Modularisierung des Simulationsmodells angestrebt, wodurch eine Anwendung des Modells als *Prüfstand* für Mehrkransteuerungssysteme ermöglicht werden soll. Neben dem Testen von Kransteuerungen vor Inbetriebnahme könnte damit erstmalig eine Vergleichbarkeit von und mit den in der Literatur beschriebenen Mehrkransteuerungsverfahren unter realbetriebsnahen Bedingungen erreicht werden.

Literatur

- Boysen, N. Briskorn, D.; Meisel, F: A generalized classification scheme for crane scheduling with interference. *European Journal of Operational Research* 258 (2017), S. 343–357.
- Briskorn, D., Emde, S. & Boysen, N. Cooperative twin-crane scheduling. *Discrete Applied Mathematics* 211 (2016), S. 40–57.
- Gharehgozli, A. H.; Vernooij, F. G.; Zaerpour, N.: A simulation study of the performance of twin automated stacking cranes at a seaport container terminal. *European Journal of Operational Research* 261 (2017), S. 108–128.
- Lim, A.; Rodrigues, B.; Xu, Z: Am-parallel crane scheduling problem with a non-crossing constraint. *Naval Research Logistics* 54 (2007), S. 115–127.
- Petering, M. E. H.; Wu, Y.; Li, W.; Goh, M.; de Souza, R: Development and simulation analysis of real-time yard crane control systems for seaport container transshipment terminals. *OR Spectrum* 31 (2009), S. 801–835.
- Schulze, F.; Könze, A.: Spieldauer bei Brückenkränen mit überlagerten Bedienbereichen. In: Krause, F.; Horn, P. (Hrsg.): *Von innovativer Krantechnik bis Virtual Reality: 16. Internationale Kranfachtagung* (2008).
- Speer, U., John, G. & Fischer, K. Scheduling Yard Cranes Considering Crane Interference. *Computational Logistics* 6971 (2011), S. 321–340.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Generierung realitätsnaher Testdaten für die Simulation von Produktionen

Generation of realistic test data for the simulation of productions

Martin Krockert, Marvin Matthes, Torsten Munkelt, Hochschule für Technik und
Wirtschaft Dresden, Dresden (Germany), martin.krockert@htw-dresden.de,
marvin.matthes@htw-dresden.de, torsten.munkelt@htw-dresden.de

Sven Völker, Technische Hochschule Ulm, Ulm (Germany), sven.voelker@thu.de

Abstract: Test data are required for the simulative validation of production planning and control (PPC) procedures. To comprehensively determine the performance of PPC methods, it is necessary to examine different productions based on corresponding test data. A generator for test data is presented, which generates realistic master data based on structure-describing characteristics of a production. In addition, a method for determining the intermediate arrival time of customer orders is presented, which ensures a desired utilization of the production, which the generated test data describe. The validation of the test data generation shows that already for relatively small productions test data are generated, whose structure-describing characteristics correspond to the given parameters. With the same parameterization, the test data generator can generate reproducible test data and provide it for the simulative test of any PPC procedure.

1 Einleitung

Für die simulative Validierung von Verfahren der Produktionsplanung und -steuerung (PPS) werden Testdaten benötigt. Testdaten können auf drei unterschiedliche Arten gewonnen werden (siehe Tabelle 1).

Jede dieser Vorgehensweisen zur Erstellung von Testdaten hat Vor- und Nachteile. Die Gewinnung von Testdaten aus Praxisdaten ist meist mit sehr aufwendiger Extraktion und Transformation verbunden. Zudem sind die Daten meist unvollständig, schlecht gepflegt oder aus Datenschutzgründen schlicht nicht für eine Veröffentlichung geeignet (D. Libes et al., 2017).

Tabelle 1: Arten der Testdatengewinnung in Anlehnung an (Thomas Döring et al., 1999)

	Praxisdaten	Daten aus Problembibliothek	Generierte Daten
Korrektheit	O	+	+
Vollständigkeit	O	+	+
Umfang	+	O	+
Festlegung von Verteilung	-	-	+
Realitätsnähe	+	-	O
Bereitstellungsaufwand	-	+	O
Datenschutzkonformität	-	+	+

- schlecht erfüllt; O wenig erfüllt; + gut erfüllt

Die Nutzung von Problembibliotheken ist dann hilfreich, wenn ein neues Verfahren auf eine konkrete Problemstellung angewendet und mit bestehenden Verfahren verglichen werden soll. Eine bekannte Bibliothek für Problemstellungen unterschiedlichster Art stellt die OR-Library von Beasley (1990) dar. Beasley stellt in der Bibliothek u.a. statische Testdatensätze für verschiedene Fertigungstypen einer Produktion zur Verfügung. Der Nachteil von Problembibliotheken ergibt sich aus der hohen Spezialisierung und dem eher geringen Umfang der Daten innerhalb eines Testdatensatzes. So bilden die Probleme nur einen kleinen Teil möglicher Probleme ab, sind zu speziell oder bilden, aufgrund des Auslassens von Nebenbedingungen, nicht genau das Problem ab, welches das zu testende Verfahren der PPS lösen soll.

Testdatengeneratoren erzeugen Testdaten hingegen einfach, vollständig und in beliebigem Umfang (Sven Völker et al., 2001). Sie können darüber hinaus nicht nur zur Validierung von Verfahren der Produktionsplanung und -steuerung, sondern auch zur Erzeugung von umfangreichen Testdatensätze genutzt werden u.a. für die Bereitstellung von Trainingsdaten für den Einsatz von künstlicher Intelligenz in der Produktion (D. Libes et al., 2017). Im Gegensatz zu Praxisdaten entsprechen die generierten Daten nur in ihren strukturbeschreibenden Kenngrößen den realen Praxisdaten. Es existieren bereits viele kommerzielle und nicht-kommerzielle Generatoren für Testdaten. Bekannte Generatoren sind u.a. Bogus in C# oder Faker in Python. Darüber hinaus gibt es viele dedizierte Lösungen zur Generierung von Testdaten, wie z.B. Mockaroo oder FINRA DataGenerator. Die Generatoren haben keinen spezifischen Anwendungsbezug und ermöglichen die Generierung von einfachen semantisch randomisierten und teilweise einfachen hierarchisch strukturierten Testdaten. Die Generierung komplexer Testdaten, wie sie in Form von aufeinander aufbauenden Produktstrukturen und Arbeitsplänen in der Produktion vorliegen, ist mit diesen Generatoren nicht möglich. Neben den generischen Testdatengeneratoren, existieren bereits Testdatengeneratoren im Umfeld der Produktion. Fernandes et al. (2020) stellt einen Generator von Problemszenarien vor, bei denen verschiedene Fertigungstypen mit variierbaren Eingabeparametern simuliert werden können. Bei diesen Simulationen werden jedoch keine Produktstrukturen generiert, womit sich der Generator auch nicht für die Validierung von verschiedenen Verfahren zur Produktionsplanung und -steuerung eignet. Darüber

hinaus gibt es Generatoren für Produktstrukturen aus dem Anwendungsgebiet des Product-Lifecycle-Management, welche auf Basis von CAD-Modellen, Produktstrukturen generieren (Adolphy et al., 2015). Die synthetische Generierung von Produktstrukturen auf Basis strukturbeschreibenden Kenngrößen einer Produktion ist mit diesen Generatoren nicht möglich.

Die bisherigen Testdatengeneratoren genügen unseren Anforderungen nicht, da sie keine Testdaten in Form von Stammdaten für die Simulation von Produktionen bereitstellen oder nur einzelne detaillierte Produktstrukturen erzeugen. Deshalb stellen wir eine Methode zur automatischen Generierung von Stammdaten vor. Die generierten Testdaten sollen keine semantische, sondern eine strukturelle Realitätsnähe aufweisen. Dazu werden durch den Testdatengenerator Erzeugnisstrukturen, Arbeitspläne, Maschinenstrukturen und Werkzeuge generiert, welche in ihren Bezeichnungen generisch sind, in ihren Strukturen jedoch den vorgegebenen strukturbeschreibenden Kenngrößen entsprechen, die auch aus Realdaten gewonnen werden können.

Des Weiteren stellen wir ein Verfahren zur Generierung von Bewegungsdaten (Kundenaufträgen) bereit, mit denen für die zuvor erstellten Stammdaten eine ausgewogene Auslastung der Produktion gewährleistet wird. Ziel ist die Bereitstellung eines Generators zum Erzeugen realitätsnaher und reproduzierbarer Testdaten für die Simulation von Produktionen und den Test von Verfahren der PPS für andere Forschende und uns.

Im folgenden Kapitel werden die zugrundeliegenden strukturbeschreibenden Kenngrößen eingeführt. Darauf aufbauend, beschreibt Kapitel drei die Vorgehensweise des Generators. Das anschließende Kapitel stellt die Validierung des Testdatengenerators vor. Dabei wird u.a. ausgewertet, wie gut die aus den Testdaten gewonnenen strukturbeschreibenden Kenngrößen den vorgegebenen entsprechen.

2 Strukturbeschreibende Kenngrößen einer Produktion

Die Daten einer Produktion werden in Stammdaten und Bewegungsdaten unterteilt. Stammdaten beschreiben u.a. Materialien, Stücklisten, Arbeitspläne und Maschinen. Bewegungsdaten werden aus Stammdaten erzeugt und umfassen beispielsweise Kundenaufträge und Produktionsaufträge. Die Generierung von Testdaten zielt auf Stammdaten ab. Zur Generierung müssen zunächst die folgenden strukturbeschreibenden Kenngrößen aus den realen Daten extrahiert oder bei unvollständiger Datenlage geschätzt werden.

2.1 Erzeugnisstruktur

Die Erzeugnisstruktur entspricht einem gerichteten, azyklischen Graphen, welcher in der Produktion als Gozinto-Graph bezeichnet wird (Kurbel, 2016). Die gerichteten Kanten (e) zwischen den als Knoten (v) repräsentierten Materialien stellen eine „geht ein in“-Beziehung zwischen Materialien dar. Das Kantengewicht beschreibt die Menge des Quellmaterials, welche in das Zielmaterial eingehen. Für den Testdatengenerator werden zur Beschreibung des Erzeugnisstruktur-Graphen zum einen die Anzahl der Knoten ohne ausgehende Kante ($V_{Endprodukt}$), der Eingangsgrad (Komplexitätsgrad) und der Ausgangsgrad (Mehrfachverwendungsgrad) angegeben. Der Längste Pfad von $V_{Endprodukt}$ zu $V_{Kaufteilen}$ ist die maximale

Erzeugnisstrukturtiefe. Die Anzahl der Knoten ohne eingehende Kante ($V_{Kaufteil}$) ergibt sich aus der Anzahl der Endprodukte sowie dem Verhältnis aus Mehrfachverwendungsgrad und Komplexitätsgrad.

$$V_{sales\ product} \leftarrow \{v \in V | e_G^-(v) = 0\} \quad (1)$$

$$V_{purchase\ product} \leftarrow \{v \in V | e_G^+(v) = 0\} \quad (2)$$

Der Mehrfachverwendungsgrad (MV) bzw. Ausgangsgrad des Graphen beschreibt, wie oft ein Produkt in anderen Produkten verwendet wird, wobei V die Menge an Produkten repräsentiert und $succ(v)$ die Menge der in das Produkt v eingehenden Produkte (Heinrich, 1987).

$$DM = \frac{\sum_{v \in V} |succ(v)|}{|V \setminus V_{sales\ product}|} \quad (3)$$

Der Komplexitätsgrad (KG) bzw. Eingangsgrad des Graphen gibt an, aus wie vielen anderen Produkten ein Produkt durchschnittlich besteht, wobei $pred(v)$ die Menge der Produkte repräsentiert, die in das Produkt p eingehen.

$$DC = \frac{\sum_{v \in V} |pred(v)|}{|V \setminus V_{purchase\ product}|} \quad (4)$$

Die maximale Erzeugnisstrukturtiefe (EST) beschreibt die Anzahl der Kanten des längsten gerichteten Pfades im Erzeugnisstrukturgraphen. Darüber hinaus werden der Erwartungswert und die Standardabweichung für das Kantengewicht $w_{a,b}$ im Erzeugnisstrukturgraphen angegeben, wobei das Kantengewicht die Menge eines in Produkt a eingehenden Materials b angibt.

$$EST = \max(est(V_{sales\ product})) \quad (5)$$

2.2 Maschinenstrukturen

Für jede zu generierende Maschinengruppe (m) werden die Anzahl der Maschinen (r), der Erwartungswert (μ) und die Standardabweichung (σ) für die Bearbeitungszeit der auf Maschinen dieser Maschinengruppe durchzuführenden Arbeitsgänge angegeben. Zusätzlich umfassen strukturbeschreibende Kenngrößen für Maschinenstrukturen auch die Anzahl der Werkzeuge (w) pro Maschinengruppe und eine Angabe zur Rüstzeit (ϑ).

$$m = (r, \mu, \sigma, w, \vartheta)^T \quad (6)$$

2.3 Arbeitsplanstrukturen

Arbeitsplanstrukturen beschreibt der Organisationsgrad (OG) als strukturbeschreibende Kenngröße. Aufgrund des Organisationsgrades werden Arbeitspläne so erzeugt, dass sie einem bestimmten Fertigungstyp entsprechen.

Der Organisationsgrad (OG) ist ein Maß für den Fertigungstyp. Mithilfe des OGs kann stufenlos zwischen einzelnen Fertigungstypen einer Produktion, von Werkstatt- über Reihen- bis hin zu Fließfertigung unterschieden werden (Corsten und Gössinger, 2012). In der nachfolgenden Formel repräsentiert m die Anzahl der

Maschinengruppen und π_{ij} die Wahrscheinlichkeit des Übergangs von der Maschinengruppe i auf die Maschinengruppe j (Thomas Döring et al., 1999).

$$OG = \frac{1}{M-1} \sum_{i=1}^M \sum_{j=1}^M \left(\pi_{ij} - \frac{1}{M} \right)^2 \tag{8}$$

3 Konzept für den Testdatengenerator

Abbildung 1 zeigt den grundlegenden Ablauf zur Generierung der Testdaten. Die Generierung der Testdaten ist in zwei Abschnitte untergliedert.

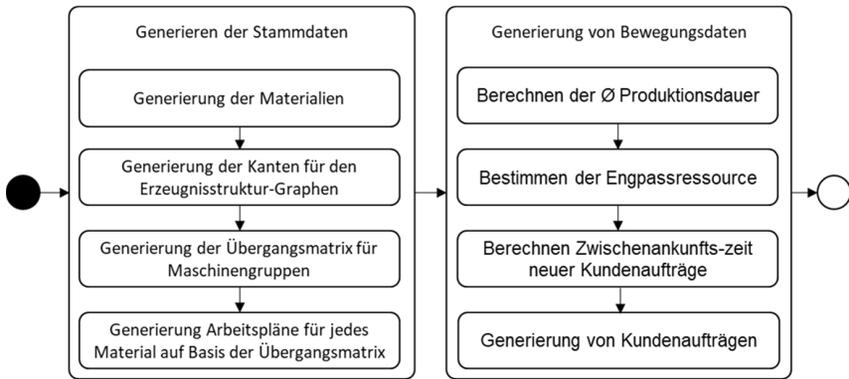


Abbildung 1: Ablauf der Testdatengenerierung

3.1 Generierung der Stammdaten

Auf Basis der strukturbeschreibenden Kenngrößen, KG , MV und EST , generiert Algorithmus 1 im ersten Schritt der Generierung die Materialien bzw. Knoten des Erzeugnisstrukturgraphen.

Algorithmus 1: Generierung der Produkte

```

begin
    V ← ∅
    for est = 1 to EST
        for i = 0 to ⌊(KG/MV)est-1⌋
            V ← V ∪ {vest,i}
        end for
    end for
End
  
```

Bei einer konvergierenden Struktur ($MV < KG$) beginnt die Generierung mit dem Hinzufügen der Kanten bei $V_{Endprodukt}$ (siehe Algorithmus 2) und bei einer divergierenden Struktur bei $V_{Kaufteil}$ (siehe Algorithmus 3). Es wird sichergestellt, dass jeder Knoten v im Erzeugnisstrukturgraphen G wenigstens über eine Kante e mit einem anderen Knoten verknüpft ist.

Algorithmus 2: Generierung der initialen Kanten für konvergierende Strukturen

```

begin
  for i = 1 increment to EST - 1
    for each k in  $V_i$ 
       $E \leftarrow E \cup \{(k, \text{beliebiges } \{v \in V_{i+1} \mid e_G^-(v) = 0\}, k)\}$ 
    end for
  end for
  for i = EST decrement to 2
    for each k in  $\{v \in V_i \mid e_G^-(v) = 0\}$ 
      beliebiges j mit  $p(j) = \frac{2^j}{i(i-1)}$  mit  $1 \leq j < i$ 
       $E \leftarrow E \cup \{(k, \text{beliebiges } \{v \in V_j\})\}$ 
    end for each
  end for
end

```

Algorithmus 3: Generierung der initialen Kanten für divergierende Strukturen

```

begin
  for i = EST decrement to 2
    for k in  $V_i$ 
       $E \leftarrow E \cup \{(k, \text{beliebiges } \{v \in V_{i-1} \mid e_G^+(v) = 0\})\}$ 
    end for each
  end for
  for i = 1 increment to EST - 1
    for each k in  $\{v \in V_i \mid e_G^+(v) = 0\}$ 
      beliebiges j mit  $p(j) = \frac{2^{j-1}}{(EST-i)(EST-i+1)}$  mit  $1 \leq j < i$ 
       $E \leftarrow E \cup \{(\text{beliebiges } \{v \in V_j\}, k)\}$ 
    end for each
  end for
end

```

Anschließend erzeugt Algorithmus 4 in Abhängigkeit vom KG und MV weitere Kanten. Die dabei entstehenden Multikanten zwischen zwei Knoten werden durch den Algorithmus bewusst nicht ausgeschlossen, da Multikanten auch in realen Produktstrukturen vorhanden sind.

Algorithmus 4: Generierung der zusätzlichen Kanten

```

begin
  for j =  $|E| + 1$  increment to  $\max(|V \setminus V_{\text{Endprodukt}}|, |KG \cdot |V \setminus V_{\text{Kaufteil}}|)$ 
     $v_{i,j} \leftarrow \text{beliebiges } \{v \in V \setminus V_{\text{Endprodukt}}\}$ 
    beliebiges k mit  $p(k) = \frac{2^{-k}}{i(i-1)}$  mit  $1 \leq k < i$ 
     $E \leftarrow E \cup \{(v_{i,j}, \text{beliebiges } \{v \in V_k\})\}$ 
  end for
end

```

Algorithmus 5 erzeugt eine Maschinenübergangsmatrix Π für die zu generierenden Arbeitspläne. Im ersten Schritt erzeugt er eine erweiterte $(|M| + 1) \times (|M| + 1)$ -Übergangsmatrix Π^A , bei welcher die Wahrscheinlichkeit π_{ij} den Übergang von Maschinengruppe m_i zu Maschinengruppe m_j beschreibt. Die Matrix ist insofern erweitert, als dass sie eine erste Zeile für die Quelle ein und eine letzte Spalte für die

Senke enthält. Eine $(|M| + 1) \times (|M| + 1)$ -Hilfsmatrix Π^B , wird mit $OG^B = 0$ generiert, wenn OG^A größer als der gewählte OG ist. Umgekehrt wird bei einem OG^A kleiner als der gewählte OG eine Hilfsmatrix Π^B mit $OG^B = 1$ erzeugt. Mittels eines Halbierungsverfahrens und der Hilfsmatrix Π^B , wird die Matrix Π^A angenähert, bis OG^A dem gewählten OG mit einer zuvor gewählten Mindestgenauigkeit übereinstimmt.

Algorithmus 5: Algorithmus zur Generierung der Übergangsmatrix für Maschinengruppen

```

begin
   $\Pi^A := \pi_{ij} \leftarrow \begin{cases} \frac{\lambda^{(j-i)}}{(j-i)!} & i \leq j \\ \frac{\lambda^{(i-j)}}{2 \cdot (i-j)!} & i > j \end{cases} \text{ mit } 0 < \lambda \leq 1$ 
   $\pi_i \leftarrow \sum_{j=1}^M \pi_{ij}$ 
   $\Pi^A := \pi_{ij} \leftarrow \frac{\pi_{ij}}{\pi_i}$ 
  if  $OG^A > OG$  then
     $\Pi^B := \pi_{ij} \leftarrow \left(\frac{1}{M}\right)_{(|M|+1) \times (|M|+1)}$ 
  else
     $\Pi^B := \pi_{ij} \leftarrow \begin{cases} 1 & j = i \\ 0 & \text{sonst} \end{cases}$ 
  end if
  while  $|OG^A - OG| > 0,001$ 
     $\Pi^C := \frac{1}{2} (\Pi^A + \Pi^B)$ 
    if  $OG^A - OG = OG^C - OG$  then
       $\Pi^A \leftarrow \Pi^C$ 
    else
       $\Pi^B \leftarrow \Pi^C$ 
    end if
  end while
   $\Pi \leftarrow \Pi^A$ 
end

```

Unter Anwendung der Maschinenübergangsmatrix Π erzeugt der Generator für alle Materialien eine Sequenz von Arbeitsplanpositionen. Für jede Arbeitsplanposition wird basierend auf dem Erwartungswert $\mu(m)$ und der Standardabweichung $\sigma(m)$ der gewählten Maschinengruppe eine Bearbeitungszeit ermittelt. Zusätzlich zur Maschinengruppe wird der Arbeitsplanpositionen ein Werkzeug aus dem Werkzeugpool der jeweiligen Maschinengruppe zugewiesen. Die Zuweisung der Werkzeuge erfolgt gleichverteilt. Mit der Erstellung der Arbeitspläne ist die Generierung der Stammdaten abgeschlossen.

3.2 Generierung von Bewegungsdaten

Nun werden für die generierten Stammdaten Bewegungsdaten erzeugt. Der Generator errechnet eine geeignete Zwischenankunftszeit für Kundenaufträge, um eine vorgegebene Auslastung u (z.B. 80%) auf den Engpassressourcen zu erzeugen. Die Berechnung erfolgt auf Basis der eingegebenen strukturbeschreibenden Kennzahlen. Mithilfe der Zwischenankunftszeit können ein kontinuierlicher Kundenauftragsstrom

oder für einen vorgegebenen Zeitraum alle Kundenaufträge erstellt werden. Formel 9 ermittelt die durchschnittliche Anzahl der Arbeitsplanpositionen AP je $V_{Endprodukt}$ über die gesamte Erzeugnisstruktur. Auf deren Basis lassen sich sowohl die zu erwartende Rüstzeit je Maschinengruppe $t_{Rüsten}(m)$ (Formel 10) und die erwartende Bearbeitungszeit je Maschinengruppe $t_{Bearbeiten}(m)$ (Formel 11) bestimmen. Mittels Formel 12 wird anhand des vorhandenen Kapazitätsangebots und der Verwendungshäufigkeit der minimale θ -normierte Durchsatz d über alle Maschinengruppen ermittelt. Das θ entspricht der verwendeten Zeiteinheit. Nach der Ermittlung des minimalen Durchsatzes d , wird in Formel 13, unter Berücksichtigung der angestrebten Auslastung u , ein λ für eine exponentialverteilte Zwischenankunftszeit für Kundenaufträge bestimmt.

$$AP = \frac{MV \cdot V \setminus V_{Endprodukt} \cdot w_{a,b} \cdot |M|}{V_{Endprodukt}} \quad (9)$$

$$t_{Rüsten}(m) = \min\left(\frac{AP}{w(m)}; \frac{AP}{|M|}\right) \cdot \vartheta(m) \quad (10)$$

$$t_{Bearbeiten}(m) = \frac{AP}{|M|} \cdot \mu(m) \quad (11)$$

$$d = \min_{m \in M} \frac{\theta \cdot r(m)}{t_{Rüsten}(m) + t_{Bearbeiten}(m)} \quad (12)$$

$$\lambda = \frac{1}{\theta \cdot \frac{d}{u}} \quad (13)$$

4 Validierung der generierten Testdaten

Für die Validierung der Testdaten wurden 450 Testdatensätze generiert. Dabei wurde für jede Strukturtiefe zwischen 2 und 10 eine Menge von Endprodukten erstellt. Die Menge der erzeugten Endprodukte wurde dabei ebenfalls stufenweise von 1 auf 50 erhöht. Die erzeugten Strukturen wurden nachfolgend analysiert und mit den Eingaben (= Erwartungswerten) hinsichtlich Mehrfachverwendungsgrad, Komplexitätsgrad und Organisationsgrad verglichen. Abbildung 2 zeigt die absolute Abweichung der erstellten Testdaten von den Vorgaben. Es ist gut zu erkennen, dass ab einer Strukturtiefe von 4 und einer Anzahl von 10 generierten Endprodukten die Abweichungen vom vorgegebenen Komplexitätsgrad und Mehrfachverwendungsgrad unter 5 % liegen. Damit die Abweichungen vom vorgegebenen Organisationsgrad ebenfalls gering ausfallen, benötigt man allerdings tiefere Produktstrukturen und/oder mehr Produkte.

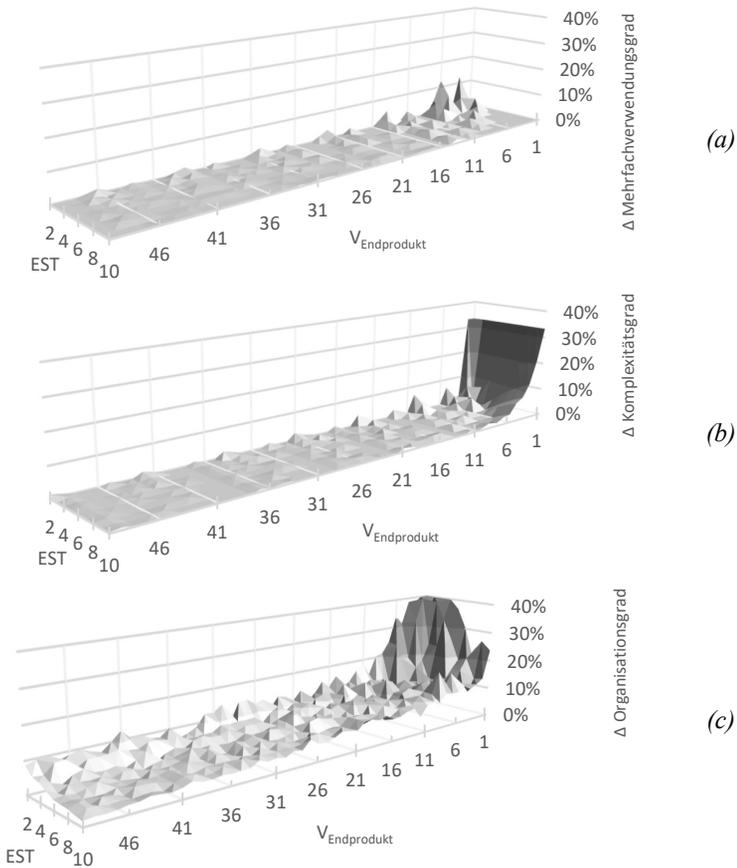


Abbildung 2: Abweichung der strukturbeschreibenden Kenngrößen

5 Zusammenfassung und Ausblick

Ziel des hier vorgestellten Generators ist das Generieren von Testdaten für die simulative Validierung von Produktionsplanung und -steuerung. Die generierten Testdaten entsprechen in ihren strukturbeschreibenden Kenngrößen bestimmten Vorgaben. Dabei können die vorzugebenden strukturbeschreibenden Kenngrößen durch Analyse realer Produktionsdaten ermittelt oder zur experimentellen Erprobung verschiedener Produkt- und Produktionsstrukturen manuell gesetzt werden. Durch die Verwendung gleicher Startwerte für die Zufallszahlengenerierung können identische Testdaten erzeugt und auf beliebige Zielsysteme übertragen werden. Für eine große Anzahl von Kombinationen aus strukturbeschreibenden Kenngrößen sind Testdaten generiert worden. Anhand dieser Testdaten wurden die Vorgehensweise bei der Testdatengenerierung evaluiert und die Grenzen des Testdatengenerators aufgezeigt, ab denen er Testdaten generiert, die in ihren strukturbeschreibenden Kenngrößen den Vorgaben entsprechen. Das vorgestellte Vorgehen bietet Raum für Erweiterungen. Beispielsweise könnten Verwendungshäufigkeit von Werkzeugen mit anderen

Verteilungen berücksichtigt werden. Auch ist es denkbar, die Einflüsse eines Produktmix bei der Generierung der Kundenaufträge zu berücksichtigen.

Danksagung

Das Forschungsvorhaben wird im Rahmen des Förderprogramms „Forschung an Fachhochschulen“ durch das Bundesministerium für Bildung und Forschung gefördert (Zeichen: 13FH133PX8).

Literatur

- Adolphy, S.; Grosser, H.; Kirsch, L.; Stark, R.: Method for Automated Structuring of Product Data and its Applications. *Procedia CIRP* 38 (2015), S. 153–158.
- Beasley, J.E., 1990: OR-Library. Online verfügbar unter <http://people.brunel.ac.uk/~mastjib/jeb/info.html>, zuletzt aktualisiert am February 2018, zuletzt geprüft am 29.04.2021.
- Corsten, H.; Gössinger, R.: *Production management (Produktionswirtschaft): Introduction to industrial production management (Einführung in das industrielle Produktionsmanagement)*. München: Oldenbourg 2012.
- D. Libes; D. Lechevalier; S. Jain: Issues in synthetic data generation for advanced manufacturing. In: 2017 IEEE International Conference on Big Data (Big Data), 2017, S. 1746–1754.
- Fernandes, E.C.; Santos, L.I.d.; Camatti, J.A.; Brown, L.; Borsato, M.: Flexible Production Data Generator for Manufacturing Companies. *Procedia Manufacturing* 51 (2020), S. 1478–1484.
- FINRA DataGenerator, 2016: The DataGenerator. Online verfügbar unter <https://finraos.github.io/DataGenerator/>, zuletzt aktualisiert am 12.05.2016, zuletzt geprüft am 04.05.2021.
- Heinrich, C.E.: *Mehrstufige Losgrößenplanung in hierarchisch strukturierten Produktionsplanungssystemen*. Berlin, Heidelberg: Springer Berlin Heidelberg 1987.
- Mockaroo, 2021: Mockaroo: Random Data Generator and API Mocking Tool | JSON / CSV / SQL / Excel. Online verfügbar unter <https://finraos.github.io/DataGenerator/>, zuletzt aktualisiert am 03.05.2021, zuletzt geprüft am 03.05.2021.
- Sven Völker; Thomas Döring; Torsten Munkelt: The Generation of Large Test Data for the Empirical Analysis of Heuristic Procedures for Production Planning and Control. In: Bernhard Fleischmann; Rainer Lasch; Ulrich Derigs; Wolfgang Domschke; Ulrich Rieder (Hrsg.): *Operations Research Proceedings 2000. Selected Papers of the Symposium on Operations Research (OR 2000)*, Dresden, September 9-12, 2000, 2001, S. 266–270.
- Thomas Döring; Torsten Munkelt; Sven Völker: Generierung komplexer Testdaten zur statistischen Analyse von Verfahren der Produktionsplanung und -steuerung. In: Martin Bösel (Hrsg.): *Amtliche und Nichtamtliche Statistiken - 12*. Ilmenauer Wirtschaftsforum, Tagungsband, 1999, S. 34–46.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Remote-Lernfabrik – Simulationsmodelle im Anlagenentstehungsprozess und deren Integration in die Virtuelle Inbetriebnahme

Remote Learning Factory – Simulation Models During the Plant Engineering Process and Their Integration into Virtual Commissioning

Benjamin Bastian Massow¹, Thomas Hausberger¹, Dominique Mathäus Geiger¹,
Matthias Klotz¹, Andreas Schlegel², Matthias Putz^{2,3},
benjamin.massow@mci.edu, thomas.hausberger@mci.edu, dm.geiger@mci4me.at,
m.klotz@mci4me.at, andreas.schlegel@iwu.fraunhofer.de,
matthias.putz@iwu.fraunhofer.de

¹Department Mechatronik, MCI - Die Unternehmerische Hochschule, Innsbruck
(Austria)

²Fraunhofer-Institut für Werkzeugmaschinen und Umformtechnik (IWU), Chemnitz
(Germany)

³Institut für Werkzeugmaschinen und Produktionsprozesse (IWP), Technische
Universität Chemnitz, Chemnitz (Germany)

Abstract: To meet the increasing demand for automated plants, engineers must learn how to utilize methods such as model-based engineering and virtual commissioning, where project- and experience-based knowledge transfer – as provided by learning factories – is identified as a key success factor in teaching a holistic understanding and problem-solving abilities. This paper presents an approach of a hybrid remote learning factory based on the concept of virtual commissioning, which removes the spatial and temporal access limitations of physical learning factories and extends the possibilities of existing virtual and hybrid approaches. In addition to intended competencies of their respective disciplines, engineers apply three principles using virtual commissioning, namely (i) the efficient reuse of simulation models for virtual commissioning, (ii) utilizing the principle of virtual commissioning during plant operation – in addition to plant engineering –, and (iii) a continuous adjustment of behavior models and kinematical models using data of the plant in operation.

1 Einleitung

Der Trend zu individuellen Produkten und kürzeren Produktlebenszyklen stellt Anlagen- und Sondermaschinenbauer sowie -betreiber vor Herausforderungen. Neue Anlagen müssen schneller entworfen und in Betrieb genommen, Bestandsanlagen

angepasst und optimiert werden – hierbei haben flexibel automatisierte Anlagen einen großen Einfluss auf die Leistungsfähigkeit eines Unternehmens (Lin et al., 2019). Der fortschreitende Bedarf an automatisierten Anlagen hat insbesondere auch Auswirkungen auf die Nachfrage an Fachkompetenzen von Ingenieuren, wie z.B. zu effizientem Entwurf und Entwicklung von Technologien, zu Automatisierungstechnik und Robotik sowie zum Projektmanagement (World Economic Forum, 2020). Dies führt zu einer Veränderung des benötigten Kompetenzprofils von Ingenieuren – was durch die fortschreitende Globalisierung und den demografischen Wandel weiter verstärkt wird (Abele et al., 2017). Neben Fachkompetenzen in ihren entsprechenden Disziplinen benötigen Ingenieure ein ganzheitliches Verständnis des Anlagenentwicklungsprozesses und aller Teildisziplinen (Hernandez-de-Menendez et al., 2020) sowie adäquater Entwicklungsmethoden.

Das Konzept der Lernfabrik hat sich hier als ein probates Mittel zur Aus- und Weiterbildung von Ingenieuren erwiesen. Hierbei haben sich insbesondere (i) ein projekt- und erfahrungsbasierter Wissenstransfer (Abellán-Nebot, 2020; Möller et al., 2020; Kovalenko et al., 2020; Pittich et al., 2020), eine kooperative Lernumgebung, (iii) interdisziplinäre und funktionsübergreifende Teams und (iv) die Bereitstellung effizienter (Arbeits-)Methoden (Vila et al., 2017; Madsen und Møller, 2017) als Schlüsselfaktoren herausgestellt. Die aktuell fortschreitende COVID-19-Pandemie hat nicht nur generelle negative Effekte auf die Aus- und Weiterbildung von Ingenieuren (Ożadowicz, 2020; Qadir und Al-Fuqaha, 2020), sondern auch auf die Verfügbarkeit und Zugänglichkeit von physischen Lernfabriken. Somit gewinnen virtuelle Lernfabriken und Labore mit Fernzugriff (Grodzki et al., 2018; Trentsiotis et al., 2020; Cimino et al., 2019) an Bedeutung. Wohingegen bei physischen Lernfabriken mit Fernzugriff die Zugangszeiten bzw. die Anzahl der Zugriffe beschränkt ist, können virtuelle Lernfabriken in einer größeren Anzahl verfügbar gemacht werden – durch fehlende reale Komponenten können hierbei jedoch nicht alle Fachkompetenzen praxisnah vermittelt sowie Ergebnisse nicht validiert werden. Ansätze mit einer hybriden Struktur aus realen und simulierten Komponenten (Orduña et al., 2018) fokussieren sich auf Teilkomponenten und -disziplinen der Anlagenentwicklung, betrachten jedoch nicht gesamte Produktionsanlagen, lassen sich nicht selbstständig erweitern und vermitteln somit nicht den Zusammenhang der Fachdisziplinen. Lernfabriken auf Basis des Konzeptes der Virtuellen Inbetriebnahme (VDI/VDE 3693 Blatt 1:2016-08; VDI/VDE 3693 Blatt 2:2018-12) ermöglichen die Betrachtung gesamter Produktionsanlagen, wobei jedoch insbesondere der Zeitaufwand und die Komplexität der zu erstellenden Kinematik- und Verhaltensmodelle (Lechler et al., 2019; Lee und Park, 2014) sowie deren aufwändige Integration Teilnehmer an der selbstständigen Erstellung der hierfür nötigen digitalen Anlagenmodelle und somit an der Zusammenführung verschiedener Fachdisziplinen hindert. Dies führt zu einer Einschränkung des Potentials von Lernfabriken auf Basis der Virtuellen Inbetriebnahme, welche somit häufig auf die Vermittlung einzelner Fachkompetenzen, etwa im Bereich der Automatisierungstechnik, begrenzt sind.

Dieser Beitrag präsentiert das Konzept einer hybriden Remote-Lernfabrik auf Basis des Konzeptes der Virtuellen Inbetriebnahme und beinhaltet ein reales sowie ein entsprechendes digitales Anlagenmodell im Labormaßstab. Die Struktur dieses Ansatzes ermöglicht es Teilnehmern, das digitale Anlagenmodell selbstständig zu erweitern bzw. zu adaptieren sowie (selbst-)entwickelte Simulationsmodelle als Kinematik- und Verhaltensmodelle zu integrieren. Somit können neben dem Konzept

der Virtuellen Inbetriebnahme verschiedene zur Entwicklung von Produktionsanlagen nötige Fachkompetenzen – u.a. in den Bereichen Mechatronik, Regelungstechnik, Simulation und Modellierung, Automatisierungstechnik, Digitale Fabrik, Robotik, Datenerfassung und -analyse sowie Digitaler Zwilling – sowie modellbasierte Entwicklungsmethoden vermittelt werden. Um Teilnehmern ein innovatives Verständnis der (i) erweiterten Einsatzmöglichkeiten sowie der (ii) ganzheitlichen Integration der Virtuellen Inbetriebnahme in die modellbasierte Entwicklung sowie den Betrieb von Produktionsanlagen zu vermitteln, wenden diese folgende drei Grundsätze

- Effiziente Wiederverwendung von Simulationsmodellen aus der Anlagenentwicklung zur Erstellung des digitalen Anlagenmodells
- Verwendung des digitalen Anlagenmodells außerhalb der Inbetriebnahme, insbesondere zu Adaption, Umbau und Optimierung der realen Anlage
- Kontinuierliche Anpassung des digitalen Anlagenmodells bzw. der Simulationsmodelle durch Daten der realen Anlage

auf den Ansatz des digitalen Anlagenmodells bzw. auf die erweiterte Virtuelle Inbetriebnahme an. Durch die hybride Struktur und die Möglichkeit des Fernzugriffs sind Teilnehmer in der praktischen Anwendung von Lerninhalten nicht räumlich oder zeitlich begrenzt – somit werden diese in der Erfassung von Zusammenhängen der Fachdisziplinen bzw. der Lerninhalte kontinuierlich unterstützt.

Nach der Vorstellung des Konzeptes der Remote-Lernfabrik (Kapitel 2) wird auf die technische Umsetzung der hybriden Struktur eingegangen (Kapitel 2.1), die effiziente Wiederverwendung von Simulationsmodellen zur Erstellung des digitalen Anlagenmodells dargestellt (Kapitel 2.2) und dessen erweiterter Einsatz im Anlagenentstehungsprozess und -betrieb erläutert (Kapitel 2.3). Abschließend werden eine Zusammenfassung und ein Ausblick gegeben (Kapitel 3).

2 Konzept der Remote-Lernfabrik auf Basis der Virtuellen Inbetriebnahme

Um die für Anlagenplanungs- und -entstehungsprozesse sowie für Anlagenbetrieb und -evolution benötigten Fachkompetenzen ganzheitlich – d.h. inklusive deren Schnittstellen und Zusammenhängen – zu vermitteln, werden diese anhand eines realen sowie entsprechendem digitalen Anlagenmodells (Abbildung 1) angewendet.

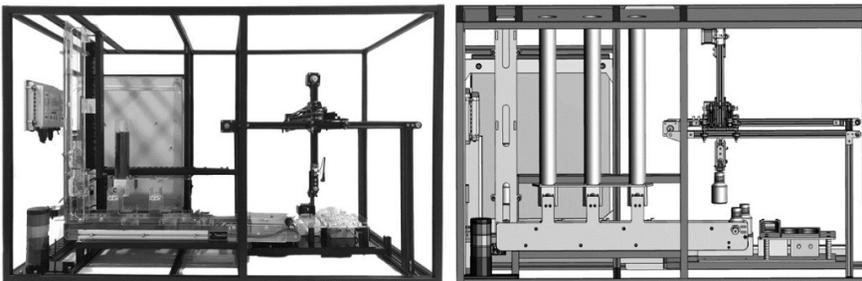


Abbildung 1: Reales Anlagenmodell (links) sowie entsprechendes digitales Anlagenmodell (rechts) der Remote-Lernfabrik.

Beide Anlagenmodelle stellen eine Abfüllanlage für Granulat im Labormaßstab dar und enthalten ein Behälterlager inklusive Vereinzlung, ein Identifikationssystem auf Basis von Radio-Frequency Identification (RFID), ein Fließband, Dosiereinrichtungen, einen kartesischen Roboter, eine Qualitäts- bzw. Gewichtskontrolle, eine Wärmebehandlungsstation sowie ein Lagersystem. Über eine Webschnittstelle oder ein Human-Machine Interface (HMI) auf Basis eines Touchscreens können Aufträge zu individuellen Granulatmischungen erstellt werden, wobei Auftragsverwaltung, Produktkonfigurationen und Fertigungsablaufpläne über ein Manufacturing Execution System (MES) verwaltet werden. Behälter und Aufträge werden über das RFID-Identifikationssystem verfolgt bzw. identifiziert und zusammen mit Qualitäts- und Gewichtsdaten abgespeichert.

Die Anwendung des erweiterten Konzeptes der Virtuellen Inbetriebnahme auf das digitale Anlagenmodell stellt das zentrale Element des Konzeptes der Remote-Lernfabrik dar. Nur mit diesem sind viele Inhalte bzw. Fachkompetenzen ohne nötige Umbaumaßen, ohne die Gefahr der Beschädigung sowie ohne direkten physischen Zugriff auf das reale Anlagenmodell zu vermitteln. Neben der Vermittlung von Fachkompetenzen zur modellbasierten Entwicklung von Anlagen werden drei Grundsätze durch die Anwendung auf die Virtuelle Inbetriebnahme bzw. das digitale Anlagenmodell vermittelt:

- 1. Effiziente Wiederverwendung von Simulationsmodellen aus der Anlagenentwicklung zur Erstellung des digitalen Anlagenmodells:** Als Hindernis zum breiten Einsatz der Methode der Virtuellen Inbetriebnahme wird insbesondere der Zeitaufwand und die Komplexität der zu erstellenden Kinematik- und Verhaltensmodelle gesehen (Lechler et al., 2019; Lee und Park, 2014). Daher werden den Teilnehmern – im weiteren Sinne der Digitalen Fabrik (Matysczok und Seewaldt, 2019; Bracht et al., 2018) – Möglichkeiten der durchgängigen Verwendung von Simulationsmodellen nähergebracht. Hierzu wenden diese insbesondere Methoden der automatischen Codegenerierung an.
- 2. Verwendung des digitalen Anlagenmodells außerhalb der Inbetriebnahme, insbesondere zu Adaption, Umbau und Optimierung der realen Anlage:** Aktuell wird das Konzept der Virtuellen Inbetriebnahme insbesondere zur Vorbereitung der Inbetriebnahme von Anlagen und nur vereinzelt auch zu anderen Zwecken im Anlagenentstehungsprozess (Schauerte, 2017; Metzner et al., 2019) eingesetzt. Insbesondere der potenzielle Einsatz im Anlagenbetrieb wird nur vorgeschlagen und angedeutet. Daher werden den Teilnehmern Möglichkeiten zur Verwendung der Virtuellen Inbetriebnahme bzw. des digitalen Anlagenmodells zur Planung und Entwicklung von Adaptionen und Erweiterungen im Betrieb befindlicher Anlagen nähergebracht. Hierzu wenden diese insbesondere die unter Punkt 1. genannte Wiederverwendung von Simulationsmodellen und eine effiziente Methode zur Implementierung des gleichen Steuerungsprogramms auf realer sowie digitaler Anlage an.
- 3. Kontinuierliche Anpassung des digitalen Anlagenmodells bzw. der Simulationsmodelle durch Daten der realen Anlage:** Zur der unter Punkt 2. genannten Verwendung des Konzeptes der Virtuellen Inbetriebnahme zur Planung und Entwicklung von Adaptionen und Erweiterungen im Betrieb befindlicher Anlagen muss die digitale Anlage das gleiche Verhalten wie die reale Anlage aufweisen. Da die Kinematik- und Verhaltensmodelle zumeist durch analytische Modellierung unter Verwendung theoretischer Parameter (etwa aus Datenblättern)

(Süß et al., 2016), werden die Verhaltens- bzw. Simulationsmodelle der Komponenten – mit entsprechend kürzeren Zykluszeiten – auch in der Echtzeit-Laufzeitumgebung ausgeführt. Die Visualisierung der digitalen Anlage läuft in der Laufzeitumgebung Unity und enthält kinematisierte Konstruktionsdateien, virtuelle Sensoren sowie – zur Simulation des Stauverhaltens – eine physikbasierte Simulation. Beide Laufzeitumgebungen laufen auf einer virtuellen Maschine und kommunizieren über das Automation Device Specification (ADS) Protokoll. Über diese Struktur kann eine sowohl (a) einfach für Teilnehmer zu bedienende als auch (b) echtzeitfähige Umsetzung der Virtuellen Inbetriebnahme realisiert werden.

Durch die hybride Struktur dieses Ansatzes können die Anwendungs- und Adaptionmöglichkeiten bestehender virtueller und hybrider Ansätze erweitert, hierdurch Probleme simuliert und u.a. eine problembasierte Lernstrategie (Tisch et al., 2016) angewendet werden. Hierbei wird (i) ein Problem in das digitale Anlagenmodell eingebracht – z.B. unvorhergesehene Störgrößen in Regelkreisen –, (ii) den Teilnehmern entsprechende modellbasierte Lösungsmethoden vermittelt, (iii) das Problem in der digitalen Anlage gelöst und auf der realen Anlage implementiert sowie (iv) die angewendeten Lösungsmethoden diskutiert und reflektiert.

2.2 Wiederverwendung von Simulationsmodellen aus der Anlagenentwicklung zur Erstellung des digitalen Anlagenmodells

Im weiteren Sinne des Begriffs der Digitalen Fabrik nimmt die Anwendung verschiedener (Simulations-)Modelle – sowie insbesondere deren Wiederverwendung zur Erstellung des digitalen Anlagenmodells – eine zentrale Rolle im Ansatz der Remote-Lernfabrik ein. Durch eine einfach zu bedienende Struktur und Werkzeugkette wird der Zeitaufwand und die Komplexität zur Integration der Kinematik- und Verhaltensmodelle für Teilnehmer stark herabgesetzt und ihnen dadurch die Möglichkeiten der durchgängigen Verwendung von Simulationsmodellen nähergebracht. Hierdurch erfahren sie – neben fachspezifischen modellbasierten methodischen Vorgehensweisen – das Innovationspotential ganzheitlicher modellbasierter Ansätze.

Um hierzu effektiv – bereits in der Anlagenentwicklung verwendete – Simulationsmodelle in das digitale Anlagenmodell integrieren zu können, wird ein besonderer Fokus auf die Verwendung von Methoden zur automatischen Codegenerierung gelegt. Hierzu werden die zur modellbasierten Entwicklung erstellten Simulationsmodelle in Verhaltens- oder Kinematikmodelle überführt. Mit Hilfe automatischer Codegenerierung werden diese in entsprechenden C/C++ Code bzw. entsprechende TwinCAT Object Model (TeCOM) überführt und innerhalb der TwinCAT 3 Echtzeit-Laufzeitumgebung ausgeführt. Hierdurch ist es Teilnehmern möglich, die Kinematik- und Verhaltensmodelle effektiv zu erstellen bzw. das digitale Anlagenmodell selbstständig zu ändern, zu erweitern und zu optimieren.

Abbildung 3 zeigt am Beispiel des modellbasierten Entwurfs der Positions- und Geschwindigkeitsregelung des Fließbandantriebes die Verwendung o.g. Vorgehensweise zur Wiederverwendung des dynamischen Fließbandmodells bzw. dessen Integration in die Virtuelle Inbetriebnahme. Die Teilnehmer setzen hierfür folgende Schritte um: (1) Analytische Modellierung des Fließbandes, (2) modellbasierter Reglerentwurf in MathWorks MATLAB, (3) Model-in-the-Loop

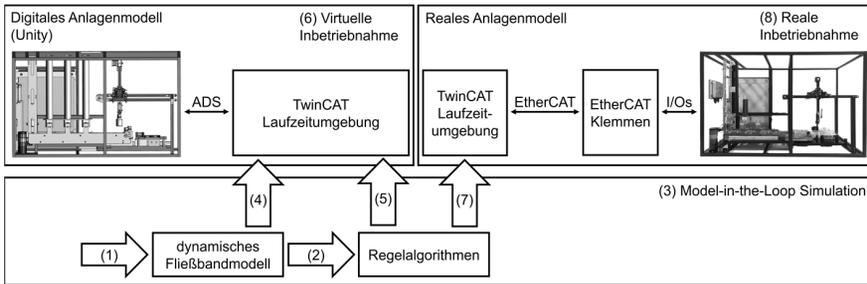


Abbildung 3: Wiederverwendung des zur modellbasierten Auslegung der Positions- und Geschwindigkeitsregelung erstellten dynamischen Fließbandmodells als Verhaltensmodell der Virtuellen Inbetriebnahme bzw. des digitalen Anlagenmodells.

(MiL) Simulation durch Implementierung des Fließbandmodells sowie der Positions- und Geschwindigkeitsregelung in MathWorks Simulink, (4) Überführung des Fließbandmodells in ein entsprechendes TcCOM und Implementierung der Verhaltensmodelle innerhalb der Echtzeit-Laufzeitumgebung des digitalen Anlagenmodells, (5) Überführung der Regelalgorithmen in entsprechende TcCOM und Integration in die zu testende Steuersoftware innerhalb der Echtzeit-Laufzeitumgebung, (6) Virtuelle Inbetriebnahme bzw. Hardware-in-the-Loop (HiL) Simulation und iterative Optimierung der Regelalgorithmen, (7) Integration der Regelalgorithmen in Steuersoftware bzw. Echtzeit-Laufzeitumgebung des realen Anlagenmodells und (8) Inbetriebnahme des realen Anlagenmodells bzw. Testen der Regelalgorithmen in Echtzeit gegen das reale Fließband über das EtherCAT Protokoll.

Abbildung 4 zeigt beispielhaft die Model-in-the-Loop (MiL) Simulation zum Test der modellbasierten Auslegung der Positions- und Geschwindigkeitsregelung des Fließbandes sowie die Wiederverwendung des Simulationsmodells als Verhaltensmodell in der Virtuellen Inbetriebnahme.

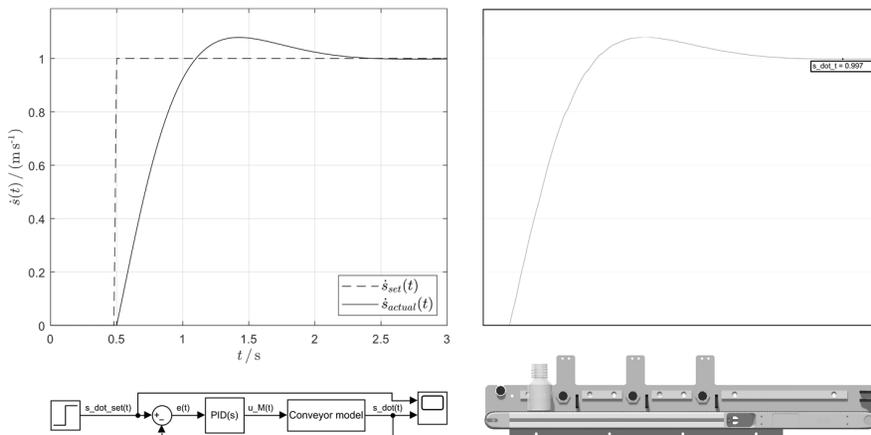


Abbildung 4: Wiederverwendung des Simulationsmodells des Fließbandes: Model-in-the-Loop (MiL) Simulation zur modellbasierten Auslegung der Positions- und Geschwindigkeitsregelung (links) sowie Integration als Verhaltensmodell der Virtuellen Inbetriebnahme bzw. des digitalen Anlagenmodells (rechts).

2.3 Verwendung des digitalen Anlagenmodells in Anlagenentwicklung und -betrieb

Abbildung 5 zeigt die erweiterte Verwendung der Virtuellen Inbetriebnahme in der Remote-Lernfabrik. Teilnehmer (1) integrieren Simulationsmodelle der modellbasierten Entwicklung als Verhaltens- und Kinematikmodelle in das digitale Anlagenmodell und verwenden dieses zur Virtuellen Inbetriebnahme, (2) modellbasiert entwickelte Komponenten werden auf dem digitalen Anlagenmodell getestet und in die reale Anlage integriert, (3) das Steuerungsprogramm wird am digitalen Anlagenmodell getestet, optimiert und an der realen Anlage implementiert, (4) Verhaltens- und Kinematikmodelle des digitalen Anlagenmodells werden wiederholt durch Daten der realen Anlage angepasst, um (5) Anlagenoptimierungen und (6) Anlagenänderungen am digitalen Anlagenmodell zu entwickeln, zu testen und wiederum an der realen Anlage zu implementieren. Somit wird Teilnehmern durch die Arbeit an der Remote-Lernfabrik insbesondere ein innovatives Verständnis der (i) Einsatzmöglichkeiten sowie der (ii) ganzheitlichen Integration der Virtuellen Inbetriebnahme in Entwicklung sowie Betrieb von Produktionsanlagen vermittelt.

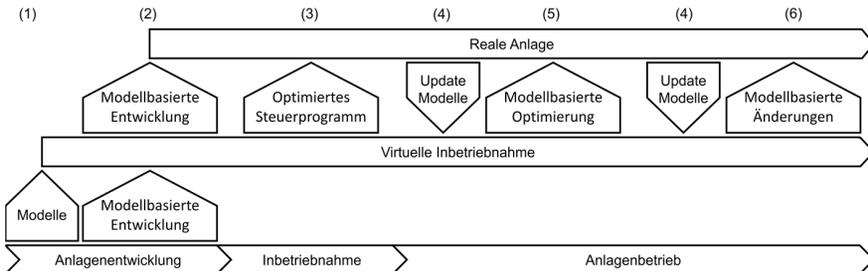


Abbildung 5: Verwendung der Virtuellen Inbetriebnahme in der Remote-Lernfabrik.

3 Zusammenfassung und Ausblick

Ingenieure im Anlagen- und Sondermaschinenbau benötigen ein ganzheitliches Verständnis aller beteiligten Fachkompetenzen sowie adäquater Entwicklungsmethoden. Hierbei hat sich das Konzept der Lernfabrik als ein probates Mittel zur Aus- und Weiterbildung erwiesen. Der vorgestellte Ansatz einer hybriden Remote-Lernfabrik auf Basis des Konzeptes der Virtuellen Inbetriebnahme hebt die räumlichen und zeitlichen Zugangsbegrenzungen physischer Lernfabriken auf und erweitert die Anwendungs- und Adaptionmöglichkeiten bestehender virtueller und hybrider Ansätze. Somit kann Ingenieuren durch die Anwendung von drei Grundsätzen – (i) effiziente Wiederverwendung von Simulationsmodellen zur Erstellung des digitalen Anlagenmodells, (ii) Verwendung des digitalen Anlagenmodells außerhalb der Inbetriebnahme sowie (iii) kontinuierliche Anpassung des digitalen Anlagenmodells durch Daten der realen Anlage – ein innovatives Verständnis der erweiterten Einsatzmöglichkeiten sowie der ganzheitlichen Integration des Konzeptes der Virtuellen Inbetriebnahme in die modellbasierte Entwicklung sowie den Betrieb von Produktionsanlagen vermittelt werden.

Zur breiteren Integration des Ansatzes der Virtuellen Inbetriebnahme sollten Konzepte zum standardisierten Austausch von Bauteil- und Komponentenbibliotheken entwickelt und in industrielle Softwareumgebungen integriert werden.

Literatur

- Abele, E.; Chryssoulouris, G.; Sihh, W.; Metternich, J.; ElMaraghy, H.; Seliger, G.; Sivard, G.; ElMaraghy, W.; Hummel, V.; Tisch, M.; Seifermann, S.: Learning factories for future oriented research and education in manufacturing. *CIRP Annals* 66 (2017) 2, S. 803–826.
- Abellán-Nebot, J.V.: Project-based experience through real manufacturing activities in mechanical engineering. *International Journal of Mechanical Engineering Education* 48 (2020) 1, S. 55–78.
- Bracht, U.; Geckler, D.; Wenzel, S.: Einleitung, Definition und Stand der Umsetzung sowie der Bezug zu Industrie 4.0. In: Bracht, U.; Geckler, D.; Wenzel, S. (Hrsg.): *Digitale Fabrik: Methoden und Praxisbeispiele*. Berlin, Heidelberg: Springer Berlin Heidelberg 2018, S. 1–25.
- Cimino, C.; Negri, E.; Fumagalli, L.: Review of digital twin applications in manufacturing. *Computers in Industry* 113 (2019), S. 103130.
- Grodotski, J.; Ortelt, T.R.; Tekkaya, A.E.: Remote and Virtual Labs for Engineering Education 4.0. *Procedia Manufacturing* 26 (2018), S. 1349–1360.
- Hernandez-de-Menendez, M.; Escobar Díaz, C.A.; Morales-Menendez, R.: Engineering education for smart 4.0 technology: a review. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 14 (2020) 3, S. 789–803.
- Kovalenko, I.; Balta, E.C.; Qamsane, Y.; Koman, P.D.; Zhu, X.; Lin, Y.; Tilbury, D.M.; Morley Mao, Z.; Barton, K.: Developing the Workforce for Next-Generation Smart Manufacturing Systems: A Multidisciplinary Research Team Approach. *Smart and Sustainable Manufacturing Systems* 5 (2020) 2.
- Lechler, T.; Fischer, E.; Metzner, M.; Mayr, A.; Franke, J.: Virtual Commissioning – Scientific review and exploratory use cases in advanced production systems. *Procedia CIRP* 81 (2019), S. 1125–1130.
- Lee, C.G.; Park, S.C.: Survey on the virtual commissioning of manufacturing systems. *Journal of Computational Design and Engineering* 1 (2014) 3, S. 213–222.
- Lin, B.; Wu, W.; Song, M.: Industry 4.0: driving factors and impacts on firm's performance: an empirical study on China's manufacturing industry. *Annals of Operations Research* (2019).
- Madsen, O.; Møller, C.: The AAU Smart Production Laboratory for Teaching and Research in Emerging Digital Manufacturing Technologies. *Procedia Manufacturing* 9 (2017), S. 106–112.
- Matysczok, C.; Seewaldt, M.: VDI Fachausschuss Digitale Fabrik - Darstellung der aktuellen Aktivitäten und Richtlinien. In: Putz, M.; Schlegel, A. (Hrsg.): *Simulation in Produktion und Logistik 2019*, 2019, S. 39–48.
- Metzner, M.; Krieg, L.; Merhof, J.; Ködel, T.; Franke, J.: Intuitive Interaction with Virtual Commissioning of Production Systems for Design Validation. *Procedia CIRP* 84 (2019), S. 892–895.
- Möller, D.P.; Jehle, I.A.; Hou, W.: Engineering Education in Intelligent Manufacturing. In: *2020 IEEE International Conference on Electro Information Technology (EIT)*, Chicago, IL, USA, 31.07.2020 - 01.08.2020, 2020, S. 7–12.

- Negri, E.; Fumagalli, L.; Macchi, M.: A Review of the Roles of Digital Twin in CPS-based Production Systems. *Procedia Manufacturing* 11 (2017), S. 939–948.
- Orduña, P.; Garcia-Zubia, J.; Rodriguez-Gil, L.; Angulo, I.; Hernandez-Jayo, U.; Dziabenko, O.; López-de-Ipiña, D.: The WebLab-Deusto Remote Laboratory Management System Architecture: Achieving Scalability, Interoperability, and Federation of Remote Experimentation. In: Auer, M.E.; Azad, A.K.M.; Edwards, A.; Jong, T. de (Hrsg.): *Cyber-Physical Laboratories in Engineering and Science Education*. Cham: Springer International Publishing 2018, S. 17–42.
- Ożadowicz, A.: Modified Blended Learning in Engineering Higher Education during the COVID-19 Lockdown—Building Automation Courses Case Study. *Education Sciences* 10 (2020) 10, S. 292.
- Pittich, D.; Tenberg, R.; Lensing, K.: Learning factories for complex competence acquisition. *European Journal of Engineering Education* 45 (2020) 2, S. 196–213.
- Qadir, J.; Al-Fuqaha, A.: A Student Primer on How to Thrive in Engineering Education during and beyond COVID-19. *Education Sciences* 10 (2020) 9, S. 236.
- Schauerte, H., 2017: Plug & Work Konzept der SMS group – Schnelle, professionelle Inbetriebnahme mit Hilfe von Prozess- und Produktionssimulation im klassischen Anlagenbau. *MATLAB EXPO 2017*, Juni 2017.
- Shen, W.; Hu, T.; Yin, Y.; He, J.; Tao, F.; Nee, A.: Digital twin based virtual commissioning for computerized numerical control machine tools. In: *Digital Twin Driven Smart Design*: Elsevier 2020, S. 289–307.
- Süß, S.; Magnus, S.; Thron, M.; Zipper, H.; Odefey, U.; Fassler, V.; Strahilov, A.; Klodowski, A.; Bar, T.; Diedrich, C.: Test methodology for virtual commissioning based on behaviour simulation of production systems. In: *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, Berlin, Germany, 06.09.2016 - 09.09.2016, 2016, S. 1–9.
- Tisch, M.; Hertle, C.; Abele, E.; Metternich, J.; Tenberg, R.: Learning factory design: a competency-oriented approach integrating three design levels. *International Journal of Computer Integrated Manufacturing* 29 (2016) 12, S. 1355–1375.
- Trentsios, P.; Wolf, M.; Frerich, S.: Remote Lab meets Virtual Reality – Enabling immersive access to high tech laboratories from afar. *Procedia Manufacturing* 43 (2020), S. 25–31.
- Vila, C.; Ugarte, D.; Ríos, J.; Abellán, J.V.: Project-based collaborative engineering learning to develop Industry 4.0 skills within a PLM framework. *Procedia Manufacturing* 13 (2017), S. 1269–1276.
- VDI/VDE 3693 Blatt 2:2018-12: Virtuelle Inbetriebnahme - Einführung der virtuellen Inbetriebnahme in Unternehmen.
- VDI/VDE 3693 Blatt 1:2016-08: Virtuelle Inbetriebnahme - Modellarten und Glossar.
- World Economic Forum: *The Future of Jobs Report 2020*. Geneva, Switzerland 2020.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Konzept zur Integration eines kontinuierlichen Materialflussmodells in die Virtuelle Inbetriebnahme durch Signalgenerierung

***Concept for the integration of a continuous material flow model in
virtual commissioning by generation of signals***

Annika Kienzlen, Alexander Verl, Institut für Steuerungstechnik der
Werkzeugmaschinen und Fertigungseinrichtungen, Universität Stuttgart, Stuttgart
(Germany), annika.kienzlen@isw.uni-stuttgart.de, alexander.verl@isw.uni-
stuttgart.de

Abstract: Simulation models for virtual commissioning of plants are of increasing importance in the digital factory. Considering whole plants, there are many piece goods whose movement has to be modelled. The existing models for the material flow are usually either too unprecise or too complex to be used for virtual commissioning. The complexity is leading to high computation times which exceed the short cycle of the control system. Therefore, a macroscopic flow model was developed in previous work which doesn't depict single goods but a continuous density distribution instead. Hence, it is effective for material flow intensive plants. However, it doesn't provide signals for the sensors of the production system and therefore cannot interact with the control system directly. This paper shows possibilities to compute signals from the density distribution to integrate the macroscopic flow model in virtual commissioning. Beforehand, the macroscopic flow model and signals in control engineering are presented.

1 Motivation

Mit steigender Komplexität von Anlagen und wachsendem Anteil an mechatronischen Komponenten, wird insbesondere das Steuerungssystem einer Anlage aufwendiger. Für die Inbetriebnahme eines Steuerungssystems müssen nicht nur einzelne Maschinen, sondern auch deren Interaktion berücksichtigt werden. Die Maschinen sind über den Materialfluss verbunden. Um ein Steuerungssystem in einer Anlage mit vielen Stückgütern auszulegen und zu testen, muss die genaue Bewegung der Stückgüter bekannt sein, die durch die Interaktion untereinander und mit der Umgebung entsteht. Beispielsweise in Anlagen zur Flaschenbefüllung bewegen sich die einzelnen Flaschen mit hoher Geschwindigkeit (Al-Hawari et al. 2010), trotzdem ist die exakte Be-

wegung der Flaschen entscheidend für den korrekten Prozess, der durch das Zusammenspiel aus Sensoren und Aktoren wiedergespiegelt wird. Hier kann die Virtuelle Inbetriebnahme (VIBN) eingesetzt werden (VDI/VDE 3693), um die Funktionsweise des Steuerungssystems vor der realen Inbetriebnahme und der Produktion zu testen und zu optimieren (Barth et al. 2015). Es existieren verschiedene Materialflussmodelle (vgl. Abb. 1), die je nach Anforderung im Kontext der Produktion eingesetzt werden können (Scheifele et al. 2016; Kienzlen und Verl 2020). Das logische Materialflussmodell stellt die Ablaufreihenfolge dar (Hoher 2017), beim ereignisbasierten Modell wird diese durch zeitliche Schritte ergänzt (Rybicka 2017). Bei beiden Modellen handelt es sich um ein- bis zweidimensionale Modelle, die nicht die exakten Trajektorien der einzelnen Stückgüter abbilden, sondern eine Abfolge der Verteilungen. Das kinematische Materialflussmodell (KMM) bewegt die Stückgüter anhand vordefinierter Trajektorien (Scheifele et al. 2016) und berücksichtigt keine Kollisionen. Das physikbasierte Modell (PMM) berechnet die höchste Detailgenauigkeit inklusive Kollisionen. Aus diesem Grund ist jedoch der größte Rechenaufwand notwendig, der mit der Anzahl der Stückgüter weiter zunimmt (Spitzweg 2009). Je mehr Kollisionen zwischen Stückgütern auftreten, desto aufwendiger wird die Berechnung. Darüber hinaus ist diese nicht deterministisch (Scheifele et al. 2016). Es ist somit nicht möglich, ein Szenario mit sehr vielen Stückgütern im Steuerungstakt zu berechnen. Hier setzt das makroskopische Flussmodell (MFM) an, bei dem der Materialfluss über die räumliche und zeitliche Verteilung der Stückgüter berechnet wird (Hoher et al. 2012). Der Berechnungsaufwand ist abhängig von der Größe des betrachteten Bereichs und der Diskretisierung, jedoch unabhängig von der Stückgüteranzahl. Das Modell stellt folglich eine Alternative bei der Betrachtung der räumlichen Bewegung von großen Mengen an Stückgütern dar. Vorige Arbeiten zeigen gute Ergebnisse beim Vergleich mit Realdaten (Hoher et al. 2012; Göttlich et al. 2014). Allerdings existieren bisher kaum Methoden zur Generierung von diskreten Signalen aus der kontinuierlichen Dichteverteilung des MFMs, ohne die eine Simulation nicht geeignet für die VIBN ist.

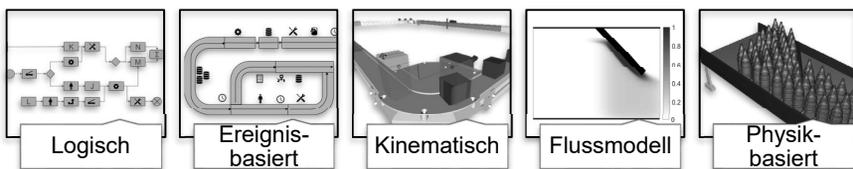


Abbildung 1: Materialflussmodelle nach (Scheifele et al. 2016; Hoher 2017; Kienzlen und Verl 2020)

2 Modellierung mit dem Flussmodell

Die Modellierung des MFM ist in (Göttlich et al. 2014) beschrieben. Es basiert auf den Annahmen, dass nur identische Stückgüter vorhanden sind und dass die Stückgüteranzahl konstant bleibt. Mathematisch wird die Dichte $\rho(x, t)$ mit Zeit t und Ort x über eine partielle Differentialgleichung beschrieben:

$$\partial_t \rho + \nabla \left(\rho (\mathbf{v}^{dyn}(\rho) + \mathbf{v}^{stat}(x)) \right) = 0, \text{ mit } \rho(x, 0) = \rho_0(x), \quad (1)$$

wobei $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2, t \geq 0$.

Die Umgebung wird als ein statisches Geschwindigkeitsfeld \mathbf{v}^{stat} berücksichtigt, beispielsweise mit positiven Geschwindigkeiten durch Bandförderer und negativen durch Hindernisse. Im Bereich der maximalen Dichte ρ_{max} , bei der die Stückgüter kollidieren, wirken Abstoßkräfte als dynamisches Geschwindigkeitsfeld

$$\mathbf{v}^{dyn}(\boldsymbol{\rho}) = H(\rho - \rho_{max}) \cdot I(\boldsymbol{\rho}) \text{ mit } H(\rho - \rho_{max}) = \begin{cases} 0, & \rho < \rho_{max} \\ 1, & \rho \geq \rho_{max} \end{cases} \quad (2)$$

Die Abstoßkraft $I(\boldsymbol{\rho})$ setzt erst ein, wenn die Heavisidefunktion $H(\rho)$ aktiv wird. Dies wiederum geschieht, wenn die Dichte in einem Bereich nahe der maximalen Dichte kommt abhängig von der stetigen, numerischen Realisierung der Heavisidefunktion. Die Abstoßkraft wird mittels einer Glättung beschrieben, wobei der Faltungskern $\boldsymbol{\eta}$ und der Faktor ϵ (Regulierung der Stärke der Abstoßkraft) verwendet werden:

$$I(\boldsymbol{\rho}) = -\epsilon \frac{\nabla(\boldsymbol{\eta} * \boldsymbol{\rho})}{\sqrt{1 + \|\nabla(\boldsymbol{\eta} * \boldsymbol{\rho})\|_2^2}}, \epsilon > 0. \quad (3)$$

In der Regel wird die Ausgangsdichteverteilung $\rho_0(x)$ aus den Positionen der Stückgüter erstellt wie sie auch für das KMM oder das PMM genutzt werden. Ein Stückgut wird dabei als bivariate Normalverteilung skaliert mit der maximalen Dichte in die Dichteverteilung überführt. Aufgrund der Nichtlinearität und Nichtlokalität in der Gleichung werden numerische Methoden zur Berechnung der Lösung benötigt. Es wird daher $\mathbf{x} = (x_1, x_2)$ in Zellen mit Breite $\Delta x_1, \Delta x_2$ und mit Mittelpunkt $x_{j,k}$ diskretisiert. In der hier verwendeten Notation wird davon ausgegangen, dass für x_1 und x_2 die gleiche Standardabweichung eingesetzt wird. Die Ausgangsdichteverteilung ergibt sich aus der Summe aller Normalverteilungen der Stückgüter mit der jeweiligen Position $x_{i,0}$ zum Zeitpunkt $t = 0$ (Göttlich und Pflirsching 2018):

$$\rho_0(x) = \sum_{i=1}^N \frac{1}{2\rho_{max}\pi\sigma_0^2} e^{-\frac{1}{2\sigma_0^2}(x-x_{i,0})^2}. \quad (4)$$

Die maximale Dichte beschreibt die maximal mögliche Anzahl an Stückgütern pro Flächeneinheit (Göttlich und Pflirsching 2018). Sie ist daher abhängig von der Form der Stückgüter. Das MFM kann bisher nur Stückgüter mit quadratischem (q) oder rundem (z) Querschnitt abbilden. Für einen runden Querschnitt mit Radius r berechnet sich die maximale Dichte nach (Degond et al. 2017)

$$\rho_{max,z} = \frac{1}{2\sqrt{3}r^2}, \quad (5)$$

da bei dichtester Packung von Zylindern ein Stückgut eine hexagonale Fläche mit Seitenlänge r besetzt (vgl. Abb. 2a). Bei quadratischem Querschnitt (Validierung in (Prims et al. 2019)) können die Stückgüter ohne Lücken positioniert werden. Die maximale Dichte ist demzufolge wie in Abbildung 2a skizziert bei der Seitenlänge a

$$\rho_{max,q} = \frac{1}{a^2}. \quad (6)$$

Neben den Positionen der einzelnen Stückgüter (als Erwartungswert in der Normalverteilung) beeinflussen folglich die Standardabweichung der verwendeten Normalverteilung und die Diskretisierung die Ausgangsdichteverteilung. Je größer die Standardabweichung, desto größer ist der Bereich, in dem ein Stückgut berücksichtigt wird. Voraussetzung ist, dass die Standardabweichung so gewählt wird, dass die Auswirkungen der Normalverteilung über ein Stückgut hinausgehen, sonst entsteht keine kontinuierliche Dichteverteilung, sondern punktuelle Spitzen je Stückgut (Göttlich und Pfrsching 2018). Überdies sollte die Standardabweichung größer als die Diskretisierung Δx sein. Konkret bedeutet dies für zylindrische Stückgüter

$$4\sigma_0 > 2r > \Delta x, \tag{7}$$

damit höchstens 95,4 % der Normalverteilung den Bereich des zylindrischen Stückguts umfassen. Bei unterschiedlicher Diskretisierung kann es sinnvoll sein, auch unterschiedliche Standardabweichungen zu wählen. Je feiner die Diskretisierung, desto genauere Werte können in den einzelnen Zellen abgelesen werden (vgl. Abb. 2b).

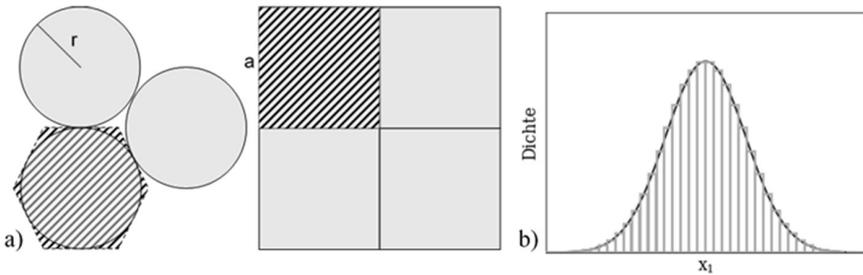


Abbildung 2: a) Skizze zur Bestimmung der maximalen Dichte b) Eindimensionale Normalverteilung eines Stückguts inklusive Diskretisierung (grau)

Die Dichteverteilung im zweidimensionalen entspricht einer Matrix, deren Werte die Wahrscheinlichkeit eines Stückguts in der entsprechenden Zelle beschreiben. Anschaulich lässt die Dichteverteilung sich als Heatmap abbilden (vgl. Abb. 3 rechts).

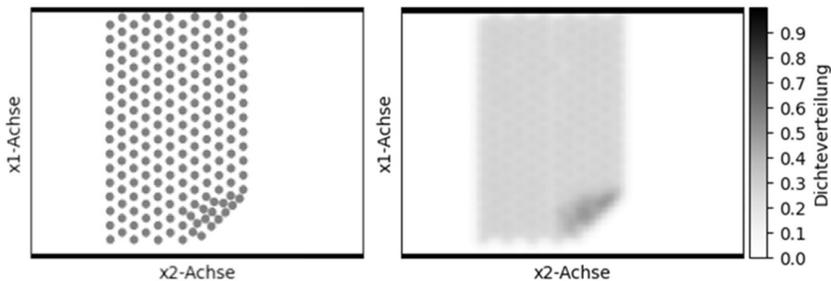


Abbildung 3: Positionen der Stückgüter und Ausgangsdichteverteilung als Heatmap

3 Signalgenerierung aus dem Flussmodell

Eine Steuerung sendet Aktorsignale an eine Maschine oder Anlage, um einen Prozess zu steuern. Die Steuerung erhält wiederum Informationen über den aktuellen Zustand des zu steuernden Prozesses über Sensorsignale, auf welche sie reagieren muss (vgl. Abb. 4). Das Zusammenspiel zwischen Aktor- und Sensorsignalen muss stimmen, damit die Steuerung ihre Funktion erfüllt. Bei der VIBN soll dies anhand einer Simulation getestet werden. Dazu wird in einer Hardware-in-the-Loop Simulation die Maschine oder Anlage durch eine Simulation ersetzt. Die Simulation muss daher Aktorsignale verarbeiten und Sensorsignale zur Verfügung stellen (Strigl 2009). Dabei müssen die Taktzeiten der Steuerung eingehalten werden. In der Materialflusstechnik werden verschiedene Sensoren zur Detektion eingesetzt (Li und Liu 2019), die auch in der Simulation abgebildet werden müssen (Spitzweg 2009).

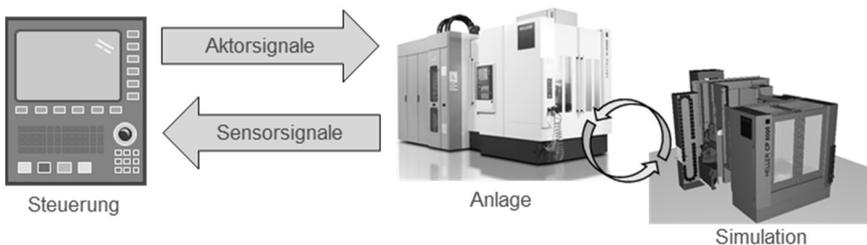


Abbildung 4: Signale zwischen Steuerung und Anlage bzw. Simulation

Aktorsignale für den Materialfluss sind insbesondere Geschwindigkeiten von Bandförderern oder ähnlichem und Geometrieänderungen wie beispielsweise die Winkeländerung eines Hindernisses. Diese Aktorsignale spiegeln sich im MFM im statischen Geschwindigkeitsfeld wieder. Die Änderungen der Geschwindigkeit in der Stärke des Feldes und die Geometrieänderungen in der Richtung des Feldes. Da die Schritte im MFM unabhängig von den vorherigen sind, kann prinzipiell in jedem Zustand die Ausgangssituation verändert und neu gestartet werden. Die Auswirkungen durch dynamische Änderungen wurden bisher jedoch nicht untersucht.

Für das MFM wurden bisher kaum Methoden zur Berechnung von Sensorsignalen beschrieben, daher sollen die Möglichkeiten zur Generierung aus dem MFM hier betrachtet werden. Die Funktionsweise der Sensoren ist dabei unerheblich für die Simulation. Es werden lediglich Informationen über die Positionen der Stückgüter betrachtet, andere Daten wie Temperatur oder Qualität werden im MFM nicht abgebildet. Da nicht einzelne Stückgüter erfasst, sondern kontinuierliche Dichteverteilungen beschrieben werden, können diese nicht analog zu KMM oder PMM berechnet werden. In den folgenden Unterkapiteln wird beschrieben wie verschiedene Sensorinformationen aus dem MFM generiert werden können. Der Fokus liegt dabei auf zylindrischen Stückgütern. Darüber hinaus wird die Darstellung in der Visualisierung für die VIBN adressiert, die ebenfalls aufgrund der kontinuierlichen Dichteverteilung erschwert wird. Die Berechnung der Signale wird anhand eines einfachen Beispiels an einem Bandförderer mit Verengung im folgenden Kapitel veranschaulicht.

3.1 Bereichsinformationen

Am einfachsten ist die Bestimmung von Bereichsinformationen, hier ist das MFM besonders vorteilhaft. Signale mit Bereichsinformationen können beispielsweise Zähl- oder sonstige Waagen, Stauwippen oder Kamerazählssysteme sein. Stauwippen detektieren das Aufstauen von Objekten; als Signal wird der Neigungswinkel der Wippe übergeben. Zählwaagen berechnen aus dem Gewicht direkt die Stückgutanzahl. Diese Signale lassen sich im MFM direkt aus der Dichtebeschreibung des Materialflusses ableiten.

Der reine *zeitliche Verlauf* hinter einer bestimmten Stelle wurde bereits in (Göttlich et al. 2014) als Qualitätskriterium berechnet. Dafür werden die Ausgangsdichteverteilung ρ_0 und die Anzahl der Stückgüter N als Ausgangspunkt verwendet und darauf basierend die Anteile hinter einem Hindernis im Bereich $\Omega_b \subseteq \mathbb{R}^2$ berechnet. Die Dichteverteilung, die sich in dem Bereich Ω_b befindet, wird aufsummiert und ins Verhältnis zu der Summe über die gesamte Dichteverteilung gesetzt:

$$N_{\Omega_b}(t) = \frac{\sum_{x_{k,l} \in \Omega_b} \rho(x_{k,l}, t)}{\sum_{x_{k,l} \in \mathbb{R}^2} \rho_0(x_{k,l})} \cdot N. \quad (8)$$

Mit dem Wissen über die Gesamtanzahl an Stückgütern kann daraus die Anzahl der Stückgüter berechnet werden, was dem Ergebnis der Zählwaage oder eines Kamerazählsystems entspricht. Ergänzt durch die Einzelmasse kann die Gesamtmasse bestimmt werden und daraus der Neigungswinkel der Stauwippe. Dies kann auf andere Bereiche übertragen und erweitert werden.

Eine Sonderform der Signale in einem Bereich ist die *Kollisionserkennung*. Hierfür gibt es in der Regel in der Realität keinen Sensor, in der Simulation kann die Information aber hilfreich sein. Gemeint ist damit eine Detektion, ob Stückgüter in einem bestimmten Bereich aufeinanderstoßen. Im MFM entspricht dies einem Bereich, in dem die Heaviside-Funktion die dynamische Geschwindigkeit aktiv schaltet. Der Bereich sollte dabei größer als die Fläche eines Stückguts sein.

3.2 Informationen zu einzelnen Stückgütern

An *Lichtschranken* werden einzelne Güter detektiert, deren Signale nicht direkt in einem kontinuierlichen Modell abgelesen werden können. Um trotzdem mit dem MFM Aussagen über einzelne Güter zu treffen, können stochastische Informationen genutzt werden. Die Dichteverteilung kann dazu als eine normalisierte Wahrscheinlichkeitsverteilung betrachtet werden (Göttlich und Pfirsching 2018). Für eine Lichtschranke werden die durchkreuzten Zellen der Dichteverteilung betrachtet, für einen Flächensensoren wiederum die Zellen der Dichteverteilung in diesem Bereich. Um aus der Wahrscheinlichkeitsverteilung Signale zu generieren, muss die räumliche Diskretisierung berücksichtigt werden. Die Wahrscheinlichkeit, dass sich ein Stückgut auf einer Linie befindet, setzt sich aus der Summe der Wahrscheinlichkeiten der einzelnen Zellen zusammen. Daher ergibt sich für die Wahrscheinlichkeit $P(x, t)$ auf der Linie L bei $x_2 = c$ zum Zeitpunkt t

$$P_L(t) = \sum_k \rho(x_{k,c}, t) \cdot \Delta x_1 \cdot \frac{1}{\sqrt{3r}}. \quad (9)$$

Falls die Wahrscheinlichkeit $P(t) \geq 1$ ist, bedeutet dies nicht zwangsläufig, dass sich ein Stückgut mittig auf der Linie befindet. Dafür kann es mehrere Ursachen geben:

1. Hierbei sind die Fläche der Stückgüter bzw. die projizierte Fläche berücksichtigt. Für eine Lichtschranke ist dies irrelevant, da diese auf den Umfang und nicht den Mittelpunkt eines Stückguts anschlägt
2. Zwei halbe Stückgüter wirken sich genauso auf die Wahrscheinlichkeit aus wie ein ganzes. Dies kann sich auf den bestimmten Wahrscheinlichkeitswert auswirken, wenn sich viele Stückgüter nahe der Lichtschranke befinden.
3. Die Anfangsdichteverteilung wird insgesamt geglättet. Daher ist zumindest zu Beginn der Simulation die Dichte eines Stückguts über ein größeres Gebiet verteilt.

Wird stattdessen ein *Flächensensor* im Bereich $\Omega_A(a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2) \in \mathbb{R}^2$ eingesetzt, ergibt sich analog zu (9) bezogen auf die Fläche A statt die Linie

$$P_A(t) = \sum_{k=a_1}^{b_1} \sum_{l=a_2}^{b_2} \rho(x_{k,l}, t) \cdot \Delta x_1 \cdot \Delta x_2. \quad (10)$$

Liegt der Wert der Dichteverteilung in diesen Zellen unter einem Schwellwert ϕ , so kann angenommen werden, dass kein Stückgut vorhanden ist. Dies gilt analog für eine Lichtschranke. Ein sinnvoller Wert für den Schwellwert könnte beispielsweise mit der doppelten Standardabweichung σ_0 aus der Normalverteilung in (4) bestimmt werden:

$$\phi = \frac{1}{2\rho_{max}\pi\sigma_0^2} e^{-2}. \quad (11)$$

Aus makroskopischer Sicht verhält sich eine *Abstandsmessung* ähnlich wie eine Lichtschranke. Es müssen wiederum die Zellen in einer Linie betrachtet werden. Allerdings muss eine Wahrscheinlichkeit definiert werden, bis zu der die Abstandsmessung betrachtet wird. Dann werden der Abstand bzw. die Zellenanzahl bis zu dieser Wahrscheinlichkeit betrachtet.

3.3 Informationen über spezifische Stückgüter

Im MFM werden alle Stückgüter zusammen in eine Dichteverteilung umgerechnet, ohne dass die einzelnen eine Markierung haben. Dadurch können keine Aussagen über spezifische Stückgüter getroffen werden. Die Erfassung spezifischer Güter wie über RFID-Sensoren, Barcode-Scanner oder kamerabasierte Codeleser kann mit Hilfe des MFMs nicht wiedergegeben werden. Werden diese Informationen benötigt, ist das MFM nicht geeignet. An dieser Stelle sollten eine Kopplung oder Überführung mit einem PMM angedacht werden. Dies soll in zukünftigen Arbeiten betrachtet werden.

3.4 Visualisierung

Die Visualisierung stellt keine Sensorsignale zur Verfügung, sondern dient zur direkten Interaktion des Benutzers mit der Simulation. Sie gibt einen ersten, anschaulichen Eindruck der aktuellen Situation für die VIBN. Dies ist für das MFM besonders wichtig, da die reine Wiedergabe als Matrix mit hunderten Einträgen unübersichtlich ist. Die einfachste Visualisierung der Dichteverteilung erfolgt mit Hilfe eines *Heatmap-Diagramms* (vgl. Abb. 3). Diese Visualisierung wurde bereits

von (Hoher et al. 2012) für die Dichteverteilung eingesetzt. Die Farbgebung kann dabei je nach Anwendungszweck variiert werden.

Für die VIBN wird in der Regel eine *dreidimensionale Visualisierung* basierend auf den CAD-Daten (Computer-Aided Design) eingesetzt. Darüber hinaus ist eine dreidimensionale Darstellung häufig intuitiv verständlicher als die zweidimensionale Visualisierung. Die einzelnen Zellen der Dichteverteilung können als Quader dargestellt werden (vgl. Abb. 5). Je höher der Quader, desto größer der Dichtewert. Die Skalierung sollte in Abhängigkeit der Umgebung gewählt werden. Je nach Umgebungsgröße und Diskretisierung kann für die Visualisierung der Durchschnitt mehrerer Zellen zu einem Quader zusammengefasst werden. Die Farben aus der zweidimensionalen Darstellung können weiterhin eingesetzt werden, evtl. ergänzt durch eine Variation der Transparenz. Der Transparenzwert α variiert in der RBGA-Darstellung zwischen Null (komplett durchsichtig) und eins (komplett undurchsichtig), er kann daher direkt aus der Dichteverteilung übernommen werden.

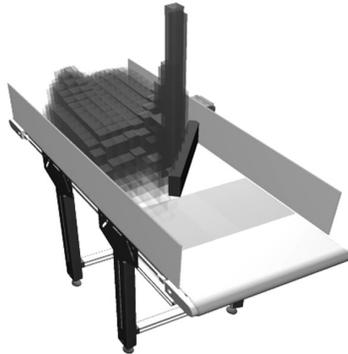


Abbildung 5: Dreidimensionale Visualisierung der Dichteverteilung

4 Veranschaulichung der Generierung von Sensorsignalen

Zur Veranschaulichung soll ein Förderband mit Verengung betrachtet werden, auf dem Flaschen bewegt werden. Die Visualisierung des zugehörigen makroskopischen Modells ist in Abbildung 5 zu sehen. Nach der Verengung befindet sich eine Zählwaage (Abb. 6 links, dunkel grob schraffiert), welche die Anzahl der Stückgüter misst. Am Übergang zu der Zählwaage befindet sich eine Lichtschranke (Abb. 6 links, helle Linie), die erkennt, wenn Stückgüter auf die Zählwaage transportiert werden. Zusätzlich gibt es einen Flächensensor (Abb. 6 links, hell fein schraffiert), der die Fläche vor der Verengung überwacht. Es befinden sich 192 Flaschen auf dem Förderband, deren Startpositionen in Abbildung 3 gezeigt sind. Die Signale werden in einem Bedienfeld gezeigt, das in Abbildung 6 rechts nach 1,4 Sekunden dargestellt ist. Die Lichtschranke wird mit Formel (9) berechnet, der Flächensensor mit Formel (10) skaliert mit der Stückgüteranzahl und der maximalen Dichte und die Zählwaage mit Formel (8). Für die Lichtschranke wurde der Schwellwert (11) mit dem Durchmesser der Flaschen als maximale Dichte verwendet.

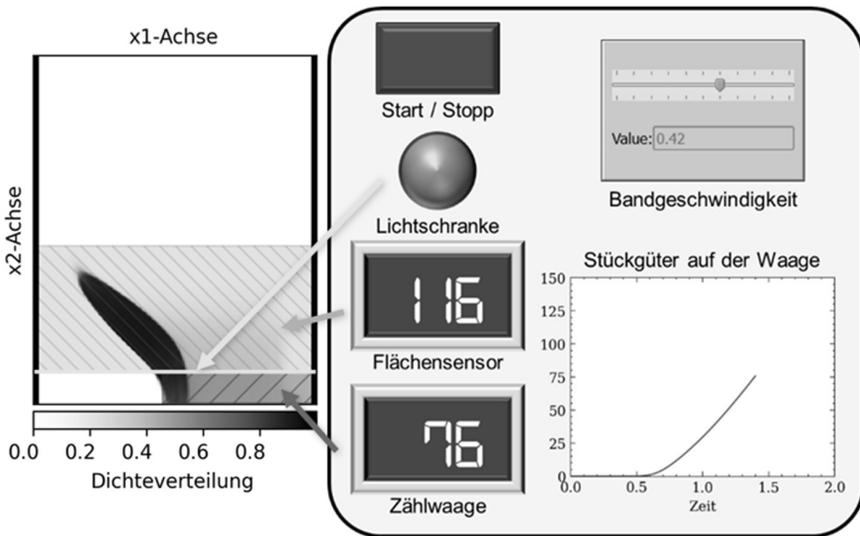


Abbildung 6: Dichteverteilung und schematisches Bedienfeld nach 1,4 Sekunden

5 Zusammenfassung und Ausblick

Der Materialfluss in einer Produktionsanlage kann mit unterschiedlichen Modellen dargestellt werden; bei großen Stückzahlen ist das MFM besonders geeignet. Um ein MFM in die VIBN zu integrieren, müssen Signale aus der kontinuierlichen Dichteverteilung generiert werden. Am einfachsten zu berechnen sind Bereichssignale wie der zeitliche Verlauf hinter einer Stelle, diese können auf die Ausgangsdichteverteilung bezogen werden. Für Lichtschranken und Flächensensoren können die stochastischen Wahrscheinlichkeitsinformationen in der Dichteverteilung genutzt werden. Aus der diskretisierten Darstellung wird die Wahrscheinlichkeit als Summe mit zusätzlichem Faktor bestimmt. Demzufolge handelt es sich nicht um absolute Ergebnisse, sondern um Ergebnisse mit Wahrscheinlichkeit. Die Visualisierung kann in Heatmap-Diagrammen oder dreidimensional mit Quadern unterschiedlicher Größe und Farbintensität erfolgen. Es können bisher keine dreidimensionalen Einflüsse berücksichtigt werden, sondern lediglich die Bewegung in einer Ebene. Dies soll in zukünftigen Arbeiten betrachtet werden. In Zukunft soll zusätzlich die Kopplung zwischen einem physikbasierten und einem MFM untersucht und realisiert werden, so dass spezifische Stückgüter ebenfalls betrachtet werden können. Außerdem sollen Änderungen im statischen Geschwindigkeitsfeld untersucht werden, so dass Aktorsignale einbezogen werden können.

Danksagung

Die Autoren bedanken sich für die Förderung durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 327964174. Vielen Dank außerdem an Frau Jennifer Weißen, M. Sc. vom Institut für Mathematik der Universität Mannheim für die Unterstützung und Zusammenarbeit.

Literatur

- Al-Hawari, T.; Aqlan, F.; Al-Buhaisi, M.'a.; Al-Faqeer, Z.: Simulation-Based Analysis and Productivity Improvement of a Fully Automatic Bottle-Filling Production System: A Practical Case Study. In: Second International Conference on Computer Modeling and Simulation, Sanya, China, 2010, S. 195–199.
- Barth, M.; Puntel Schmidt, P.; Hoernicke, M.; Oppelt, M.; Wolf, G.; Hundt, L.; Stern, O.: Methoden und Modelle der Virtuellen Inbetriebnahme - Eine Übersicht der Richtlinienarbeit des GMA FA 6.11. *Automation* 2015 (2015) 16.
- Degond, P.; Ferreira, M.A.; Motsch, S.: Damped Arrow–Hurwicz algorithm for sphere packing. *Journal of Computational Physics* 332 (2017), S. 47–65.
- Göttlich, S.; Hoher, S.; Schindler, P.; Schleper, V.; Verl, A.: Modeling, simulation and validation of material flow on conveyor belts. *Applied Mathematical Modelling* 38 (2014) 13, S. 3295–3313.
- Göttlich, S.; Pfirsching, M.: A micro-macro hybrid model with application for material and pedestrian flow. *Cogent Mathematics & Statistics* 5 (2018) 1, S. 1–20.
- Hoher, S.: Ein gekoppeltes Materialflussmodell zur durchgängigen Entwicklungsunterstützung von Materialflusssteuerungen: Stuttgarter Beiträge zur Produktionsforschung. Stuttgart, Universität Stuttgart, Dissertation, 2017.
- Hoher, S.; Schindler, P.; Göttlich, S.; Schleper, V.; Röck, S.: System Dynamic Models and Real-time Simulation of Complex Material Flow Systems. In: ElMaraghy, H.A. (Hrsg.): *Enabling Manufacturing Competitiveness and Economic Sustainability*. Berlin, Heidelberg: Springer 2012, S. 316–321.
- Kienzlen, A.; Verl, A.: Comparison of Material Flow Models and Acceleration of the Macroscopic Flow Model for Virtual Commissioning. In: Nketsa, A.; Baron, C.; Foucher, C. (Hrsg.): *ESM 2020, Toulouse, 2020*, 168–175.
- Li, P.; Liu, X.: Common Sensors in Industrial Robots: A Review. *Journal of Physics: Conference Series* 1267 (2019), S. 12036.
- Prims, D.; Kötz, J.; Göttlich, S.; Katterfeld, A.: Validation of Flow Models as New Simulation Approach for Parcel Handling in Bulk Mode. *Logistics Journal* (2019) 6.
- Rybicka, J.: *Modelling Flexible Manufacturing Systems through Discrete Event Simulation*. Cranfield, Großbritannien, Cranfield University, School of Aerospace, Transport and Manufacturing, PhD, 2017.
- Scheifele, C.; Lechler, A.; Verl, A.: Materialflussmodelle für die HiL-Simulation. *wt Werkstattstechnik online* 106 (2016) 3, S. 119–124.
- Spitzweg, M.: *Methode und Konzept für den Einsatz eines physikalischen Modells in der Entwicklung von Produktionsanlagen*. München: Utz 2009.
- Strigl, T.: Materialflusssimulation zur schnellen Inbetriebnahme realer Steuerungen. *ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb* 104 (2009) 1-2, S. 76–80.
- VDI Verein Deutscher Ingenieure VDI/VDE 3693: *Virtuelle Inbetriebnahme*. Berlin: Beuth, 2016.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Virtuelle Inbetriebnahme eines Leitsystems für die roboterbasierte automatische Kommissionierung in der Automobilindustrie

Virtual Commissioning of a Control System for automated robot-based Picking in automotive Environment

Xinyi Xie, Thomas Büttner, Marko Friedemann, Ken Wenzel, Fraunhofer Institut für Werkzeugmaschinen und Umformtechnik IWU, Chemnitz (Germany),
xinyi.xie@iwu.fraunhofer.de, thomas.buettner@iwu.fraunhofer.de,
marko.friedemann@iwu.fraunhofer.de, ken.wenzel@iwu.fraunhofer.de

Abstract: Robot-based commissioning in combination with automated guided vehicles (AGVs) for material provisioning is a solution for coping with changing requirements due to increasing product variety. To control such a system a software is required that optimizes the picking sequence and that coordinates the robot, the AGV and the safety equipment. Due to the complexity of the technical system, it is challenging to test the control software with real hardware. Therefore, this paper presents a virtual commissioning approach for such a software system called AutoLog. A simple programmed model is used for unit and regression tests and a more complex material flow model with stochastic behaviour for integration tests.

1 Motivation und Problemstellung

Um die Wettbewerbsfähigkeit der Produktion auch bei steigender Variantenvielfalt vor allem in Hochlohnländern zu sichern, müssen manuelle Tätigkeiten in der Fahrzeugmontage zunehmend automatisiert werden. Dies hat auch Auswirkungen auf damit verbundene Logistikprozesse. Besonders die Kommissionierung von Bauteilen entsprechend der Perlenkette oder deren Zusammenstellung in einem Warenkorb wurden bisher manuell durchgeführt. Durch Fortschritte in der Robotertechnik und den Einsatz von fahrerlosen Transportsystemen (FTS) können hier aber vor allem bei der Kommissionierung von großen Bauteilen Potentiale erschlossen werden. Dies gilt sowohl im Hinblick auf die Entlastung der Werker von körperlich schweren Arbeiten als auch im Hinblick auf die Reduzierung der notwendigen Fläche zum Aufbau der Anlagentechnik und zur Lagerung der Materialbehälter.

Dazu wurden vom Fraunhofer IWU zusammen mit einem Automobilhersteller entsprechende technische Lösungen wie Greifmittel, Sensortechnik und Roboterprogramme zum Aufbau einer automatischen Kommissionierungsanlage für

Fahrzeugteile entwickelt, die nach dem Ware-zur-Person-Prinzip vollautomatisch, unter Nutzung von fahrerlosen Transportsystemen, arbeitet.

Die Anlage ist für die Kommissionierung von Großbauteilen, wie Frontscheiben oder Getrieben, vorgesehen und hält jeweils eine gewisse Teilmenge in Behältern auf Pufferplätzen und in direkter Reichweite des Roboters vor. Bei großer Variantenzahl kann es erforderlich sein, einige dieser Plätze mit wechselnden Teiletypen zu belegen. Die dynamische Versorgung dieser Pufferplätze mit den entsprechenden Teiletypen durch ein FTS, der wahlfreie Zugriff auf Teile- und Sequenzbehälter sowie die Bestimmung einer Kommissionierungsreihenfolge stellen ein mehrdimensionales Optimierungsproblem dar. Deshalb wurde ein Softwaresystem entwickelt, das die Operationen von FTS zur Materialbereitstellung und Roboter zur Bauteilentnahme anhand der Aufträge vorausplant und deren Ausführung koordiniert.

Da durch die Nutzung eines FTS auf schwankende Transportzeiten und auch Ereignisse, wie notwendige Ladevorgänge reagiert werden muss, erweist sich ein umfassender Test des Softwaresystems als schwierig. Auch wenn einzelne Sonderfälle durch programmatische Tests abbildbar sind, können komplexe Testszenarien nur durch ein Simulationsmodell mit entsprechend stochastischem Verhalten durchgeführt werden. Mit Inbetriebnahme einer ersten Version des Softwaresystems hat sich gezeigt, dass im Realbetrieb sehr viele unerwartete, zufällige Ereignisse auftreten, die durch die Steuerungsalgorithmen in der Software nicht berücksichtigt wurden. Dies führte zu Fehlern und letztendlich zu einer langen Inbetriebnahme. Unterstützen kann hierbei eine virtuelle Inbetriebnahme (VIBN). Sie „wird zum Aufdecken und Beheben von Fehlern aus dem Engineering des Automatisierungssystems durchgeführt [...] und bedeutet das entwicklungsbegleitende Testen einzelner Komponenten und Teilfunktionen des Automatisierungssystems mithilfe von auf die jeweilige Aufgabenstellung abgestimmten Simulationsmethoden und -modellen“ (Verein Deutscher Ingenieure e.V. 2016). Um die Anlaufphase zu verkürzen, wurde die Methode der VIBN durch Kopplung des Leitsystems mit einer Materialflusssimulation gewählt.

2 Einordnung in das Themenfeld und wissenschaftlicher Beitrag

Das Anwendungsumfeld in diesem Beispiel ist gekennzeichnet durch einen Entwicklungsprozess in Anlehnung an die Richtlinie VDI 2221 - Methodik zum Entwickeln und Konstruieren technischer Systeme und Produkte (Verein Deutscher Ingenieure e.V. 2019). Die Hauptschritte können gegliedert werden in Anforderungsanalyse, Konzeptentwicklung, Machbarkeit, Umsetzung der priorisierten Lösung und iterative Anpassungen, Implementierung in ein bestehendes System mit Anlaufphase und Betrieb.

Sowohl die Konzeptentwicklung als auch anschließende Machbarkeitsuntersuchungen werden heutzutage standardmäßig mit digitalen Werkzeugen aus Produktentwicklung und Fabrikplanung unterstützt. Beide letztgenannte Betrachtungsbereiche sind dabei unabdingbar miteinander verknüpft (Westkämper 2008; Scheer 2015).

Mit Entwicklung spezieller Fertigungssysteme im Automotive-Bereich kommen in der Praxis ereignisdiskrete Simulationsmodelle, wie bspw. Siemens Plant Simulation oder Process Simulate, zum Einsatz. Diese dienen unter anderem der Absicherung des Materialfluss- und Anlagenkonzepts. Die Ergebnisse der Simulationsstudien werden anschließend in technische Ausschreibungsunterlagen überführt, in diesem Fall jeweils für die Softwareentwicklung und den Anlagenbau. Durch die Abkapselung der Simulationsmodelle in den Konzeptphasen und abhängig vom Komplexitätsgrad des technischen Systems sind die Validierungsergebnisse für das reale System schwierig annehmbar. Vielmehr ist es notwendig, die Modelle als Emulation der realen Welt, für die technische Umsetzung mitzuführen, was wiederum selten oder gar nicht praktiziert wird. Mit der Befähigung bzw. Erweiterung, die Modelle über Schnittstellen miteinander kommunizieren zu lassen, können ebendiese für eine VIBN genutzt werden.

Gerade in komplexen Systemwelten, wie in diesem Anwendungsfall bestehend aus Roboter, FTS, Anlagen-SPS, kundenseitige Auftrags- und Lagerverwaltungssysteme sowie der Interaktion untereinander, fehlt aktuell eine mittels VIBN umfassend qualifizierte Leitsteuerung als Software-in-the-Loop.

Im Themengebiet der virtuellen Inbetriebnahme existieren bereits viele Arbeiten. Hauptsächlich werden maschinennahe, Hardware-in-the-Loop-Anwendungen thematisiert, die auf die Entwicklung und Qualifizierung von Verhaltensmodellen von Maschinen und Anlagen ausgerichtet sind. Software-in-the-Loop sowie Model-in-the-Loop-Anwendungen bilden dabei den kleineren Anteil.

Grzona et al. (2019) beschreiben ein praxisnahes Vorgehen innerhalb der Fabrikplanung, bei dem physische Komponenten einer realen Versuchsanlage emuliert werden. Die Simulationsumgebung übernimmt Funktionalitäten in Form von Prozessbausteinen und steuert das Realsystem ereignisbasiert über eine reale Steuerung. Für die Kommunikation untereinander wurden innerhalb der Simulation Bausteine für einen Datenaustausch mit einem OPC UA-Server erstellt.

Süß et al. (2015) konzentrieren sich auf das Verhaltensmodell von Komponenten beziehungsweise dessen Erstellung, durch Austausch mit bestehenden Modellinformationen aus dem Entwicklungsprozess der Komponentenhersteller. Damit sollen Abschätzungen hinsichtlich der Integrierbarkeit neuer Produkte im Brown-Field einer Fabrik verbessert werden, um Entscheidungen in der Beschaffung zu unterstützen. Eine analoge Entwicklungsrichtung wurde im Projekt ENTOC (Festo SE & Co. KG) am Demonstrator einer Hochzeitsstation in der Automobilproduktion verfolgt.

Makris et al. (2012) beschreiben einen Workflow zur virtuellen Inbetriebnahme einer Roboterzelle mit zwei arbeitsteiligen Robotern. Der Fokus liegt hier auf der Abstimmung von Roboterbewegungen für einen gemeinsam benutzten Bereich. Dementsprechend sind detaillierte Eingangsdaten, wie beispielsweise Werkstückgeometrie, Kinematik, Layout der Roboterzelle und technologischer Prozessfluss, notwendig.

Wallner (2021) schlägt generell ein modellbasiertes Entwicklungsvorgehen vor und beschreibt, eher kurz, wie ein etablierter Workflow aus der Automobilbranche auf die Maschinenbaubranche übertragen werden kann. Dies stellt er exemplarisch mit den Werkzeugen Simulink und Simscape vor. Weiterhin gibt er einen kurzen Überblick

über die Architektur des Hardware-in-the-Loop-Prinzips und wie der Austausch der echtzeitfähigen Maschinen-Codes mit der SPS verbunden werden.

Bernard (2019) berichtet, analog wie zuvor, vom prinzipiellen Vorgehen bei einer virtuellen Inbetriebnahme aus einem Forschungsprojekt zwischen Volkswagen und Siemens. Im Beispiel der Heckklappenmontage existierten bereits mehrere ablauffähige Simulationsmodelle einzelner Anlagenkomponenten in jeweils unterschiedlicher Simulationsumgebung. Herausforderung war es, diese Modelle zusammenzuführen, wofür ein Framework entwickelt wurde, um ein Gesamtmodell, inklusive Steuerungssoftware, zu erhalten. Die Zusammenführung verschiedener Modelle zu einer Co-Simulation entspricht laut Einschätzung der Autoren den heutigen Anforderungen, ist allerdings nicht Fokus dieses Beitrags.

Kaffka (2019) beschreibt die Entwicklung eines Emulations-Controllers, welcher als konfigurierbare Schnittstelle zwischen Simulationsmodell eines Hochregallagers und einem Warehouse-Management-System als Produktivsystem dient. Mehrere emulierte Teilmodelle kommunizieren über diese Schnittstelle miteinander und tauschen Befehle und Status aus. Das Szenario ähnelt dem im hier vorliegenden Beitrag beschriebenen Szenario. Im Unterschied zu Kaffka liegt der Fokus auf einer Befähigung eines eigenständig ablauffähigen Simulationsmodells hin zu dessen externer Steuerung durch ein übergeordnetes Leitsystem, AutoLog. Des Weiteren sind auch Auftragsverwaltung und Behälterverwaltung integraler Bestandteil der Simulation, so dass jegliche, auch physische, Systemelemente in ihrem Ablaufautomatismus voneinander zu entkoppeln sind. Letztendlich dient die Simulation im vorliegenden Beitrag dazu, das Leitsystem als Software-in-the-Loop zu qualifizieren.

Weitere Praxisberichte sind in Best Practices zur Virtuellen Inbetriebnahme (Riedel et al. 2020) zu finden. Hier werden prinzipielle Vorgehensweisen für unterschiedliche Einsatzzwecke in Kürze dargestellt. Unter anderem werden bausteinorientierte Simulationsplattformen mit Vereinigung von Software-in-the-Loop, Hardware-in-the-Loop und Model-in-the-Loop für durchgängige virtuelle Inbetriebnahmen vorgeschlagen. Der hier vorliegende Bericht konzentriert sich hingegen auf die spezifische Entkopplung von Prozessen innerhalb der Materialflusssimulation, welche über Socket-Schnittstellen von einem externen Leitsystem angestoßen werden und gleichzeitig Rückkopplungen aus dieser emulierten Umgebung zurückgespielt werden.

3 Aufbau der Systemarchitektur

3.1 Übersicht der Systemkomponenten

Die Systemkomponenten für die virtuelle Inbetriebnahme lassen sich in die drei Hauptkomponenten: das Simulationsmodell, den Adapter als Vermittler zwischen Modell und Leitsystem sowie das Leitsystem selbst, einteilen.

Das zur Simulation der realen Umgebung verwendete Modell wurde vom Auftraggeber als eigenständiges Modell zur Verfügung gestellt. Es besteht aus mehreren Komponenten: dem fahrerlosen Transportfahrzeug (FTF), inklusive dessen Auftragssteuerung, dem Roboter mit Steuerung und dem Auftragswesen für die

Erstellung der Kommissionier-Sequenz. Das FTF versorgt den Roboter bedarfsgerecht mit den angeforderten Teilen und tauscht die Behälter dynamisch aus. Für die Verwendung des Modells zur virtuellen Inbetriebnahme des Leitsystems mussten Anpassungen vorgenommen werden (siehe Kapitel 4). Über eine Schnittstelle werden einfache Textnachrichten für Aufträge und Rückmeldungen in Richtung Leitsystem gesendet. Ein Adapter wandelt die Nachrichten in das Format der Schnittstellen der Produktivsysteme um, mit denen im Realbetrieb interagiert wird. Das Leitsystem sendet Aufträge an den Roboter und das FTS zurück, die im Adapter wieder in Textnachrichten umgewandelt werden und im Simulationsmodell entsprechende Methoden zur Steuerung der Komponenten auslösen.

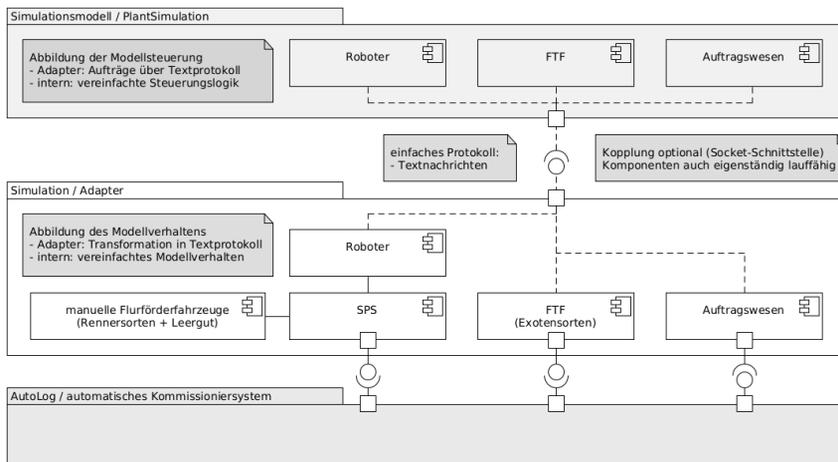


Abbildung 1: Übersicht der Systemarchitektur der VIBN

3.2 Prozesse und Funktionalität des Leitsystems

Zentrale Aufgabe des Leitsystems ist die Steuerung des automatischen Kommissionierungsprozesses entsprechend der eingehenden Aufträge. Neben dem Roboter für die eigentliche Kommissionierung sind daran mehrere FTF für die Zuführung von Exotensorten sowie den Transport fertig beladener Sequenzwagen zur Montagelinie beteiligt. Während für letztere nur ein Abfahrtssignal gesendet werden muss und sie ansonsten unabhängig vom Leitsystem arbeiten, ist die Erzeugung und Überwachung von Transportaufträgen zur zeitlich optimierten Bereitstellung von Exotensorten eine wichtige Teilaufgabe des Leitsystems. Unter Berücksichtigung von Behälterfüllständen, Anstellungssituation und auch der Auftragsreihenfolge muss das System Transportaufträge einplanen, an das FTF übergeben und ihre Ausführung überwachen. Behältertransport und -tausch erfolgen dabei parallel zu Roboterbewegungen (dynamische Anstellung). Die dafür notwendige Synchronisation wurde in Form einer virtuellen Schranke am Behälteranstellplatz umgesetzt, nachfolgend Behältersperre genannt, die einen Behälterwechsel so lange blockiert, wie der Roboter mit der Teileentnahme beschäftigt ist. Die Synchronisation auf Seiten des Roboters wird gewährleistet durch den Erhalt der Steuerbefehle erst bei Verfügbarkeit des Materials.

Zur Minimierung von Stillstandszeiten des Roboters wird ein Genetischer Algorithmus verwendet, der die vom FTS durchzuführenden Tauschoperationen sowie Leergutfahrten für Behälter mit Exotensorten optimiert. Dabei kann gegebenenfalls auch die Reihenfolge von Greifoperationen innerhalb der Aufträge angepasst werden, soweit die Anlage bzw. der Sequenzwagen eine freie Ablagereihenfolge unterstützt. Für die Einhaltung der Gesamttaktzeit der Pilotanlage ist das erforderlich, da die Fahrzeiten des FTS die Prozesszeiten des Roboters um ein Mehrfaches überwiegen.

Der entstehende Plan ist die Basis für die Roboter- und FTF-Auftragssteuerung sowie die Synchronisation mit der Teileentnahme und der Behältersperre. Er dient auch zur Ankündigung manuell durchzuführender Ver- und Entsorgungsprozesse außerhalb der Kommissionierungszelle.

3.3 Anforderungen für die virtuelle Inbetriebnahme

Durch die Kommunikation mit den unterschiedlichen Systemen innerhalb der Anlage (SPS, Roboter, Steuerungen der FTF) und in Verbindung mit den äußeren Ver- und Entsorgungsprozessen (Flurförderfahrzeuge, Routenzüge etc.) ist das Leitsystem einer Vielzahl von dynamischen Einflüssen und Fehlerquellen ausgesetzt. Diese können mit automatisierten Tests der einzelnen Teilkomponenten nicht ausreichend abgedeckt bzw. nachgebildet werden.

Ziel war es daher, mit einer virtuellen Nachbildung der Anlage, inklusive der äußeren Prozesse, das Leitsystem virtuell in Betrieb zu nehmen. Das heißt, die Kommissionierung einer großen Teilmenge zu simulieren sowie den Optimierungsalgorithmus bei vorgegebener Taktzeit und die FTF-Steuerung zu validieren. Um das System selbst dabei möglichst unverändert zu lassen, sollten alle virtuellen Komponenten die Schnittstellen der realen Anlage nachbilden. Die Simulation sollte also die Übermittlung von Auftragsdaten an das übergeordnete Leitsystem, das Verhalten von Roboter und Sensorik (SPS), das Verhalten der FTF sowie die äußeren Prozesse (Behältertausch) beinhalten.

3.4 Aufbau und Ablauflogik des Simulations-Modells

Die Programmsteuerung des ursprünglichen Modells besteht aus mehreren Teilen: der Sequenzgenerierung, der Auftragsabwicklung, der Behälterverwaltung, der FTF-Auftragssteuerung, der Robotersteuerung und der FTF-Routenverwaltung.

Bei der Sequenzgenerierung generiert das Programm eine zufällige Abfolge von Kundenaufträgen entsprechend einer vorgegebenen Häufigkeitsverteilung. Es werden sowohl Renner-Bauteile als auch Exoten-Bauteile unterschieden. Exoten-Bauteile werden vom FTF am Kommissionierungs-Roboter bereitgestellt. Der Versorgungsprozess für die Renner-Bauteile steht nicht im Fokus. Hier wird eine unendliche Quelle angenommen. Die Auftragsabwicklung prüft dieses Kriterium in Verbindung mit der Materialverfügbarkeit an den Lagerplätzen und erstellt einen FTS-Auftrag mit FTF-Auftragsnummer, Lager- und Zielort. Die FTF-Auftragssteuerung ordnet den FTF die entsprechenden Aufträge zu und plant die Auftragsabwicklung. Die FTF-Routenverwaltung berechnet die FTF-Fahrstrecke zwischen aktuellem Entnahme- und Abstellplatz und plant das Fahrzeug entsprechend zeitlich ein. Die Behälterverwaltung steuert den Nachschub voller Behälter mit

Exoten-Bauteilen im Austausch mit leeren Behältern ein. Die Robotersteuerung bestimmt, wann und wo der Roboter das angeforderte Bauteil aufnimmt und in Sequenz auf die Ablage legt. Hat das FTF den Bereitstellplatz erreicht, kann der Roboter mit der Bauteilentnahme aus dem Behälter beginnen. Der Übergabeprozess zwischen FTF und Roboter ist für den Zweck der Anwendung unbedeutend. Der Roboter selbst wird lediglich mit einer Handhabungszeit von circa 30 Sekunden zwischen Entnahme und Ablage modelliert. Nach Entnahme des Bauteils wird der Behälter für weitere Transporte durch das FTF freigegeben. Sind aktuell keine Transportaufträge vorhanden, kehrt es zur nahegelegenen Warteposition zurück. Kommissionierungsaufträge werden aufgrund der Perlenkette strikt nach dem FIFO-Prinzip eingesteuert. Abbildung 2 zeigt die Beziehung zwischen den einzelnen Steuerungen.

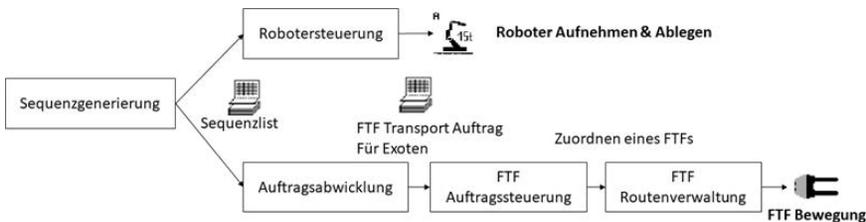


Abbildung 2: Relationen-Modell der Simulationskomponenten

4 Methodik

4.1 Kommunikation zwischen Leitsystem und Simulationsmodell

Bereits während der Entwicklung des Softwaresystems wurden automatisierte Tests erstellt, um sicherzustellen, dass Änderungen oder neue Funktionen keine Regressionen verursachen. Diese Tests wurden zunächst getrennt für die einzelnen Bestandteile der Software ausgeführt. Die angesprochene Komplexität im Zusammenspiel zwischen eingehenden Aufträgen, der Planung entsprechender Roboter- und FTS-Operationen sowie deren Abarbeitung durch die einzelnen, unabhängigen Systeme machte es jedoch erforderlich, vollständige Testläufe mit allen Komponenten durchführen zu können. Dazu wurden die in der frühen Entwicklungsphase, für die initialen Testsznarien entwickelten Bausteine zu einer vorerst deterministisch simulierten Umgebung kombiniert. Mit ihnen wurden das Verhalten von Roboter und FTS vereinfacht nachgebildet und die realen Schnittstellen von Auftragswesen, programmierbarer Steuerung (PLC) und FTS implementiert. Auf diese Weise war es möglich, frühzeitig vollautomatische Testläufe ausführen zu können.

Im Sinne der Qualitätssicherung wurde die Software anschließend gegen ein stochastisches, detaillierteres Simulationsmodell in Siemens Plant Simulation getestet. Es erzeugt ERP-, SPS- und Sensordaten und übermittelt diese an das Leitsystem. Dieses wiederum übermittelt die Steuerungsbefehle an die Systemelemente im Modell und steuert dort Roboter und FTF (siehe Abbildung 3).

Für die Kopplung beider Systeme wurde die vereinfachte Simulation zu einem Adapter mit Socket-Schnittstelle erweitert, der die im Realsystem verwendeten Schnittstellen für die Anlagensteuerung und das FTS sowie in Richtung Plant Simulation ein einfaches Textprotokoll anbietet. Dadurch ist es möglich, das Softwaresystem ohne Anpassungen, sowohl zusammen mit dem Simulationsmodell als auch in der realen Anwendung, auszuführen.

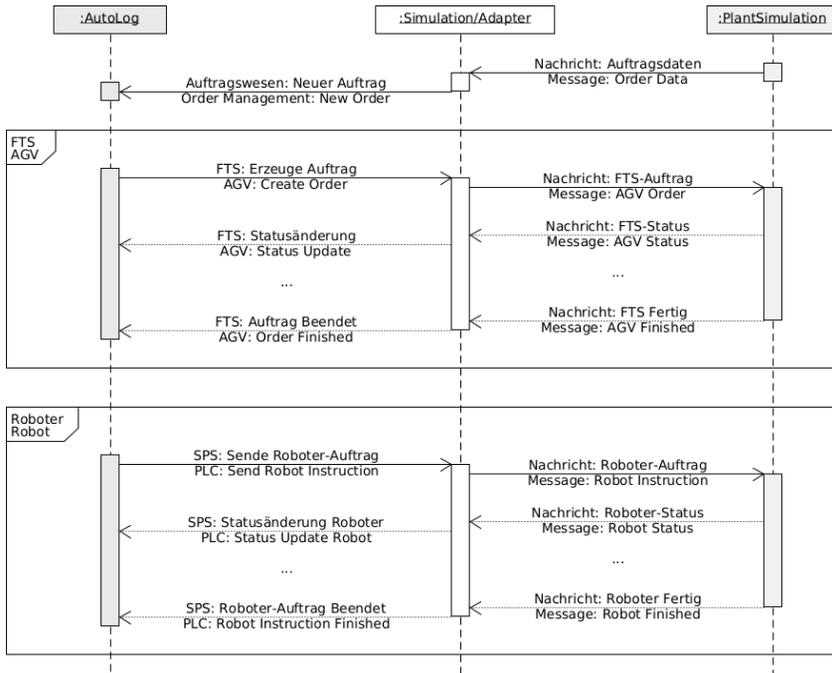


Abbildung 3: Sequenzdiagramm der VIBN (Ausschnitt)

4.2 Befähigung des Logistikmodells als emuliertes Realsystem

Um die Kontrolle des FTS und des Roboters im Modell durch eine externe Steuerung zu ermöglichen, müssen die Steuerungsmethoden der entsprechenden Funktionen im initialen Simulationsmodell angepasst werden.

Aus der Abbildung im obigen Abschnitt ist ersichtlich, dass das externe Leitsystem AutoLog eine Sequenzliste aus dem Modell erhält, auf deren Basis es Aufträge für die Beladung der Sequenzwagen generiert. Das Leitsystem bestimmt die optimale Teilereihenfolge unter Berücksichtigung notwendiger Transporte von Exoten-Behältern, um die Effizienz des FTS und dessen Transportleistung zu erhöhen. Dies wird unter anderem durch eine gezielte und vorausschauende Verwendung der Sequenzliste gewährleistet. Ein Beispiel ist ein Verbleib eines bestimmten Behälters am Roboter, falls Teile daraus demnächst wieder angefordert werden. So greift AutoLog gezielt in die Beauftragung des FTS im Simulationsmodell ein und eine strikte FIFO-Reihenfolge kann zum Teil aufgebrochen werden. Gleichzeitig erfasst das Leitsystem zu jeder Zeit den Behälterfüllstand und erteilt bei Bedarf rechtzeitig

Entsorgungsaufträge. Dies erfordert, dass die ursprüngliche FTF-Auftrags erzeugungsmethode des Modells vollständig inaktiv ist und diese Funktion dem Leitsystem überlassen wird. Darüber hinaus bestimmt es die Reihenfolge beziehungsweise Entnahmeposition der Teile im Behälter, was eine Steuerung der Entnahme des Roboters nach sich zieht. Um eine flexiblere Steuerung des Roboters zu ermöglichen, muss das ursprüngliche Modell in den Steuerungsmodus "Kommissionieren anhand der vom Leitsystem vorgegebenen Position des Bauteils" geändert werden. Gegen das vorzeitige Tauschen von Behältern am Entnahmeplatz und als Synchronisierung mit den Roboterbewegungen, wurde im Modell die erwähnte Behälterschranke umgesetzt. Nach Quittierung der Teileentnahme und Freigabe durch das Leitsystem wird der „pausierte“ Transportauftrag im Simulationsmodell fortgesetzt. So kann das FTF Transportaufträge eher starten, anstatt auf das Prozessende des Roboters zu warten. Der neue Programmablaufplan des Simulationsmodells sieht wie folgt aus:

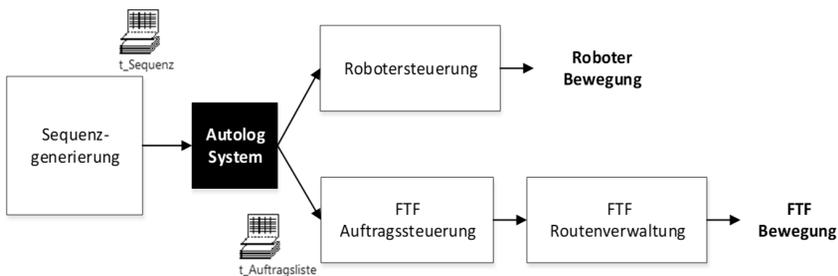


Abbildung 4: Sequenzdiagramm Simulationsmodell nach verkoppelt mit AutoLog

Der Datenaustausch zu AutoLog erfolgt durch einfache Textnachrichten über eine Socket-Schnittstelle. Über eine Schlüssel-tabelle werden die unterschiedlichen Objektnamen von Simulationsmodell und Leitsystem gepaart sowie das Format und die Interpretation der von beiden Seiten zu übertragenden Nachrichten festgelegt. Entsprechend der in der Schnittstelle definierten Steuerungsmethode wird der übergebene Informationssatz in einzelne Anweisungen für FTF, Roboter oder SPS der Kommissionierungszelle zerlegt. Prozessereignisse, wie Start, Fortschritt und Ende, werden zurück an das Leitsystem gemeldet. Wie in Abbildung 5 links dargestellt, zeigt die Konsole des Modells die vom Adapter eingegangenen Befehlsnachrichten an. Die Rückruf-Methode des Sockets teilt die Nachricht auf und weist sie auf verschiedene Objekte (FTF, Roboter, Schranke) zu. Analog werden über die Socket-Schnittstelle auch Antworten versendet. Der rechte Teil der Abbildung zeigt die Rückmeldung, die der Adapter vom Modell erhält.

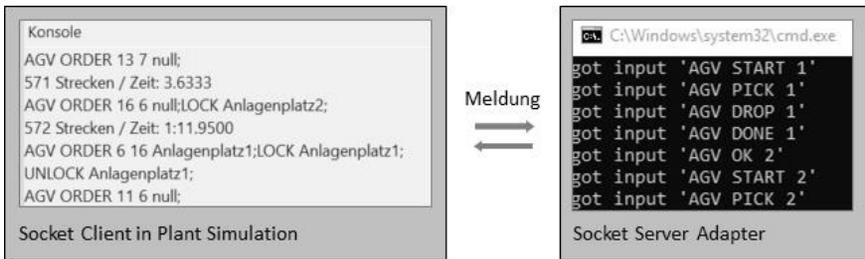


Abbildung 5: Beispiele für Kommunikationsnachrichten zwischen Modell und Adapter

5 Ergebnisse

Die beschriebene Anlagensimulation hat umfangreiche Tests des Systems AutoLog mit mehreren zehntausend Teilen ermöglicht. Dabei wurde sowohl auf das vereinfachte (zeitbasierte) Systemverhalten in der Komponente Simulation/Adapter als auch auf die umfangreichere Variante in Plant-Simulation-Modell zurückgegriffen. Durch das dynamische Zusammenspiel der Komponenten konnte eine Vielzahl von Unzulänglichkeiten und Fehlern im Planungs- und Optimierungsalgorithmus, in den Ausführungseinheiten für Roboter- und FTS-Operationen sowie deren Synchronisation (Behältersperre) identifiziert und behoben werden. Hauptfehlerklassen waren dabei Stillstände durch fehlerhafte Abhängigkeiten zwischen Operationen (Planung) und Wettlaufsituationen bei deren Ausführung und Synchronisation. Insgesamt sind aus der Analyse von fehlgeschlagenen Simulationsläufen 15 neue Regressionstests entstanden, die die jeweilige Fehlersituation nachbilden und ihre korrekte Handhabung sicherstellen.

Als Rückkopplung zum Simulationsmodell und als Anpassung dessen wurde die, bereits erwähnte, Behältersperre am Bereitstellplatz integriert. Damit konnten Bewegungen des Roboters und des FTF im Modell teilweise parallelisiert und zeitlich optimiert werden. Die Wartezeit des FTF am Behälteranstellplatz verringerte sich um circa zwei Drittel. Die Leerfahrten zur erwähnten Warteposition fallen abhängig vom Zielort und der Auftragsreihenfolge zeitlich unterschiedlich aus. Hier wurden keine genauen quantitativen Zahlen erfasst. Dennoch wurde eine merkbliche Verringerung der Leerfahrten beobachtet. Eine Quittierung von Entnahme- und Transportvorgängen zwischen den Komponenten Roboter, Behälter und FTF ist damit anforderungsgerecht umgesetzt worden.

6 Zusammenfassung und Ausblick

Durch Modularisierung und Flexibilisierung von Kommissionierungssystemen skaliert deren technische Entwicklung, da sie Standort unabhängig und in jeweils benötigter Modulanzahl eingesetzt werden kann. Damit skaliert auch das entwickelte Leitsystem. Durch eine entwicklungsbegleitende VIBN, welche unterschiedliche Kombinationen solcher Kommissionierungsmodule berücksichtigen kann und deren Einbindung in die informationstechnische Systemwelt testet, kann die Anlaufphase bis zum Serienbetrieb kontrolliert klein gehalten werden. Zudem sind Testszenarien virtuell schneller zu evaluieren als am realen System.



Das Simulationsmodell mit seiner aktuellen Anbindung an das Leitsystem ist geeignet für zukünftige Erweiterungen. So sollen demnächst, ausgehend vom einzelnen Kommissionierungssystem, größere Produktionsbereiche oder auch ganze Produktionssysteme emuliert werden und mit den entwickelten Schnittstellen einfach an AutoLog anzubinden sein.

Literatur

- Bernard, A.: Vollgas in der Fertigungsstraße. Ein aktuelles Forschungsprojekt von Volkswagen und Siemens. Siemens AG 2019. Online verfügbar unter <https://new.siemens.com/de/de/unternehmen/stories/forschung-technologien/digitaler-zwilling/digital-twin-cars-production-line.html>, zuletzt geprüft am 17.05.2021.
- Festo SE & Co. KG: Forschungsprojekt ENTOC. Verhaltensmodelle von Komponenten für virtuelle Inbetriebnahme. Online verfügbar unter <https://www.festo.com/group/de/cms/12827.htm>, zuletzt geprüft am 18.05.2021.
- Grzone, P.; Knüpfer, P.; Wilsky P.: Anwendungsorientierte Simulationsbausteine für die schlanke virtuelle Inbetriebnahme verketteter Produktionssysteme. In: Matthias Putz und Andreas Schlegel (Hg.): Simulation in Produktion und Logistik 2019. Auerbach: Verlag Wissenschaftliche Scripten.
- Kaffka, J.: Virtuelle Inbetriebnahme von Logistiksystemen. Der Einsatz des Emulation Controllers und von Demo3D als Erfolgsfaktor. SimPlan Integrations GmbH. 2019. Online verfügbar unter https://www.simplan.de/wp-content/uploads/Pr%C3%A4sentation_ASIM_2019_Jan_Kaffka_Virtuelle-Inbetriebnahme-von-Logistiksystemen-Der-Einsatz-des-Emulation-Controllers-und-von-Demo3D-als-Erfolgsfaktor_16zu9.pdf, zuletzt geprüft am 18.05.2021.
- Makris, S.; Michalos, G.; Chryssolouris, G.: Virtual Commissioning of an Assembly Cell with Cooperating Robots. In: *Advances in Decision Sciences* 2012, S. 1–11.
- Riedel, O.; Verl, A.; Kübler, K.; Jaensch, F.; Deuse, J.; Nöhling, F.; Borggräfe, T.; Richter, R.: Best Practices zur Virtuellen Inbetriebnahme. Hg. v. Verband Deutscher Maschinen- und Anlagenbau e.V. VDMA. Frankfurt am Main, 2020. Online verfügbar unter https://industrie40.vdma.org/documents/4214230/57246917/VDMA_Leitfaden_VIBN_1607503338494.pdf/c6b632a6-ebc2-113e-5753-073e4b28eb34, zuletzt geprüft am 19.05.2021.
- Scheer, A.-W. : Industrie 4.0: Von der Vision zur Implementierung. Whitepaper. 2015. Online verfügbar unter https://www.scheer-group.com/Scheer/uploads/2016/03/Scheer-Whitepaper_Industrie-4-0_DE.pdf, zuletzt geprüft am 14.04.2021.
- Süß, S.; Strahilov, A.; Diedrich, C.: Behaviour Simulation for Virtual Commissioning using Co-Simulation. In: 20th IEEE Conference on Emerging Technologies and Factory Automation (ETFA). September 8 - 11, 2015, Luxembourg. Institute of Electrical and Electronics Engineers; IEEE International Conference on Emerging Technologies and Factory Automation; ETFA. Piscataway, NJ: IEEE.
- Verein Deutscher Ingenieure: VDI-Richtlinie 2221 Blatt 1, 2019-11: Entwicklung technischer Systeme und Produkte. 2019.
- Verein Deutscher Ingenieure: VDI/VDE-Richtlinie 3693 Blatt 1, 2016-08: Virtuelle Inbetriebnahme. 2016.

Wallner, P.: Virtuelle Inbetriebnahme im Maschinenbau. In: *Computer und Automation* (1/2021), S. 18–20. 2021.

Westkämper, E.: Digitales Engineering von Fabriken und Prozessen. In: *Fertigungstechnisches Kolloquium (FTK) // Schriftliche Fassung der Vorträge zum Fertigungstechnischen Kolloquium am 10. und 11. September in Stuttgart ; [Tagungsband]*. Stuttgart: Ges. für Fertigungstechnik, 2008, S. 427–452.





*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Continuous Validation and Updating for High Accuracy of Digital Twins of Production Systems

Kontinuierliche Validierung und Aktualisierung für hohe Realitätsnähe von Digitalen Zwillingen von Produktionssystemen

Leonard Overbeck, Arthur Le Louarn, Oliver Brützel, Nicole Stricker, Gisela Lanza,
wbk Institute of Production Science, KIT, Karlsruhe (Germany),
leonard.overbeck@kit.edu, arthur.louarn@student.kit.edu,
oliver.bruetzel@kit.edu, nicole.stricker@kit.edu, gisela.lanza@kit.edu

Abstract: Despite continuous improvements in modelling, software tools and data availability, simulation projects of production systems still require a lot of manual effort, expertise in various disciplines and time. In many projects the high initial invest for building the simulation model is followed by a rather short period of experimentation and analysis. As production systems have to be adapted at an increasing pace to respond to rapidly changing markets and business environments, simulation models of these systems become outdated earlier, reducing their useful time window. One way to extend this time window would be the implementation of a method of automated comparison with the current production systems and subsequent self-adaption of the model to reality to maintain and even improve its accuracy over time. This approach will be presented and validated at a real world use case. Such an enhanced simulation model can be called a digital twin of the production system.

1 Introduction

Discrete-event simulation models (DES) permit the in-depth analysis and evaluation of improvement ideas on existing production systems without having to interfere with running production, which makes them a powerful tool for efficiency improvement of production (Mayer et al. 2020). Yet, in most companies simulation models of production systems are still built and used only in temporary projects (VDI, 2014). This leads to limited benefits by high initial costs, since simulation models require a lot of expertise and time to be creates and implements and even more to obtain satisfying accuracy. A longer usability would improve the return on investment of simulation models. But once a model is created, it constantly has to be adapted to changes in the real production system, if it shall be used over the whole life cycle of the production system for ongoing analysis and improvement. Since manual adaption is extremely time consuming, an approach of continuous validation of simulation

models and automated updating was developed. Validation is by VDI (VDI, 2014) defined as the “examination of the model as to whether the real behaviour of the modelled system is sufficiently well rendered with regard to the examination target” (part 1, p. 21). The continuous validation and update from real production data turn the simulation model into a real digital twin of the production system (Kritzinger et al., 2018).

2 Literature review

2.1 Modell generation and maintenance

Splanemann (1995) was one of the first to try semi-automated simulation model generation. His approach primarily uses CAD data in STEP-format (STandard for the Exchange of Product model data) to model the layout of the production system automatically. Focussing more on model parameters, Werner and Weigert (2002) proposed an approach to parametrize a model template, which was developed a-priori by experts, with data from ERP (Enterprise-Resource-Planning) and PDA (Production Data acquisition) systems and performed an analysis of model convergence to reality.

Bergmann and Straßburger (2020) presented different tools and methods to automatically generate simulation models which help designing a high-automated update process. One important step in this research field is the dissertation of Bergmann (2013) which uses the Core Manufacturing Simulation Data standard to create simulation models. Kotiades (2016) introduced the concept of a Self-Adaptive Discrete Event Simulation (SADES) but did not provide an exemplary implementation.

A recent and more elaborate overview of existing approaches is given by Reinhardt et al. (2019).

2.2 Data input for simulation models

Robinson and Perera (2002) provide an early discussion of chances and obstacles to automated data input, but IT-systems in production have changed a lot in the last 20 years. Skoogh et al. (2012) show how automated input data management can lead to time reduction and enhanced performance.

Several models and system architectures have been proposed to model the data exchange between physical and digital production systems. Those models are the foundation of the optimization and updating process of digital twins (Redelinghuys et al., 2020; Uhlemann et al., 2017). These works focus on the input side of the digital twin and updating, but do not discuss output validation and related automated update triggering in greater detail. Recent work of Goodall et al. (2019) presents a use case for data input in a remanufacturing facility.

2.3 Open research topic

Most of the existing approaches focus on automatic model generation. Some end up in a model translation, where the production system is modelled in a certain modelling style and then translated into an DES, which only decreases the modelling effort, if a model in the original modelling environment already exists (Terkaj and Urgo, 2015).

To tackle the problem of the need for initial modelling and because commercial simulation tools permit the easy and intuitive creation of simple models even for beginners, the presented approach chooses a different path: An existing model, which is manually modelled and implemented in a commercial simulation software, shall be enhanced by validation and update modules to turn it into a digital twin, which permits its use over the entire life time of the production system. The hypothesis is that the automated validation and updating can improve the initial models performance in terms of prediction accuracy.

3 Own Approach

The presented approach is explicitly aimed at simulation models of existing production systems, which shall be improved or controlled. It does not work for planning simulation models of production systems, which are not yet existing, since a comparison to reality and real data-based updates are impossible. Nevertheless, the approach can be used to transform planning simulation models into process accompanying simulation models during the building and commissioning phase of the production system.

3.1 Process flow

Production lines evolve over time and thus the input data needed for the simulation model, such as process times, availabilities, quality rates etc., change. Therefore, it is necessary to ensure that the digital twin always stays up-to-date and offers a close representation of reality in a given time period. The presented solution is composed of a two parts iterative process (Fig. 1): the validation and the automated updating procedure.

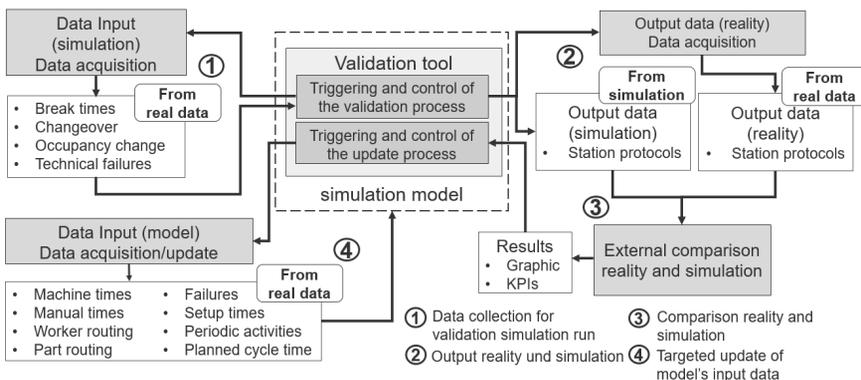


Figure 1: Iterative process of validation and automated updating

The simulation model itself is built and validated beforehand by simulation experts, following (VDI, 2014).

3.2 Validation

The objective of the validation is to automatically compare the simulation model with reality on different levels. The first step is to compare the output of the simulation model and reality by using carefully chosen key performance indicators (KPIs) and boundary values.

To evaluate the deviation of simulation runs to reality, the relative variation (see Eq. 1) and the NRMSE (Normalized Root Mean Square Error) (see Eq. 2) are used. The variation quantifies the final state of production of the studied period whereas the NRMSE quantifies the difference between reality and simulation during the course of the studied period.

$$\textit{Variation} = \frac{\|N_{real} - N_{sim}\|}{N_{real}} * 100 \quad (1)$$

with N_{real} , N_{sim} being the total amount of produced part at the end of the studied period respectively in reality and in simulation.

$$\textit{NRMSE} = \frac{1}{\bar{x}_{real}} * \sqrt{\frac{\sum_{i=1}^N (x_{real} - x_{sim})^2}{N_{real}}} \quad (2)$$

with x_{real} and x_{sim} representing the total amount of produced parts at each point in time t_i of the studied period, respectively for reality and simulation.

If the model output values deviate from the real output less than a predefined degree, this means that the digital twin satisfies the expectations and represents the reality to a satisfactory extent. In the case that outputs do not match, input values of the digital twin have to be examined in order to differentiate between input parameters that are still up-to-date and obsolete ones. According to these analysis results, the automated updating will be triggered precisely for the relevant parameters.

3.3 Update

In order for the automated updating process to be efficient, two prerequisites have to be fulfilled. A digital twin where the most effective input parameters are characterized as well as a data pipeline between data sources and simulation system are indispensable. Furthermore, the automated updating process allows replacing outdated data.

Once the update is performed, a simulation run is realized and the validation process is repeated to check the validity of the updated model. The whole process is repeated until the output is within the boundaries or until the digital twin cannot be further improved. In this case feedback is given to the user that an appropriate level of closeness could not be reached automatically and a manual intervention is necessary. An important outcome of this iterative process is to choose an appropriate time period for the data acquisition, that consequently gives the best compromise between data meaningfulness and acquisition effort while satisfying the performance criteria of the digital twin.

4 Use Case

The described approach was developed in a research partnership between of the wbk Institute for Production Science at the Karlsruhe Institute of Technology (KIT) and the central department Connected Manufacturing of the Bosch Powertrain Solutions division with the goal to develop an agile production system. Its application and validation are also part of this joint research project.

4.1 Production system

The exemplary production system, for which the digital twin is implemented, assembles car engine components in high volume and is composed of two areas which are connected via a conveyor. The two areas are assembly and testing, each semi-automated, following the Chaku-Chaku principle. This means that the machines perform their processes mainly in an automated manner and the workers are primary required for loading and unloading of machines and transporting parts between them. The line produces various product types with differing material flows, processing times, etc. The number of workers in each area varies due to external factors as vacations, sick days, reduced customer demand, trainings, etc. This has to be considered in the validation of the model. Historic production data from various sources is stored in a central data lake, including process times, change over delays, errors, scrap rates, etc. The software “Tecnomatix Plant Simulation” by Siemens is used to implement the digital twin.

4.2 Implementation

Using the approach described above, a validation tool that enables the validation and automated updating process is implemented. Before running the simulation model, the validation has to gather information about the system status at each point of time of the validation period from existing information systems such as MES and ERP. This includes the number of workers, produced product types and exceptionally long downtimes (more than one hour), that appear very rarely. If the simulation run would not consider this information, its comparison to reality would not be meaningful. The information about the number of workers in the production system at a certain period in time is not stored in the data lake, but in a different IT-System which is not accessible and therefore has to be added manually.

A python script preprocesses the real and simulation output data and compares them automatically. In the use case the chosen characteristic KPIs are: the progression of produced parts over time, the variance of the hourly overall equipment effectiveness (OEE) as well as the total OEE within the analyzed time period. These KPIs give an overview over the systems performance and keep track of the behavior of the digital twin during the whole simulation run. The permitted deviation of each KPI is decided according to the company’s performance goals and the systems inherent fluctuation. In the use case the corresponding threshold of permitted deviation shall not exceed 3% for the variation and 5 for NRMSE.

If the validation process results in higher deviations, another Python script performs the automated update by directly accessing the IT systems and data warehouses to obtain the latest input data. The data pipeline is composed of SQL queries and then filtered and processed into exploitable update data for the Plant Simulation software.

5 Results

Three experiments were conducted on two different weeks of production. The first experiment validates an input data set and the model's behaviour while the second and third experiment highlight the use of an automated targeted update to correct the input data and enable a better fitting of simulation with reality.

5.1 Automated validation

The first experiment was conducted for a production period of one week. To model the non-deterministic behaviour of the simulation, five simulation runs with different random seed values were conducted for each experiment with Plant Simulation Tecnomatix to get a statistical confidence of the results. Those five runs were considered sufficient as they well reflect the statistical repartition of the model while ensuring an acceptable optimized run time of the experiment. The automated validation compares each simulation run with reality and on the one hand returns figures (Figure 2) to help the user visualize the part's production during the production period. On the other hand, it generates key values to quantify the production systems behaviour (Table 1). Figure 2 shows a good fit of line output between the simulation runs and the reality for both the assembly and testing lines. This visual analysis is confirmed by the calculated key values from Table 1. The mean variation for both lines is under 3% and the mean NRMSE is under 5.

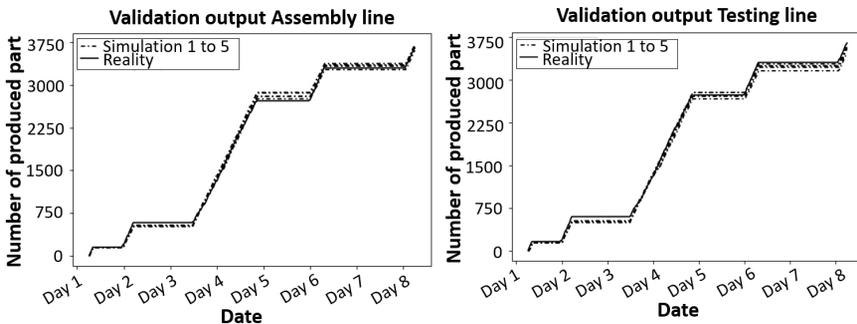


Figure 2: Validation output of assembly and testing line

Table 1: Results of automatic validation

Line	Produced parts reality	Mean produced parts simulation	Mean variation (%)	Mean NRMSE
Assembly	3650	3677	1.13	2.69
Testing	3639	3545	2.58	3.57

The fixed criteria from section 4.2 are therefore fulfilled and the input data is considered still up-to-date. The focus of the analysis lies on the number of produced parts since the OEE follows this number linearly.

5.2 Targeted update of the input parameters

For the second and third experiment, simulation and validation were conducted for another week of production. In the second experiment, the same input parameters as in section 5.1 were used. However, the obtained results before any update (Table 2, Figure 3) from the validation process exceeded the fixed threshold. Therefore, an update of input parameters is triggered. The first step of the update process is to determine which data must be replaced and if the line is partially or totally concerned by the update. The mean variation on both assembly and testing line are bigger than 3%, furthermore the NRMSE of the testing line is above 5. Consequently, both lines have to be updated.

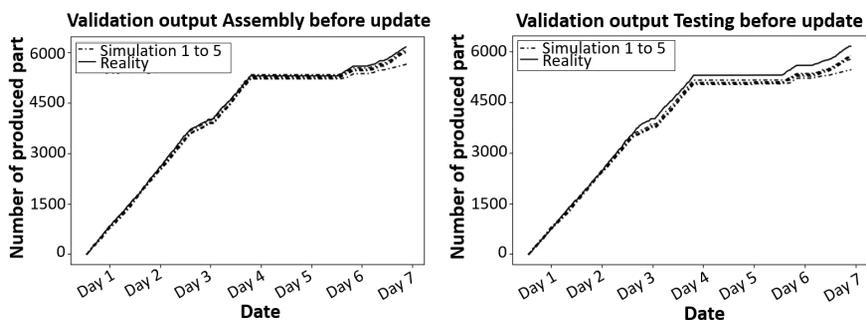


Figure 3: Validation output before update on assembly and testing line

Among the input data, it is possible to update the following parameters: Part routing, worker routing, failures, machine process times, manual process times, setup time and planned cycle times. Nonetheless, among those parameters few register notable variations during the chosen time period. In this paper, the focus was put on the machine process times, which encountered consequent variation over the studied week. After recalculating the probability density function of the machine process times from real data with a python script, the targeted update process compares the new calculated values with the old values for each machine. The machine process times are modelled by a normal distribution through mean and standard deviation. If the mean differs more than 0.15 seconds and the standard deviation more than 0.2, the old value is replaced with the new value. As mentioned above, in this use case the other input data did not change significantly and did not need any update.

Once the input parameters are updated, a third experiment with the newly calculated input data is conducted. Figure 4 depicts the output validation after the update for assembly and testing lines. Figure 4 shows improvement compared to Figure 3. The behaviour of the simulation is closer to reality and shows less variability. Those observations are verified through the key values in Table 2. For the assembly line, the mean variation of simulation went down from 3.03% to 0.92% and the mean NRMSE went from 2.99 to 1.77. For the testing line the mean variation went down from 6.32% to 3.2% and the NRMSE from 5.69 to 3.22. A net improvement is indeed realized. The behaviour of the assembly line is now completely validated whereas the testing line still has a mean variation barely above 3%. But the NRMSE has been improved and is now below 5. The machine process times could not be further improved for the

testing line. In a next step, other parameters of the simulation models, i.e. availabilities, scrap rate, etc. should be updated. For these parameters an automated update process is not yet implemented.

The capability of the targeted update process was nonetheless proved but still needs further improvement particularly concerning the threshold values and the trigger conditions for the targeted update mechanism.

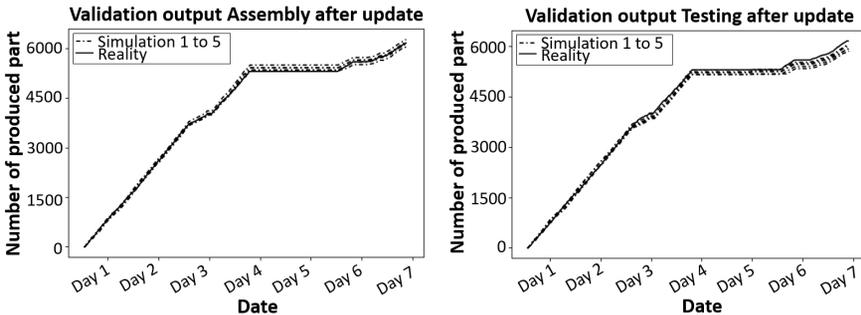


Figure 5: Validation output after update on assembly and testing line

Table 2: Validation metrics before and after automatic update

Experiment number	Line	Mean difference output (%)	Mean NRMSE output	Validation passed?
Before update	Assembly	3.03	2.99	No
After update	Assembly	0.92	1.77	Yes
Before update	Testing	6.32	5.69	No
After update	Testing	3.2	3.22	No

6 Conclusion and Outlook

Motivated by the ever-changing structure of modern production systems, an approach to enable simulation models to mirror these changes was developed. The approach contains a module for continuous validation which compares simulation KPIs to real historic KPIs. If a certain deviation threshold is surpassed, this module triggers an automated update module which changes the simulation model to better reflect reality. The application of this approach at a semi-automated production line of automotive components leads to a convergence of the simulation model to reality, turning it into a digital twin.

Further research has to be done to evaluate the behaviour of the digital twin in different scenarios of changes in the production system as well as its robustness to incomplete and/or biased data. Another line of research would be the extension of the

available update mechanism of the digital twin. This could be combined with a thorough examination of the validation KPIs and their thresholds.

References

- Bergmann, S.: Automatische Generierung adaptiver Modelle zur Simulation von Produktionssystemen. Illmenau, TU Illmenau, Dissertation, 2013.
- Bergmann, S.; Straßburger, S.: Automatische Modellgenerierung – Stand, Klassifizierung und ein Anwendungsbeispiel. In: Mayer, G.; Pöge, C.; Spieckermann, S.; Wenzel, S. (Hrsg.): Ablaufsimulation in der Automobilindustrie. Berlin, Heidelberg: Springer Berlin Heidelberg 2020, pp. 333–347.
- Goodall, P.; Sharpe, R.; West, A.: A Data-Driven Simulation to Support Remanufacturing Operations. *Computers in Industry* (2019) 105, pp. 48–60.
- Kotiades, K.: Towards self-adaptive discrete event simulation (SADES). In: Anagnostou, A.; Hoad, K.; Kunc, M. (Hrsg.): Operational Research Society Simulation Workshop 2016 (SW16) 2016,
- Kritzinger, W.; Karner, M.; Traar, G.; Henjes, J.; Sihm, W.: Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine* 51 (2018) 11, pp. 1016–1022.
- Mayer, G.; Pöge, C.; Spieckermann, S.; Wenzel, S. (Hrsg.): Ablaufsimulation in der Automobilindustrie. Berlin, Heidelberg: Springer Berlin Heidelberg 2020.
- Redelinghuys, A.J.; Basson, A.H.; Kruger, K.: A six-layer architecture for the digital twin: a manufacturing case study implementation. *Journal of Intelligent Manufacturing* 31 (2020) 6, pp. 1383–1402.
- Reinhardt, H.; Weber, M.; Putz, M.: A Survey on Automatic Model Generation for Material Flow Simulation in Discrete Manufacturing. *Procedia CIRP* 81 (2019) 2, pp. 121–126.
- Robinson, N.; Perera, T.: Automated data collection for simulation? *Simulation Practice and Theory* 9 (2002) 6-8, pp. 349–364.
- Skoogh, A.; Johansson, B.; Stahre, J.: Automated input data management: evaluation of a concept for reduced time consumption in discrete event simulation. *SIMULATION* 88 (2012) 11, pp. 1279–1293.
- Splanemann, R.: Teilautomatische Generierung von Simulationsmodellen aus systemneutral definierten Unternehmensdaten. Bremen, Universität Bremen, Institut für Betriebstechnik und angewandte Arbeitswissenschaften, Dissertation, 1995.
- Terkaj, W.; Urgo, M.: A Virtual Factory Data Model as a Support Tool for the Simulation of Manufacturing Systems. *Procedia CIRP* 28 (2015) 1, pp. 137–142.
- Uhlemann, T.H.-J.; Lehmann, C.; Steinhilper, R.: The Digital Twin: Realizing the Cyber-Physical Production System for Industry 4.0. *Procedia CIRP* 61 (2017), pp. 335–340.
- VDI: VDI-Richtlinie 3633: Simulation von Logistik-, Materialfluß- und Produktionssystemen, Berlin, Beuth, 2014,
- Werner, S.; Weigert, G.: Process accompanying simulation - a general approach for the continuous optimization of manufacturing schedules in electronics production. In: *Proceedings of the Winter Simulation Conference, San Diego, CA, USA, 8-11 Dec. 2002, 2002*, pp. 1903–1908.





*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Investigation and evaluation of 3D scanning methods for use cases in production planning

Erhebung und Evaluation von 3D-Aufnahmeverfahren für Anwendungsfälle in der Produktionsplanung

Maximilian Metzner, Svenja Schießel, Siemens AG, Erlangen (Germany),
maximilian.metzner@siemens.com, svenja.schiessel@siemens.com

Lukas Grünhöfer, FAU Erlangen-Nürnberg, Erlangen (Germany)

Jörg Franke, FAU Erlangen-Nürnberg, Lehrstuhl für Fertigungsautomatisierung und
Produktionssystematik (FAPS), Erlangen (Germany), joerg.franke@faps.fau.de

Abstract: Whilst production planning is mostly conducted using virtual design and simulation tools, most requirements and restrictions involve the physical shopfloor conditions. Hence, it is crucial for both greenfield and brownfield applications, as well as manual and automated production systems, to realistically capture the relevant features and to efficiently use them in the planning process. While both scanning hardware and processing software is widely available, a comprehensive collection of use-cases for digital production planning and simulation is not yet available. This contribution aims to close the gap by identifying use-cases for the application of 3D scanning methods in production systems, deriving requirements for the scan itself and comparing those to available technologies and procedures. The theoretical suitability is then validated using the identified approaches on real-world examples from an electronics manufacturing plant. From this, benefits of the implementation of 3D-scanning for production planning are exhibited and further research areas are identified.

1 Motivation for 3D reality capturing in production planning

Digital tools are a near compulsory part in production system planning. They are used throughout the process from the early concept design and material flow simulations to work cell design, ergonomics and kinematics simulation, robot offline programming, and virtual commissioning of the system. All these design and simulation methods are depending on an accurate digital spatial model of the system. Hence, all deviations between the virtual planning data and the real conditions inevitably lead to planning errors. Furthermore, these errors cannot be uncovered by even the most sophisticated simulation, validation or verification technique, as the

simulation itself is also based on the biased digital data (Rabe et al., 2008). Hence, there is a need for approaches to systematically verify the digital spatial planning data with the real world for simulations in production planning related use cases.

2 Related work

As the contribution focuses on the applicability of different scanning approaches for diverse use cases in production planning, related work in both scanning methods and technology as well as presented application scenarios is referred and discussed.

2.1 3D scanning and postprocessing

To capture 3D data in a production system relevant scale, mostly time-of-flight (ToF), stereoscopic and structure-from-motion (SfM) technologies are used (Seßner and Ramer, 2019; Singh et al., 2019). ToF systems can be distinguished into ToF depth cameras, repeatedly capturing a field of view and depth, often real-time in combination with an RGB image, and ToF scanners, which over time capture a panoramic hull around the scanner with large ranges (Gregor et al., 2008). Stereoscopic sensors use at least two 2D (RGB) cameras with defined offset to capture the same scene from two points of view and calculating the disparity of features in both frames to deduct the spatial position. (Tränkler, H.-R. and Reindl, L. 2014)

All the above scanning methods by themselves only create one depth frame from their current viewpoint, which is why other techniques for combining multiple scans are necessary. For static ToF scanners, the measuring points are either calibrated to each other via physical measurement, or the resulting scans are stitched together based on a minimal error fitting of overlapping areas (Besl and McKay, 1992). For dynamic depth cameras, simultaneous localization and mapping techniques (SLAM) are commonly used (Bailey and Durrant-Whyte, 2006).

SfM bases on feature matching from numerous RGB images of a scene from different viewpoints (Lowe, 2004; Nister and Stewenius, 2006). By this, the viewpoint of the camera for each image is estimated (Kneip et al., 2011). This knowledge is then used to spatially locate the matched features to construct a sparse 3D cloud (Farenzena et al., 2009; Schönberger et al., 2016). This sparse model is then densified using the additional information from the RGB images (Toldo, 2013). It should be noted that contrary to the other methods, SfM reconstructions are not scaled, so an object of known dimensions is required to adjust the scan's scale.

The result of either scanning method is usually an unordered, sometimes coloured, point cloud that needs further post processing for some use-cases. Common applied methods include filtering, segmentation, registration, and surface reconstruction.

2.2 Applications in production planning

Applications for production system planning have been discussed previously, for example by Gregor et al. (2008) and Lindskog et al. (2016). These mostly address the use case of comparing a planned layout with the available space in a brownfield application scenario. Newer approaches however also show promising usages of such technologies such as the integration in VR-based rough concept planning (Schäffer, 2021), for live-teleoperation of robots (Kohn et al., 2018) and calibration of offline generated robot programs for plug-and-produce robot systems (Metzner et al., 2021).

2.3 Research gap

The field of 3D scanning has made a technological leap in recent years. Scanning technology has evolved from costly, inflexible devices to a broad array of accessible, often cheap, or free alternatives that still yield acceptable results for some use cases. On the other side, potential for usage in planning and simulation has also evolved greatly, through capabilities such as VR but also a (r)evolution of production systems towards intelligent robotics.

However, neither the use cases, their requirements nor the appropriate scanning and processing methods have been systematically described. Hence this contribution aims to close this gap by gathering and analysing use cases, deriving their respective requirements and matching them with available technologies and methods. This provides a valuable overview to both scientists and practitioners on applications in their field and open need for action.

3 Use-cases and requirements for 3D scanning in production planning

First, an overview of exemplary use cases of 3D scanning in production planning is given, and the resulting requirements are derived.

3.1 Use-cases in the production lifecycle

Use-cases for 3D-scanning can be found in production planning, production engineering and during production or for adaptation. For some examples, see Fig. 1.

In production planning, two directions for data capture can be distinguished: the digitalization of the (brownfield) shopfloor as boundary conditions for new production systems and the digitalization of simplified concepts, e.g. carton mock ups, for further detailing or fitting with real-world conditions. Both concepts can be combined to allow a virtual fit analysis of a planned production concept.

Considered use cases in production include the design of a manual or hybrid work cell (UC1). This use-cases incorporates the digitalization of a cardboard simulation mock-up as basis for the mechanical design and digital ergonomics simulation, the combination of the mock-up with active kinematics simulation, e.g. robots, in VR-based workshops as well as the validation of the workshop results in a digital factory layout or through AR on the actual shopfloor.

The second use case is a line or factory scale scan to update shopfloor layout plans, highlight deviations between building plans and actual situation or plan infrastructure (UC2). These scans can be used for plant-scale material flow simulation, e.g. for automated guided vehicles. It is also a basis for the validation of use case one.

The third use case is the adaption and simulation of a brownfield work cell or line for which no comprehensive digital model is available (UC3). This can be the case for manual work cells basing on standard worktables with custom attachments or older machines for which no 3D model is provided.

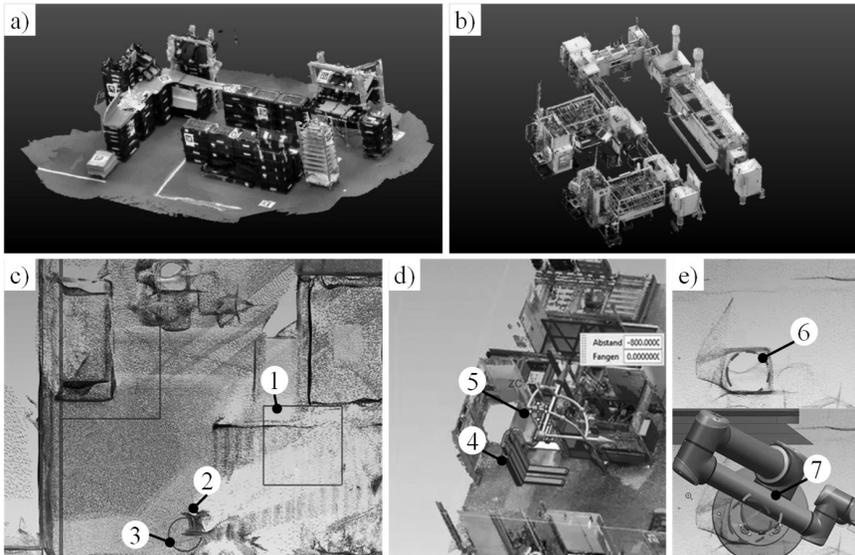


Figure 1: Application of 3D Scanning in the production lifecycle. (a) Scan of a carton mock-up. (b) Virtual model of a production line. (c) Verification of a 2D layout drawing featuring a pillar (1) which is really located at (2) and a non-existent hole (3). (d) Accessibility check for autonomous guided vehicle (4) to loading port in brownfield application (5). (e) Validation of “as planned”(7) and “as is”(6) state of a robot cell after setup.

A use-case in production engineering is the planning, simulation, and offline programming of robots (OLP) using a model adjusted through 3D scanning (UC4). This can be used to check the layout, to adjust interaction points or to check for collision potentials with surrounding objects. (Metzner et al., 2021) It can be furthermore used to simulate sensor data, e.g. to check coverages of safety LiDAR scanners incorporating the system’s surroundings. These use-cases are also relevant in production execution and system adaption, as those calibrated models can be used for OLP of new variant without the need for online-teach in (Metzner et al., 2021).

The last use case is the small-scale scanning of parts to either create spare parts using additive manufacturing or to use the realistic 3D model with texture and surface characteristics as a basis for synthetic training data generation for machine-learning based applications, e.g. for object detection, pose estimation or classification (UC5). Realistically textured parts are also necessary for simulation of vision-based processes, like software-in-the-loop simulation of bin picking systems.

3.2 Requirements for reality capturing

The requirements for each use case vary concerning necessary detailing/resolution, accuracy, completeness, handling of complexity and occlusions, readability of signs or notes, accessibility, and allowable data size. The highest requirements for resolution and accuracy are set by UC4 and UC5. Both the training data for ML models and the simulation model for robot programming require sub-millimetre accuracy and a resolution high enough to allow for smooth and detailed model

reconstruction. For UC3, accuracy should be below 1 centimetre since the model can be used to design interfaces with other system components or attachments to the system. Resolution is also very important here, since also fine structures are often relevant (cables and tubes, profiles, etc.). A similar requirement for accuracy can be stated for UC3. UC1 required the least accurate scans since the cardboard models themselves are mostly coarse replicas of the real system and a very accurate capturing does not provide additional value. However, since models in UC1 are often labelled and enhanced by notes, specifications or other written data, scan resolution and resulting readability of this information is crucial. The resolution in UC2 on the other hand is technically capped due to the size of the regions scanned. Accessibility restrictions are most common in UC2-4, as scans are taken on the shopfloor. Occlusions are most problematic for UC2 and UC4 due to their additional focus on the entire machine/line/factory, while UC3 is mostly focused on one part of it.

Table 1: Comparison of the requirements for the individual use-cases

Use-case	Resolution	Accuracy	Completeness	Handling of Occlusions	Readability
UC1	> 0,1 mm	< 50 mm	+	+	+
UC2	> 1 mm	< 10 mm	+	0	-
UC3	> 0,5 mm	< 10 mm	+	+	0
UC4	> 0,1 mm	< 1 mm	-	0	-
UC5	> 0,1 mm	< 1 mm	+	-	+

Another important factor for assessment is the required scanning and computation time for the different methods. For UC2-5, computation time is rather insignificant, as the scans can be prepared beforehand. For UC1 however, computation time is crucial, e.g. when workshop mock-ups are to be connected with a VR-based simulation, since this supposedly takes place in the same session. The same goes for scanning times, except for UC2, where the scale of the scanning area requires a rapid per-area scanning pace.

4 Validation and results

For each of the described use-cases, scans with all reasonably applicable technologies are taken and, if required, further processed. The results are then compared with focus on the above defined use case requirements. For each use-case, we then suggest one, or if applicable, multiple integrative methodologies on how to acquire the needed results. Based on the findings, we argue that certain use cases may require different scanning methods. Considering the one-time costs and scanning efforts, the use cases can be sorted into “low hanging fruits”, that are easily applicable for most productions yet still results in accuracy gains of planning and simulations, and more advanced ones only suitable for larger companies or special scenarios.

4.1 Experimental setup and validation scenarios

The validation scenarios are taken in an electronics manufacturing plant. For the stereo camera and ToF camera / SLAM setup, we use an Intel Realsense® D415 and

L515 depth camera with the Dot3D scanning software. ToF LiDAR / SLAM scanning is done using a Navvis® M6 scan trolley, which incorporates multiple LiDAR scanners but also cameras to calculate point colours. SfM is based on images from a Sony Alpha 6000 camera.

UC1 is validated on a carton mock-up for both a singular manual work cell as well as an assembly line mock-up. The scenario for UC2 is an electronics manufacturing plant with > 1000 m² production area incorporating both automated lines and manual work areas as well as logistics systems. UC3 is demonstrated on a production line featuring both manual work cells as well as automated machinery connected by a conveyor system. For UC4, a robot cell for electronics component assembly is chosen, see also (Metzner et al., 2021). The results are quantitatively compared to a ground truth model derived from a FARO Quantum V2/Blue LLP scanner, see Fig. 2. UC5 is demonstrated on a plastic housing component.

It is found that for UC1, due to the low accuracy requirements, all scanning methods are applicable. However, if the readability of notes is required, only SfM and depth camera /SLAM approaches are suitable. Since SfM requires significant computing times, a workshop-integration application is best supported with depth camera / SLAM scans. UC3, which covers a similar scan area, on the other hand has a higher accuracy requirement. Here, LiDAR / SLAM, SfM and ToF LiDAR /SLAM approaches are suitable. Existing systems are automatically captured in factory scans as of UC2, so data can also be derived from there. For UC2, due to the size of the scan area, only ToF LiDAR systems are generally feasible. While for UC4 scan size generally all scan methods are applicable, only SfM satisfies both resolution and accuracy requirements if the teach-in of points is planned, see Fig. 2. With SfMFor UC5, both high accuracy and resolution are required, which is why SfM methods, best in combination with a turntable, are most suited.

Table 2: Comparison of the features and characteristics of scanning technologies

Technology	Resolution	Accuracy	Completeness	Handling of Occlusions	Readability	Total Invest
Stereo camera / SLAM	> 0,1 mm	< 40 mm	0	0	0	0
ToF camera /SLAM	> 0,1 mm	< 15 mm	+	+	0	0
ToF LiDAR / SLAM	> 5 mm	< 5 mm	+	+	-	-
SfM (RGB camera)	> 0,1 mm	< 1 mm	-	+	+	+

Concerning economic applicability, the different scanning methods require vastly different investment. SfM approaches can be achieved at little to no cost, since simple RGB cameras, found in every smartphone, suffice. Free reconstruction software rendering good results is also available online. Depth camera / SLAM approaches require small investment, as general-purpose sensors are relatively cheap. In our experiments, we found that open-source SLAM software has some shortcomings concerning data quality and size, which is why some additional investment for

commercial scanning software is advisable. Total cost can be assumed in the low thousands of euros. ToF LiDAR scanners (terrestrial or portable) are priced significantly higher, often multiple ten thousand euros.

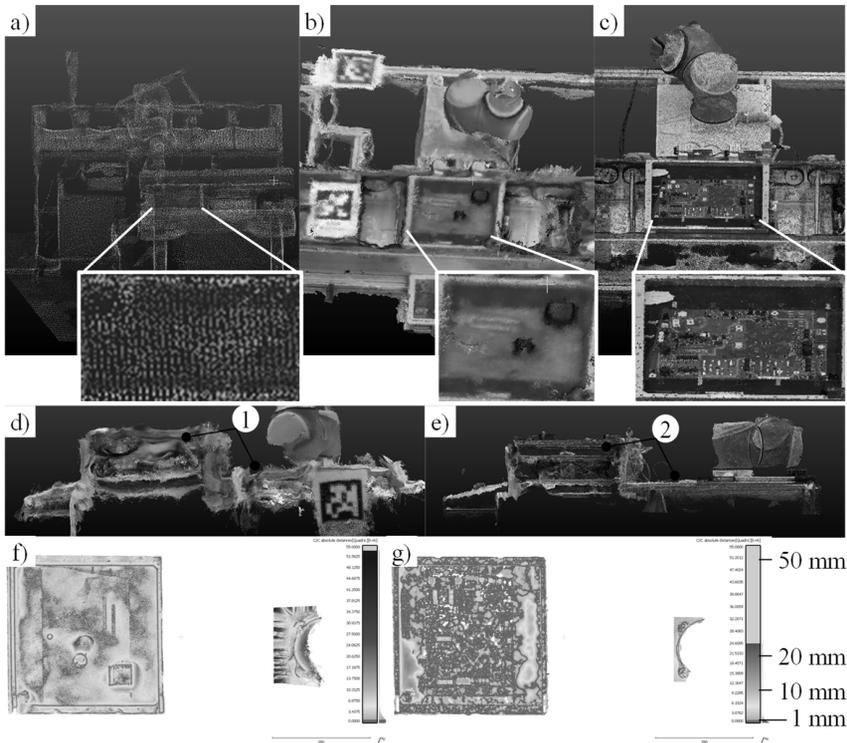


Figure 2: Comparison of scan resolution and accuracy from a ToF/SLAM factory scan (a), a local 3D camera scan (b) and SfM reconstruction (c), each with a closeup on the relevant region of interest for robot interaction. Comparison of noise levels from 3D camera scan (d) with noticeable biases (1) and SfM (e) with true geometry representation. Deviation heatmaps for 3D camera scan (f) and SfM (g) from ground truth model

4.2 Discussion and applicability for simulation

The results show that entry barriers for using 3D scanning in production planning and simulation are relatively low for most use-cases, except for UC2. However, as this use-case does not require frequent re-scanning, it can be contracted as a service for smaller enterprises. Applicability of the used data frequently requires post processing, which is mostly scanning-method independent. The scanning result as such is a point cloud, hence a collection of measured points with cartesian coordinates. While many planning and simulation tools offer native interfaces for point cloud integration, they mostly serve for visualization and interaction capabilities are very limited.

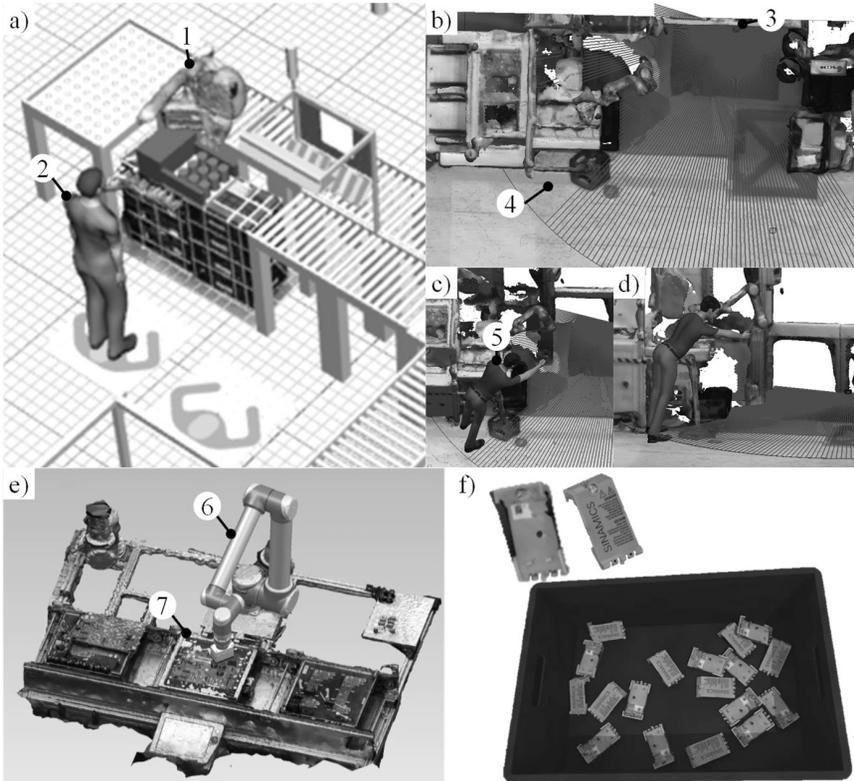


Figure 3: Examples of simulation integration of 3D scans. (a) Material flow simulation integrating a carton/box mock-up of a manual workstation (1) for layout and worker (2) walking distance evaluation. (b) Simulation of the coverage and dead zones (4) of a LiDAR safety sensor in an existing work cell. (c-d) Integration of a human mannequin (5) to simulate reachability without LiDAR detection. (e) OLP using a virtual robot (6) automatically fitted in an SfM scan (7) (Metzner et al., 2021). (f) Physically sound random box filling simulation based on 3D scanned parts as training data for 6D pose estimation convolutional neural network.

Hence it is mostly necessary to convert the scans into surface or volumetric models. For known objects, this can be done by fitting ideal CAD models into the point cloud. Rough fitting can be done manually or via matching routines, see e.g. Drost et al., 2010. For fine alignment, a quantitative error minimization is advisable, see e.g. Arun et al., 1987. The benefit of matching existing models into the scan is mainly the preservation of additional simulation-relevant data conserved in the model, such as kinematics and dynamics for process simulation or cycle times and availability for material flow simulation. Since not all objects in a system or factory are previously modelled, methods to reconstruct surfaces are necessary. For simple-shaped components such as walls, floor and pillars (planes) or pipes (cylinders), methods based on a random sample consensus can be applied, see e.g. Fischler and Bolles, 1981. For all non-modelled, complex-shaped components, such as modular manual workstations, surface reconstruction methods such as the (Screened) Poisson method

can be used (Kazhdan and Hoppe, 2013). These surface models can be used as basis for further modelling, for physics-based or kinematic simulation or for collision analysis.

5 Conclusions and outlook

In this contribution, use-cases for 3D scanning in production planning and simulation are derived. The usability of common scanning methods, as well as benefits and drawbacks are detailed. Furthermore, post-processing steps for integration into simulation tools are discussed. Further research focus should lie on the automatic reconstruction of models from scans integrating known objects matching and reconstruction methods. This will accelerate simulation model generation while also increasing validity, as the real boundary conditions from the shopfloor can be integrated into the simulation rather than ideal assumptions and models. This trend will furthermore be fuelled by more readily available scanning technology, as can be observed in first consumer smartphones and tablets that incorporate stereo-based and, more recently, LiDAR-based depth cameras. The increased distribution of such devices will likely foster the applicability of 3D scanning and as a result, more virtual planning and simulation even in smallest production scenarios.

References

- Arun, K.S.; Huang, T.S.; Blostein, S.D.: Least-squares fitting of two 3-d point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9 (1987) 5, pp. 698–700.
- Bailey, T.; Durrant-Whyte, H.: Simultaneous localization and mapping (SLAM): part II. *IEEE Robotics & Automation Magazine* 13 (2006) 3, pp. 108–117.
- Besl, P.J.; McKay, N.D.: A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14 (1992) 2, pp. 239–256.
- Drost, B.; Ulrich, M.; Navab, N.; Ilic, S.: Model globally, match locally: Efficient and robust 3D object recognition. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13.06.2010, 2010, pp. 998–1005.
- Farenzena, M.; Fusiello, A.; Gherardi, R.: Structure-and-motion pipeline on a hierarchical cluster tree. In: 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops 2009, pp. 1489–1496.
- Fischler, M.A.; Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24 (1981) 6, pp. 381–395.
- Gregor, P.I.M.; Budzel, I.F.; Štefánik, I.A.; Plinta, I.D.: 3D Laser Scanning in Digitization of Current Production Systems. *IFAC Proceedings Volumes* 41 (2008) 3, pp. 86–93.
- Kazhdan, M.; Hoppe, H.: Screened poisson surface reconstruction. *ACM Transactions on Graphics (ToG)* 32 (2013) 3, pp. 1–13.

- Kneip, L.; Scaramuzza, D.; Siegart, R.: A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation. In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011), Colorado Springs, CO, USA, 6/20/2011 - 6/25/2011, 2011, pp. 2969–2976.
- Kohn, S.; Blank, A.; Puljiz, D.; Zenkel, L.; Bieber, O.; Hein, B.; Franke, J.: Towards a Real-Time Environment Reconstruction for VR-Based Teleoperation Through Model Segmentation. 2018 IEEE International Conference on Intelligent Robots and Systems (IROS) (2018).
- Lindskog, E.; Berglund, J.; Vallhagen, J.; Johansson, B.: Layout Planning and Geometry Analysis Using 3D Laser Scanning in Production System Redesign. *Procedia CIRP* 44 (2016), pp. 126–131.
- Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* 60 (2004) 2, pp. 91–110.
- Metzner, M.; Reisinger, D.; Ortmann, J.-N.; Grünhöfer, L.; Handwerker, A.; Blank, A.: An approach for direct offline programming of high precision assembly tasks on 3D scans using tactile control and automatic program adaption. In: Tagungsband des 6. Fachkolloquiums Montage, Handhabung und Industrierobotik (accepted for publication) 2021,
- Nister, D.; Stewenius, H.: Scalable Recognition with a Vocabulary Tree. In: CVPRW '06, New York, NY, USA, 17-22 June 2006, 2006, pp. 2161–2168.
- Rabe, M.; Spieckermann, S.; Wenzel, S.: Verifikation und Validierung für die Simulation in Produktion und Logistik. Berlin, Heidelberg: Springer Berlin Heidelberg 2008.
- Schäffer, E.W.: Web- und wissensbasierter Engineering-Konfigurator für roboterzentrierte Automatisierungslösungen, FAU University Press, 2021.
- Schönberger, J.L.; Zheng, E.; Pollefeys, M.; Frahm, J.-M.: Pixelwise View Selection for Unstructured Multi-View Stereo. In: European Conference on Computer Vision (ECCV), 2016,
- Seßner, J.; Ramer, C.: Sensortechnik. In: Rainer Müller; Jörg Franke; Dominik Henrich; Bernd Kuhlenkötter; Annika Raatz; Alexander Verl (Hrsg.): Handbuch Mensch-Roboter-Kollaboration: Optische Sensoren: Carl Hanser Verlag GmbH & Co. KG 2019, pp. 71–115.
- Singh, Y.; Toldo, R.; Magri, L.; Fantoni, S.; Fusiello, A. Method for 3D modelling based on structure from motion processing of sparse 2D images, 2019: Method for 3D modelling based on structure from motion processing of sparse 2D images: Google Patents.
- Toldo, R.: Towards automatic acquisition of high-level 3D models from images. *Universita degli Studi di Verona, Verona* (2013).
- Tränkler, H.-R.; Reindl, L. (Hrsg.): Sensortechnik: Handbuch für Praxis und Wissenschaft. Berlin, Heidelberg: Springer Berlin Heidelberg 2014.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Physikalische Sensorsimulation zur Verifikation und Validierung optischer Systeme

Physical Sensor Simulation for the Verification and Validation of Optical Systems

Jörn Thieling und Jürgen Roßmann, RWTH Aachen, Aachen (Germany),
thieling@mml.rwth-aachen.de, rossmann@mml.rwth-aachen.de

Abstract: Optical systems are a key technology for highly automated production processes including inline inspection systems and driverless transport systems. All of them require comprehensive verification and validation to ensure their functionality, safety and reliability. Often this can only be fully determined too late—i.e. after implementation in the production process. To allow early and continuous system tests we propose the use of Digital Twins and Hybrid Testbeds that combine simulated and real components as available or required. To this end, a scalable physical sensor simulation is required, that on the one hand enables a flexible configuration and simulation of optical systems within their operational environments, and on the other hand, generates physically reliable sensor data for realistic simulation results. This contribution addresses the resulting complexity and presents methods that handle this by means of modular and physical models of intrinsic sensor components as well as extrinsic sensor-environment interactions.

1 Einleitung

Optische Systeme sind nicht nur die Schlüsseltechnologie zur Realisierung von Inline-Prüfsystemen, sondern auch der Wahrnehmungssinn von Fahrerlosen Transportsystemen. Damit leisten sie einen wesentlichen Beitrag zur Kontrolle und Automatisierung des Produktionsprozesses und legen den Grundstein für eine Produktion von hochwertigen Produkten zu geringen Kosten und kurzen Produktionszeiten. Allerdings ist jede Messung von optischen Systemen aufgrund von einer Vielzahl von Einflussfaktoren inhärent unsicherheitsbehaftet und durch verschiedene Störgrößen und Limitierungen des Messprozesses verfälscht. Einflüsse können dabei sowohl extrinsischer als auch intrinsischer Natur sein. Aufgrund der Komplexität der verschiedenen, miteinander korrelierenden Einflussfaktoren ist die Verifikation und Validierung von optischen Systemen und der auf diesen basierenden Gesamtsystemen stets mit einem großen Aufwand verbunden, der vielzählige Tests am realen System und in einem realitätsnahen Einsatzumfeld erfordert. Häufig kann die funktionale Validierung des realen optischen Systems erst nach der

Implementierung in den Produktionsprozess vollständig ermittelt werden, was potenziell zu kostenintensiven Iterationen zur Behebung von Design- und Implementationsfehlern führt. In diesem Beitrag wird die simulationsbasierte Verifikation und Validierung von optischen Systemen fokussiert, wohingegen die Methoden zur Überprüfung der Gültigkeit der zugrundeliegenden Modelle und Simulationen in (Thieling und Roßmann 2021) behandelt wurden.

1.1 Verwandte Arbeiten

Um bereits in einem frühen Stadium der Entwicklung von optischen Systemen einen Einblick in deren Funktionalität und Anwendbarkeit zu erlangen, bietet sich der ergänzende Einsatz von Simulationstechnik an. So liefert der aktuelle Stand der Technik verschiedenste Ansätze zur Simulation in Produktion und Logistik. Kommerzielle Beispiele für die industrielle Anwendung sind Tecnomatix Plant Simulation und FlexSim, welche prinzipiell Prozess- und Logistik-Simulationen ermöglichen aber den Sensoraspekt moderner Produktionsanlagen komplett vernachlässigen. Auch mit Blick auf robotische Anwendungen – bei denen Sensoren in Zeiten von intensiver Automation und autonomen Robotern unabdingbar sind – erlauben Systeme wie Microsoft Robotics Studio, Gazebo (Open Source Robotics Foundation 2021), Webots (Cyberbotics 2021) und USARSim (Carpin et al. 2007) nur in engen Grenzen eine detaillierte und validierte Simulation von Sensorsystemen und den zugehörigen komplexen Regelkreisen. Häufig basiert die Simulation optischer Sensoren auf Gaming Engines wie Unreal (Epic 2021), Unity (Unity 2021) oder CryEngine (Crytek 2021), welche alle zur Visualisierung von Computerspielen und Spezialeffekten entwickelt wurden. Dementsprechend sind diese Ansätze zwar echtzeitfähig und erscheinen physikalisch plausibel. Sie sind jedoch nicht physikalisch korrekt im Sinne einer ausreichend detaillierten und validierten Simulation optischer Systeme. Daher erlauben sie keine physikalisch belastbare Aussage zur Verifikation und Validierung von optischen Systemen.

Auch die Betrachtung aktueller Forschungsergebnisse in diesem Bereich liefert überwiegend Ansätze, die lediglich ein Teilproblem der Simulation optischer Systeme behandeln. So existieren beispielsweise Ansätze zur spezifischen Simulation einzelner Sensoren (Mohammadikaji et al. 2020; Elmadawi et al. 2019; Kolchaev et al. 2017; Hirsenkorn et al. 2017; Douchin et al. 2016; Liu 2013) als auch zur modell- und herstellerübergreifenden Simulation generischer Sensoren (Stolz und Nestlinger 2018; Velazquez et al. 2018; Hanke et al. 2015).

Der Nachteil nahezu all dieser Ansätze ist der Verlust an Flexibilität oder Realitätsnähe. Entweder fokussieren sie sich auf die detaillierte Simulation eines herstellereinspezifischen Sensors und erlauben keine generische Modellierung weiterer Varianten oder sie unterstützen die Simulation zahlreicher Sensoren aber die hierbei generierten Ergebnisse sind realitätsfern aufgrund nicht-physikalischer Modelle.

1.2 Überblick

In unserem Beitrag stellen wir den Einsatz von Digitalen Zwillingen und Hybriden Testbeds vor. Diese modellieren und simulieren alle relevanten Komponenten und Einflussfaktoren des optischen Systems einschließlich seines geplanten Einsatzumfeldes. Durch Kombination von simulierten und realen Komponenten schaffen Hybride Testbeds nicht nur eine frühzeitige Möglichkeit zur funktionalen

Validierung mittels simulierter Einsatzszenarien, sondern erlauben zugleich einen nahtlosen Transfer von einem ersten digitalen Prototyp zum finalen, realen System (Thieling und Roßmann 2020). In diesem Kontext umfassen Digitale Zwillinge – als 1:1-Replikat von realen Systemen in der digitalen Welt – alle relevanten Systemkomponenten einschließlich ihrer Funktionen und Schnittstellen und werden durch den Einsatz von 3D-Simulationen zum Leben erweckt.

Einen wesentlichen Beitrag zur Realisierung dieser Digitalen Zwillinge und Hybriden Testbeds liefert die in diesem Beitrag behandelte physikalische Sensorsimulation. Diese modelliert und simuliert sowohl die Sensor-Umfeld Interaktion als auch die intrinsischen Sensorkomponenten und adressiert somit den größten und komplexesten Teil eines optischen Systems. Dazu liefern die folgenden Abschnitte einen Einblick in das übergeordnete Konzept der skalierbaren physikalischen Sensorsimulation (Abschnitt 2), behandeln die im Kontext von Produktion und Logistik besonders relevanten Einflüsse und Störgrößen anhand zugehöriger extrinsischer und intrinsischer Modelle (Abschnitt 3 und 4) und zeigen deren Anwendung in bereichsübergreifenden Einsatzszenarien (Abschnitt 5).

2 Skalierbare physikalische Sensorsimulation

Zur Realisierung der hier vorgeschlagenen physikalischen Sensorsimulation wird das in (Thieling und Roßmann 2020) eingeleitete Konzept der Dreiachsigen Skalierbarkeit angewandt. Dieses verfolgt eine Skalierbarkeit von Modellen und Simulationsverfahren entlang der Dimensionen: Simulationstiefe, Sensortypen & Anwendungen und Prozessoren (vgl. Abbildung 1 links).

Grundlage hierfür ist, dass die notwendigen Sensor- und Fehlermodelle als wiederverwendbare Basiselemente implementiert werden, die sich allein durch Parametrieren und Konfigurieren auf ein spezifisches Sensorsystem münzen lassen. Realisiert wird dies durch die Verwendung von weitestgehend generischen und physikalischen Schnittstellen und Modellen (vgl. Abbildung 1 rechts), die sich allein anhand der vorliegenden Kenngrößen (Spezifikation des Sensorsystems) und/oder durch Verfahren zur Parameteridentifikation (auf Basis realer Referenzaufnahmen) passend zum spezifischen Verhalten konfigurieren lassen.

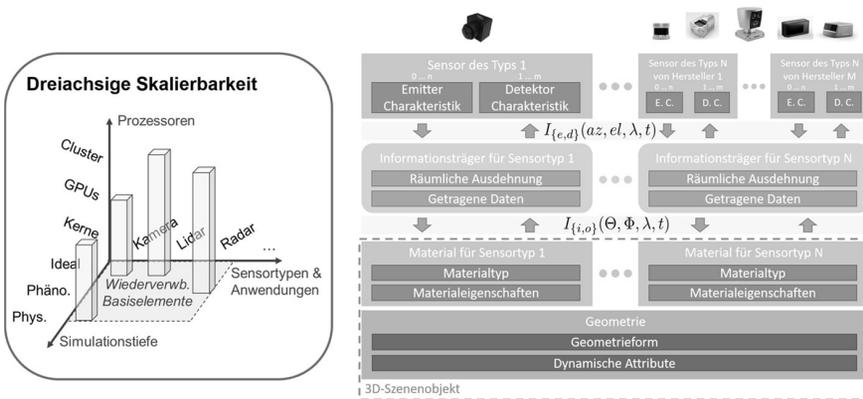


Abbildung 1: Konzept der Dreiaxialen Skalierbarkeit (links) und physikalische Architektur zur skalierbaren (multimodalen) Sensor-Umfeld Interaktion (rechts). Tiefergehende Erläuterungen finden sich in (Thieling und Roßmann 2020).

Mit Blick auf die Realisierung von Hybriden Testbeds – bestehend aus zum Teil simulierten und zum Teil realen Komponenten – ist eine weitere Grundlage die Skalierbarkeit über verfügbare Prozessoren. Diese erlaubt erst eine echtzeitfähige Simulation durch Parallelisierung des entstehenden Rechenaufwands über alle verfügbaren Rechenressourcen. Realisiert wird dies durch die im Kern parallelisierte Implementierung der wiederverwendbaren Basiselemente, die ihre Arbeitspakete als Funktionen über die sensorzentrischen Dimensionen Azimut, Elevation und Zeit beschreiben und so die Berechnungen auf tiefer Ebene voneinander entkoppeln. Im Detail wird hierbei der zur Simulation benötigte Rechenaufwand über tausende Kerne eines Grafikprozessors (GPU) und allen verfügbaren GPUs eines Servers parallelisiert, was die Generierung simulierter Sensordaten mit akzeptablen Rechenzeiten garantiert – ohne zwangsweise die Simulationstiefe zu reduzieren.

Stehen die für eine echtzeitfähige und zugleich physikalische Simulation benötigte Anzahl an Prozessoren nicht zur Verfügung oder liegen die Modelle nur als ideale oder phänomenologische Modelle vor, ist die letzte Achse der Skalierbarkeit von Interesse, die einen nahtlosen Übergang zwischen verschiedenen Simulationstiefen repräsentiert. Realisiert wird diese durch stets physikalische (d. h. simulationsverfahren-unabhängige) Schnittstellen zwischen den einzelnen Modellen. Dies umfasst die in Abbildung 1 (rechts) gezeigten intensitätsbasierten Schnittstellen, die unabhängig vom sensorspezifischen Informationsträger die Interaktion auf Basis der emittierten und detektierten Intensitäten (I_e, I_d) respektive der einfallenden und ausgehenden Intensitäten (I_i, I_o) beschreiben – stets definiert über alle Raumwinkel (az, el) bzw. (Θ, Φ) und in Abhängigkeit von der Wellenlänge λ und Zeit t .

In den folgenden Abschnitten wird unter Berücksichtigung des hier eingeleiteten übergeordneten Konzepts die Modellierung und Simulation von ausgewählten physikalischen Effekten beleuchtet und dabei der Fokus auf jene Effekte gelegt, die einen wesentlichen Einfluss auf die Verifikation und Validierung von optischen Systemen insbesondere im Bereich von Produktion und Logistik nehmen.

3 Simulation extrinsischer Einflüsse und Störgrößen

Mit Blick auf die extrinsischen Einflüsse und Störgrößen steht vor allem die Ausbreitung des Informationsträgers (z. B. Licht) und dessen Interaktion mit den Objekten der Szene im Fokus. Neben der Richtcharakteristik der relevanten Emitter und den Geometrieformen der umgebenden Objekte wird die Sensor-Umfeld Interaktion insbesondere von der materialspezifischen Streuung und der dynamischen Bewegung von Sensor und 3D-Szenenobjekten beeinflusst.

3.1 Mehrwegeausbreitung des Informationsträgers

Für eine realitätsnahe Simulation der Mehrwegeausbreitung der Informationsträger (z. B. Licht) werden Emitter (z. B. Lichtquellen) unter anderem hinsichtlich ihrer Lage und Richtcharakteristik modelliert. Um beliebige Emitter zu unterstützen, lässt sich die Richtcharakteristik durch benutzerdefinierte Emissionsverteilungsfunktionen (EDFs) integrieren, die die ursprüngliche Intensität des Emitters in Abhängigkeit von dem Raumwinkel (az, el) und der Wellenlänge λ gewichtet und so deren Modellierung generalisiert (vgl. Abbildung 2).

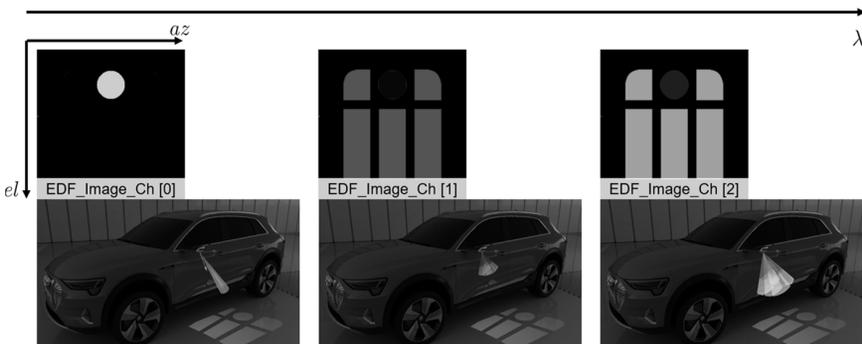


Abbildung 2: Benutzerdefinierte Emissionsverteilungsfunktion (EDF) mit drei Kanälen bestehend aus Rot, Grün und Blau des sichtbaren Lichtspektrums.

Sind die relevanten Emitter auf diese Weise entweder als Teil des Sensors (z. B. für aktive Lidar-Sensoren) oder als Teil der Szene (z. B. für passive Kameras) modelliert, wird die Ausbreitung des Informationsträgers durch die dreidimensionale Szene mittels Strahlen verfolgt. Da das direkte Umfeld eines Sensors den größten Einfluss auf die Sensordaten hat, erfolgt dies entgegengesetzt zur eigentlichen Ausbreitungsrichtung – also ausgehend vom Detektor durch die Szene bis zu den relevanten Emittlern.

3.2 Materialspezifische Streuung

Besonders wichtig für eine realitätsnahe Simulation der Mehrwegeausbreitung ist die Modellierung der materialspezifischen Streuung. Zur Generalisierung der Materialmodelle werden diese neben dem material- und wellenlängenspezifischen Brechungsindex ior und Absorptionskoeffizienten a auf Oberflächenebene durch ihre Bidirektionale Streuungsverteilungsfunktion (BSDF) definiert (vgl. Abbildung 3).

Letztere beschreibt sowohl den reflektierten, als auch transmittierten Anteil der einfallenden Intensitäten in Abhängigkeit der Wellenlänge λ und ist definiert über alle Raumwinkel (Θ, Φ) des Oberflächenreferenzkoordinatensystems – bestehend aus Oberflächenpunkt \underline{p} , Normale \underline{n} , Tangente \underline{u} und Bitangente \underline{v} . Grundlegend sind hier parametrische Funktionen zur Definition und Kombination einer diffusen, glänzenden und spiegelnden Reflexion und/oder Transmission (vgl. Abbildung 3).

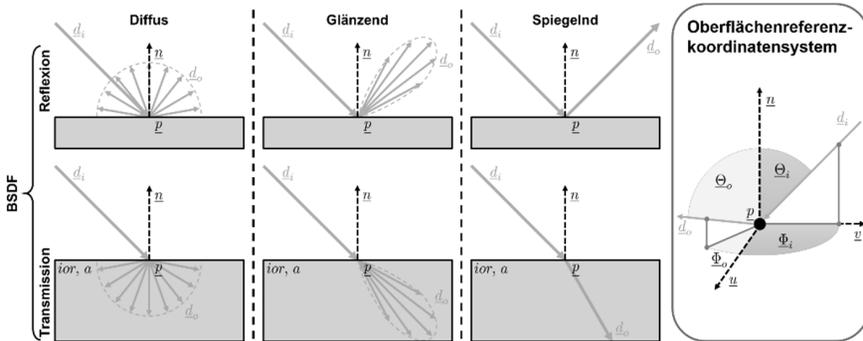


Abbildung 3: Basiselemente einer Bidirektionalen Streuungsverteilungsfunktion (BSDF) links mit zugehörigem Oberflächenreferenzkoordinatensystem rechts.

So lassen sich durch Kombination wiederverwendbarer Basiselemente beliebige Oberflächenvariationen als Materialmodelle realisieren, die wiederum zur Umhüllung beliebiger (Teil-) Geometrien dienen (vgl. Abbildung 1 rechts) und so die Schnittstelle zwischen sensorspezifischen Informationsträgern und 3D-Szenenobjekten definieren.

3.3 Dynamische Bewegung von Sensor und Umfeld

Auch die dynamische Bewegung von Sensor und Umfeld nehmen einen entscheidenden Einfluss auf die einzelnen Messwerte (z. B. Punktwolke eines Lidars). So beeinflusst die Bewegung einzelner Objekte nicht nur deren Lage und führt zu einer ggf. verzerrten Wahrnehmung aus Sicht des Sensors, sondern verursacht ebenso eine Doppler-Verschiebung in der Frequenz des reflektierten Informationsträgers.

Um beide Effekte in der Simulation zu respektieren, wird bei der Sensor-Umfeld Interaktion während eines Simulationszeitschritts sowohl die Lage der Objekte, als auch deren Geschwindigkeitsvektoren ausgewertet. Dies betrifft zum einen die Bestimmung von Schnittpunkten zwischen Informationsträgern und Objekten der Szene und zum anderen die Berechnung der Frequenzverschiebung im Zuge der materialspezifischen Streuung. Die Simulation dieser Dopplerverschiebung ist insb. für die Simulation von Frequency Modulated Continuous Wave (FMCW)-Radar- und FMCW-Lidar-Sensoren relevant.

4 Simulation intrinsischer Einflüsse und Störgrößen

Mit Blick auf die intrinsischen Einflüsse und Störgrößen steht vor allem die Abtastung des Sensorsystems im Fokus, welche sich in erster Linie aus der Lage des Detektors und dem Design einer etwaigen Scan-Mechanik (z. B. bei Lidar-Sensoren) ergibt.

Neben dem nominalen Abtastmuster in Form einer Liste an Raumwinkeln (az, el) definiert als Funktion der Zeit t wird die sensorspezifische Wahrnehmung u. a. von der Verzerrung, Tiefenschärfe und Verschmutzung des Linsensystems beeinflusst.

4.1 Linsenverzerrung

Das wohl bekannteste Beispiel einer Linsenverzerrung findet sich im Kontext einer Kamera, welche im Gegensatz zu einem idealen Lochkamera-Modell stets mit einer mehr oder weniger stark verzerrten Abbildung auf den Detektor verbunden ist. Ursächlich ist stets die Brechung des Informationsträgers (z. B. Licht) am Übergang zwischen Luft und Linsenmaterial (z. B. Glas), deren Ausmaße durch die Wölbung der einzelnen Linsen kontrolliert wird.

Da Linsen nicht nur in Kameras, sondern auch in weiteren optischen Systemen Anwendung finden (z. B. Lidar oder Radar) werden zugehörige Linsenmodelle dahingehend generalisiert, dass sie auf Basis der im Sensorkoordinatensystem gegebenen nominalen Abtastrichtung \underline{d} , eine entsprechend verzerrte Abtastrichtung \underline{d}_d berechnen (vgl. Abbildung 4 links).

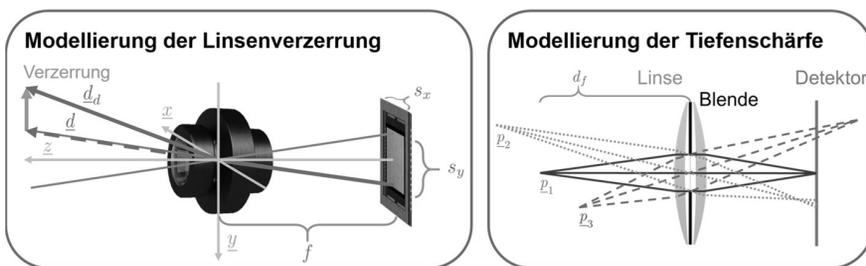


Abbildung 4: Modellierung der Linsenverzerrung (links) und Tiefenschärfe (rechts).

Grundlage hierfür bilden die Modelle von Brown-Conrady (1971) und Kannala (2006), die im Falle der Kamera, die Parameter des reinen Lochkamera-Modells (d. h. Brennweite f und Detektormaße s_x, s_y) entsprechend erweitern. Dabei ist Kannala nicht auf eine Bildebene limitiert, sondern erlaubt die Modellierung von Verzerrungen, die zu einem Sichtfeld von mehr als 180 Grad führen und insbesondere für die Simulation von Weitwinkellinsen benötigt werden. Die hierzu benötigten intrinsischen Parameter lassen sich entweder dem Sensor-Design (der Spezifikation) entnehmen und/oder mittels realer Referenzexperimente bestimmen – entsprechende Kalibrierungsprozeduren liefern die zuvor referenzierten Quellen.

4.2 Tiefenschärfe und Bewegungsunschärfe

Die eigentliche Funktion der Linse ist die Bündelung/Fokussierung von eingehenden Informationsträgern (z. B. Licht) auf einzelne Bereiche des Detektors. Dies erfordert eine endlich große Blendenöffnung, die den Eintritt einer messbaren Menge dieser erlaubt und bei der Abbildung von Objektpunkten der Szene zu mehreren möglichen Wegen durch das Linsensystem führt (vgl. Abbildung 4 rechts).

Um den hierdurch entstehenden Effekte der Tiefenschärfe zu modellieren, wird die Größe der Blendenöffnung sowie die Fokusbrennweite d_f der Linse (Distanz zu dem im

Fokus liegenden Raumpunkt; hier p_1) für ein parametrisches Modell verwendet. Dieses manipuliert die (verzerrte) Abtastrichtung \underline{d}_d dahingehend, dass der Ursprung des gedachten Strahls gleichverteilt innerhalb der Blendenöffnung verschoben und im Anschluss erneut auf den Fokuspunkt (definiert durch die Fokusbildung d_f entlang der ursprünglichen Abtastrichtung \underline{d}_d) ausgerichtet wird. Durch Mittelung über mehrere Iterationen wird so die Flächenintegration approximiert und die Entstehung der Tiefenschärfe realitätsnah simuliert. Durch Berücksichtigung des Linsengehäuses als Teil der Szene wird hierbei zugleich die Vignettierung simuliert, die aus der zu den Rändern hin zunehmenden Verschattung des Detektors resultiert.

Ähnlich zur Tiefenschärfe ist die Bewegungsunschärfe. Zur Simulation dieser findet während eines Simulationsschritts, analog zur räumlichen Unterabtastung der Blendenöffnung, eine zeitliche Unterabtastung der Detektor-Integrationszeit statt. Dadurch führen auch in der Simulation schnelle Bewegungen zwischen Sensor und Umfeld, zu einer realitätsnah „verschmierten“ Abbildung eines Objekts auf mehrere Bereiche des Detektors (z. B. Kamerapixel).

4.3 Linsenverschmutzung

Schlussendlich ist auch die Simulation der Linsenverschmutzung relevant, um die Robustheit des Systems für durch Staub belastete Einsatzumgebungen zu validieren. Die Simulation erfolgt hier sehr ähnlich zu der einer Oberflächenverschmutzung – modelliert als Material. Diesmal kommt jedoch der Transmission des Materials (Schmutz auf Linsenoberfläche) eine besondere Bedeutung zu. Verschmutzungen der Linse äußern sich vor allem durch ein verschwommenes Abbild der Szene und durch weiche Kante der Verschmutzung selbst (vgl. Abbildung 5 oben rechts).

Während sich Ersterer durch gewichtete Schichten von diffuser und glänzender Transmission in Form einer entsprechenden BSDF modellieren lassen (vgl. Abbildung 3), ist Letztere implizit durch die oben eingeführte Tiefenschärfe im Ergebnisbild enthalten. Das heißt, obgleich die Verschmutzung mit scharfer Kontur modelliert wird, ist deren Wahrnehmung vom Detektor durch die Tiefenschärfe entsprechend verschwommen.

5 Anwendungen

Durch die oben präsentierten Konzepte und Modelle zur skalierbaren, physikalischen Sensorsimulation eröffnen sich sensortypen- und anwendungsübergreifende Einsatzmöglichkeiten, die einen omnidirektionalen Technologietransfer erlauben. Im Fokus steht hier die simulationsbasierte Verifikation und Validierung von optischen Systemen als integrierter Bestandteil des entstehenden Gesamtsystems inklusive des geplanten Einsatzumfelds (vgl. Abbildung 5).

So zeigt Abbildung 5 die Ergebnisse der physikalischen Sensorsimulation für die simulationsbasierte Verifikation und Validierung einer Lidar-basierten Lageregelung beim Anflug an die Internationale Raumstation (ISS) mit der aufgrund der Bewegung verzogenen Lidar-Punktwolke (oben links) und einer Verschmutzungsdetektion für Kamera-basierte Fahrerassistenzsysteme (oben rechts). Schlussendlich liefert Abbildung 5 (unten) entsprechende Ergebnisse, die die aus der Raumfahrt stammenden und zur Absicherung von Fahrerassistenzsystemen transferierten Sensormodelle nun auch für die Produktion und Logistik präsentieren.

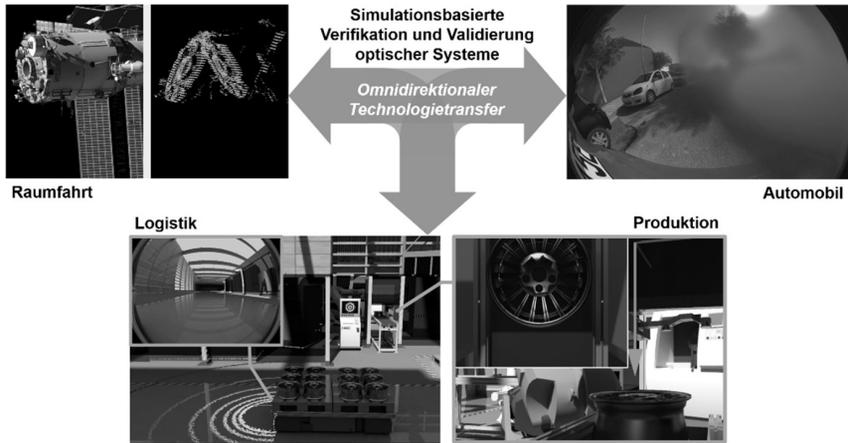


Abbildung 5: Verifikation und Validierung optischer Systeme und omnidirektionaler Technologietransfer zwischen den unterschiedlichen Anwendungsbereichen.

Dazu zeigt Abbildung 5 (unten rechts) ein simuliertes Inline-Prüfsystem (IPS), welches mittels Kamera etwaige Produktionsfehler mit Blick auf die Maßhaltigkeit der produzierten Felge prüft (simuliertes Kamerabild visualisiert als Bild im Bild). Die Simulation bietet hier die Möglichkeit unterschiedliche Hard- und Software-Konfigurationen (z. B. Sichtfeld und Verzerrung des Linsensystems, Auflösung und Integrationszeit des Detektors, Bildverarbeitungsalgorithmen und Neuronale Netze) systematisch zu evaluieren und schließlich das entstehende Gesamtsystem hinsichtlich Prüfmittelfähigkeit effizient und umfangreich zu validieren.

Ähnliches gilt für die simulationsbasierte Validierung des unten links abgebildeten Fahrerlosen Transportsystems (FTS), welches basierend auf den Sensordaten einer Weitwinkelkamera (visualisiert als Bild im Bild) und eines Multi-Layer Lidar-Sensors (visualisiert als farblich codierte 3D-Punktwolke) durch Produktionshallen navigiert und sich dabei an das dynamische Umfeld adaptiert (d. h. Hindernisse detektiert und umfährt). Hier erlaubt die Simulation eine effiziente Variation verschiedenster Szenarien (bestehend aus statischen und dynamischen Hindernissen), um eine funktionale Absicherung des entstehenden Gesamtsystems zu realisieren und so dessen Sicherheit im späteren Einsatz durch eine hohe Testabdeckung zu garantieren.

In allen Anwendungsfällen erlaubt die Simulation eine effiziente und systematische Verifikation und Validierung des entstehenden Gesamtsystems durch zielgerichtete Variation der System-Konfiguration und/oder des System-Umfelds. Die zur Beurteilung herangezogenen Metriken richten sich dabei stets nach der Funktion des betrachteten Systems (z. B. Genauigkeit des IPS und Kollisionsfreiheit des FTS). Unter Nutzung der durch die Simulationsdatenbank verfügbaren Ground-Truth Informationen (wie Geometrieformen, Materialeigenschaften und Geschwindigkeiten der Objekte) lässt sich ein jeweils passendes Fehlermaß berechnen, welches eine quantitative Beurteilung des Systemverhaltens erlaubt und so einen wesentlichen Mehrwehrt zu den sonst üblichen Tests in der realen Welt bietet.

Alle abgebildeten Anwendungen wurden gemäß der Dreiachsigen Skalierbarkeit durch wiederverwendbare Basiselemente in der gleichen Simulationsumgebung

realisiert und dazu alle oben skizzierten physikalischen Effekte modelliert. Zur Beurteilung und Sicherstellung einer akzeptablen Simulationsqualität/Realitätsnähe ist eine kontinuierliche Validierung der (Sensor-) Simulation selbst unerlässlich. Da die Anforderungen an die Simulationstiefe und der von ihr abgebildeten physikalischen Effekte stark von der betrachteten Anwendung (d. h. dem realen System) abhängt, bietet sich hier die in (Thieling und Roßmann 2021) vorgestellte Methodik an. Um ein angemessenes Verhältnis von Rechenzeit und Nutzen zu erzielen, werden Simulationskomponenten dabei so ausgelegt, dass diese so detailliert wie nötig, aber so trivial wie möglich sind. Zur Bestimmung der „richtigen“ Simulationstiefe liefert der zuvor referenzierte Beitrag ein geeignetes Verfahren.

6 Zusammenfassung

In diesem Beitrag haben wir eine skalierbare, physikalische Sensorsimulation vorgestellt, die eine effiziente, simulationsbasierte Verifikation und Validierung optischer Systeme in ihrem geplanten Einsatzumfeld erlaubt. Dazu wurden sensortypen- und anwendungsübergreifende Konzepte und Modelle vorgestellt, die die Erzeugung belastbarer, physikalischer Simulationsergebnisse erlauben.

Diese adressieren sowohl die extrinsischen Einflüsse einer Sensor-Umfeld Interaktion einschließlich Mehrwegeausbreitung, materialspezifischer Streuung und dynamischer Abläufe, als auch die intrinsischen Einflüsse des Sensorsystems mit der behandelten Linsenverzerrung, Tiefenschärfe und Bewegungsunschärfe. Auch wurde die Modellierung und Simulation oberflächlicher Verschmutzungen der Objekte und des Linsensystems (abgebildet als Materialien) präsentiert, die zur Simulation von Imperfektionen eines realen Einsatzumfeldes dienen.

Schlussendlich erlauben die hier vorgestellten wiederverwendbaren Basiselemente eine flexible und modulare Modellierung zur Erzeugung der einleitend vorgestellten Digitalen Zwillinge in Hybriden Testbeds und ermöglichen so den nahtlosen Transfer von einem ersten digitalen Prototyp zum finalen realen System – wie in (Thieling und Roßmann 2020) vorgestellt wurde.

Danksagung

Diese Arbeit ist Teil des Projektes ViTOS-II, gefördert vom Deutschen Zentrum für Luft- und Raumfahrt (DLR) mit Mitteln des Bundesministeriums für Wirtschaft und Technologie (BMWi), Förderkennzeichen 50 RA 1810.

Literatur

- Brown, D. C.: Close-range camera calibration. In: Photogrammetric Engineering Vol. 37, No. 8, 1971, S. 855–866.
- Carpin, S.; Lewis, M.; Wang, S.; Balakirsky, S.; Scrapper, C.: USARSim – a robot simulator for research and education. In: Proceedings 2007 IEEE International Conference on Robotics and Automation. IEEE, April 2002, S. 1400-1405.
- Crytek: CRYENGINE. <https://www.cryengine.com/>, letzter Zugriff: 08.04.2021.
- Cyberbotics: Webots. <https://www.cyberbotics.com/>, letzter Zugriff: 08.04.2021.

- Douchin, N.; Latger, J.; Cathala, T.; Marechal, R.: Simulating complex environments for the assessment of millimeter waves sensors. In: 2016 European Radar Conference (EuRAD), 2016, S. 13-16.
- Elmadawi, K.; Abdelrazek, M.; Elsobky, M.; Eraqi, H. M.; Zahran, M.: End-to-end sensor modeling for LiDAR Point Cloud. In: 2019 IEEE Intelligent Transportation Systems Conference (ITSC). IEEE, Oktober 2019, S. 1619-1624.
- Epic: Unreal Engine. <https://www.unrealengine.com/>, letzter Zugriff: 08.04.2021.
- Hanke, T.; Hirsenkorn, N.; Dehlink, B.; Rauch, A.; Rasshofer, R.; Biebl, E.: Generic architecture for simulation of ADAS sensors. In: 2015 16th International Radar Symposium (IRS), 2015, S. 125-130.
- Hirsenkorn, N.; Subkowski, P.; Hanke, T.; Schaermann, A.; Rauch, A.; Rasshofer, R.; Biebl, E.: A ray launching approach for modeling an FMCW radar system. In: 2017 18th International Radar Symposium (IRS), 2017, S. 1-10.
- Kannala, J.; Brandt, S.: A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 8, August 2006, S. 1335–1340.
- Kolchav, D. A.; Muratov, Y. R.; Nikiforov, M. B.; Orlov, S. V.: Real-time laser range-finder operation modeling. In: 2017 6th Mediterranean Conference on Embedded Computing (MECO), IEEE, Juni 2017, S. 1-4.
- Liu, D.: Ray-Tracing Based Interactive Camera Simulation. In: MVA2013 IAPR International Conference on Machine Vision Applications, 2013, S. 383-386.
- Mohammadikaji, M.; Bergmann, S.; Beyerer, J.; Burke, J.; Dachsbacher, C.: Sensor-Realistic Simulations for Evaluation and Planning of Optical Measurement Systems with an Application to Laser Triangulation. In: IEEE Sensors Journal, Vol. 20, No. 10, Mai 2020, S. 5336-6349.
- Open Source Robotics Foundation: Gazebo – Robot Simulation. <http://gazebosim.org/>, letzter Zugriff 08.04.2021.
- Stolz, M.; Nestlinger, G.: Fast generic sensor models for testing highly automated vehicles in simulation. In: e&i Elektrotechnik und Informationstechnik, Vol. 135, No. 4-5, August 2018, S. 365-369.
- Thieling, J.; Roßmann, J.: Scalable Sensor Models and Simulation Methods for Seamless Transitions within System Development: From First Digital Prototype to Final Real System. In: IEEE Systems Journal, 2020, S. 1-10.
- Thieling, J.; Roßmann, J.: Modulare Validierung simulierter Sensorsysteme für interagierende Digitale Zwillinge. In: Making Connected Mobility Work. Springer, 2021, S. 223-240.
- Unity: 3D development. <https://www.unity3d.com/>, letzter Zugriff: 08.04.2021.
- Velazquez, J. M. R.; Mailly, F.; Nouet, P.: A generic model for sensor simulation at system level. In: 2018 Symposium on Design, Test, Integration & Packaging of MEMS and MOEMS (DTIP), 2018, S. 1-4.





*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Entwicklung und Erprobung eines Vorgehens zur Validierung von Simulationsbausteinen komplexer Werkzeugmaschinen

Development and Testing of a Procedure for Validation of Simulation Building Blocks of Complex Machine Tools

Greta Tjaden, Carina Mieth, TRUMPF Werkzeugmaschinen GmbH & Co. KG,
Ditzingen (Germany), greta.tjaden@trumpf.com, carina.mieth@trumpf.com

Abstract: Verification and Validation are crucial for gaining sufficient confidence to transfer results from experiments within a simulation environment to real production systems. This paper shows the development and testing of a procedure for the validation of simulation building blocks, which represent machine tools for sheet manufacturing. This procedure has been developed based on a literature research for verification and validation methods considering surveys and applications. The developed validation procedure is applied on a simulation model consisting of several building blocks, and the approach and results are illustrated. Lastly, the limitations of the validation results are explained in detail.

1 Einleitung

Die stetig steigenden Anforderungen an die Produktivität von Werkzeugmaschinen und Fabriken führen zu immer komplexeren Abläufen innerhalb der Maschine und bei deren Verkettung. Simulation als wichtiger Bestandteil der Digitalen Transformation kann helfen, diese Komplexität besser beherrschbar zu machen. Dazu werden Produktionsanlagen und Prozesse virtuell analysiert, sowie Produktionsoptimierungen ohne Unterbrechung der laufenden Produktion im Simulationsmodell vorab getestet (Verein Deutscher Ingenieure (VDI), 2014, S. 4-6). Risiken und Fehler können so bereits in der Planung identifiziert werden. Voraussetzung hierfür ist die erfolgreiche Verifikation und Validierung (V&V) des Simulationsmodells.

1.1 Problemstellung

Modellbibliotheken haben sich als geeignet erwiesen, um modulare wiederverwendbare Bausteine bereitzustellen, mit denen man schnell neue Simulationen bauen kann. Bevor die Bausteine zur Modellbibliothek hinzugefügt werden, durchlaufen sie einen Prozess zur Verifikation. In diesem wird die interne

Entscheidungslogik der Maschine mit qualitativen V&V-Methoden getestet. Um die Bausteine tatsächlich für aussagekräftige Simulation nutzen zu können, müssen sie jedoch auch validiert werden. Validierung wird als Bewertung der Genauigkeit des Modells durch den Vergleich des simulierten Verhaltens mit dem Verhalten des realen Systems beschrieben (Balci, 1998, S. 41; Moreno-Díaz, Pichler, & Quesada Arencibia, 2009, S. 53). Kurz gesagt wird die Genauigkeit des Simulationsmodells überprüft (Detering & Schnieder, 2011, S. 5). Da in der Literatur kein passendes Verfahren zur Validierung der vorliegenden Simulationsbausteine identifiziert werden konnte, muss ein generalistisches und adaptierbares Validierungsvorgehen entwickelt werden, welche die zuvor erfolgte qualitative Verifikation der Simulationsbausteine optimal ergänzt.

Für dieses Validierungsvorgehen wird ein datenbasierter Abgleich von Simulation und Realität angestrebt. Hierfür müssen Daten erhoben werden, die anschließend aufbereitet und analysiert werden müssen. Werden diese Daten dann in das Simulationsmodell eingespeist, kann das Verhalten des Modells, mit dem des realen Systems verglichen werden. Eine abstrakte Veranschaulichung der Beziehung zwischen Prozess- und Simulationsmodell ist in Abbildung 1 erkennbar. Im Simulationsmodell gibt es an der Oberfläche Ein- und Ausgänge für die Aufträge (Jobs) und Blechtafeln (Sheets). Der Materialfluss, der zwischen diesen Ein- und Ausgängen nachgestellt werden muss, ist daneben zu sehen. Dabei muss auch die Kompatibilität der Maschinen mit den Automatisierungseinheiten berücksichtigt werden.

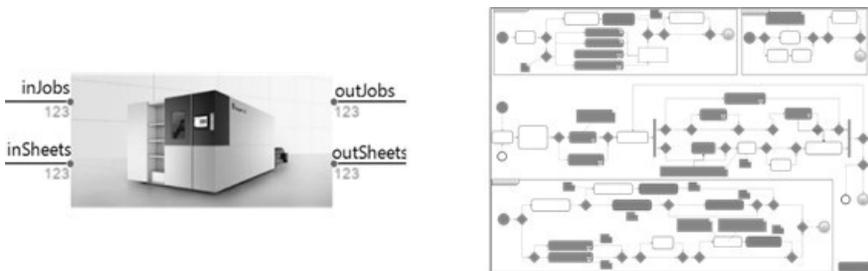


Abbildung 1: Simulationsbaustein einer komplexen Werkzeugmaschine (linkes Bild) mit Informations- (inJobs/outJobs) und Materialfluss (inSheets/outSheets) und beispielhaft ein Prozessmodell als UML Aktivitätsdiagramms (rechtes Bild)

1.2 Wissenschaftliche Abgrenzung

Zuerst wurde eine Literaturrecherche zur Identifikation der wesentlichen Methoden zur V&V durchgeführt. Die Ergebnisse dieser Literaturrecherche sind in Tabelle 1 zusammengefasst. Die Veröffentlichungen sind in Übersicht und Anwendung kategorisiert sowie eine Einordnung in qualitative und quantitative Methoden wurde hinzugefügt. Die Veröffentlichungen unter *Übersicht* nennen vor allem V&V-Methoden und erklären diese. Die Veröffentlichungen unter *Anwendung* erwähnen deutlich weniger V&V-Methoden.

Hintergrund, konnte in der Literatur nicht identifiziert werden. Auch bei TRUMPF gab es bisher noch kein Vorgehen für die Validierung von Simulationsmodellen.

Um in die Modellbibliothek aufgenommen zu werden, muss jeder Simulationsbaustein erfolgreich verifiziert werden. Bestandteil dieses Verifikationsprozess sind nicht nur Methoden wie z. B. Operational Graphics oder Desk-Checking, sondern auch die Erfüllung der „Definition of Done“-Kriterien. Diese definieren beispielsweise Anforderungen an die Dokumentation. Erst wenn der Verifikationsprozesses abgeschlossen ist, kann die Validierung erfolgen. Weil bei der Verifikation eher qualitative Methoden verwendet werden, wird in der Validierung auf quantitative Methoden gesetzt. Diese sind in Abbildung 2 hervorgehoben. Die entwickelte Methode nutzt Prozesssignale der Werkzeugmaschine zur Validierung von Simulationsbausteinen. Sie wird zwar am Anwendungsfall von TRUMPF erprobt, ist jedoch aufgrund der hohen Komplexität der einzelnen Simulationsbausteine generisch gehalten. So bietet sie aufgrund der daraus resultierenden allgemeinen Anwendungsmöglichkeiten auch Mehrwert für den wissenschaftlichen Kontext und ist übertragbar auf andere Werkzeugmaschinen. Besonders interessant wird dieser Ansatz auch bei fortschreitender Standardisierung von Maschinensignalen, z. B. über Standards wie umati (umati, 2018). Die Validierungsmethoden könnten dann kontinuierlicher und automatischer ausgeführt werden und benötigen keine aufwendigen Anpassungen an herstellerepezifische Datenmodelle.

1.3 Vorgehen und Aufbau des Beitrags

Das Vorgehen des Beitrags ist in Abbildung 2 visualisiert. Zuerst werden Methoden zur V&V in der Literatur identifiziert. Dabei werden sowohl Übersichten über Methoden als auch Anwendungen berücksichtigt. Anschließend wird aus diesen Methoden das Validierungsvorgehen entwickelt, wobei der Fokus auf den quantitativen Methoden liegt. Das Validierungsvorgehen wird mit Simulationsexpert*innen und in der Literatur identifizierten Qualitätskriterien für eine gute Validierungsmethode evaluiert. Nachdem das Verfahren auf ein erstes Simulationsmodell angewendet worden ist, werden die Einschränkungen der Ergebnisse erläutert.

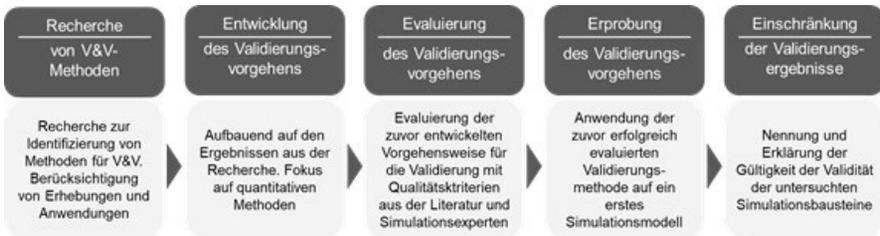


Abbildung 2: Vorgehen des Beitrags (V&V: Verifikation und Validierung)

2 Entwicklung der Validierungsmethode

Nach der zuvor beschriebenen Literaturrecherche wurde eine Auswahl der recherchierten Methoden getroffen. Da die zuvor erfolgte Verifikation der

Simulationsbausteine eher qualitativer Natur ist, wurde der Fokus vor allem auf quantitative Methoden gelegt. Zudem wurden Methoden berücksichtigt, die die Sensitivität von Parametern berücksichtigen. Die so entwickelte Methode wurde von Simulations-Expert*innen und in der Literatur identifizierten Qualitätskriterien für Validierungsmethoden evaluiert. Die identifizierten Kriterien sowie eine Beurteilung deren Erfüllung von der Validierungsmethode ist in Tabelle 2 dargestellt. Die Validierungsmethode soll sicherstellen, dass die stochastische Natur des Prozesses berücksichtigt wird, das Vorgehen sowie die Ergebnisse nachvollziehbar dokumentiert werden, eine kontinuierliche V&V erfolgt, und der betrachtete Simulationsbaustein valide ist. Letzteres bedeutet, dass alle relevanten Aspekte des Bausteins untersucht werden müssen. Außerdem muss die Validierungsmethode durchführbar sein und sollte unabhängig von den Erstellenden der Simulationsbausteine sein, um eine hohe Qualität der Ergebnisse zu gewährleisten.

Tabelle 2: Qualitätskriterien für Validierungsmethoden nach (Rabe et. al. 2008)

Kriterium	Grad der Erfüllung	
Stochastik	Voll	Die Stochastik wird in der Datenerfassung und -analyse berücksichtigt. Die Sicherstellung der Berücksichtigung von Stochastik in den Simulationsergebnissen liegt in der Verantwortung der Validierenden.
Dokumentation	Voll	Ein Beispiel für die Dokumentation wurde mittels GitExtension und Azure angelegt.
Kontinuität	Teilweise	V&V wird in allen Phasen der Bausteinerstellung durchgeführt, aber es wurde keine Methode etabliert, die eine kontinuierliche und automatische Validierung über die komplette Nutzungsdauer des Bausteins sicherstellt.
Validität	Voll	Eine Reihe von Validierungsmethoden ist enthalten, und die Berücksichtigung aller relevanten Aspekte wird bei der Validierung abgefragt.
Durchführbarkeit	Voll	Die Validierungsmethode wurde beispielhaft umgesetzt, wodurch die Durchführbarkeit bewiesen wurde.
Unabhängigkeit der Prüfenden	Voll	Die Prüfenden erstellen keine Simulationsbausteine. Somit ist die Unabhängigkeit gegeben.

In Abbildung 3 ist das entwickelte mehrstufige Validierungsvorgehen abgebildet. Zuerst werden Prozess- und Simulationsmodell verglichen, um die relevanten Prozesszeiten zu identifizieren. Anschließend wird eine Datenerhebung an der Maschine durchgeführt, um die Datengrundlage für die Validierung zu erhalten. In dieser Datenerhebung müssen ausreichend viele Werte für die Parameter erhoben werden, um mögliche statistische Abweichungen innerhalb des Produktionsprozesses abzubilden. Um die Datengrundlage für die jeweiligen Parameter zu erhöhen, werden die Werte aus unterschiedlichen Experimenten zusammengefasst, die als unabhängig voneinander betrachtet werden.

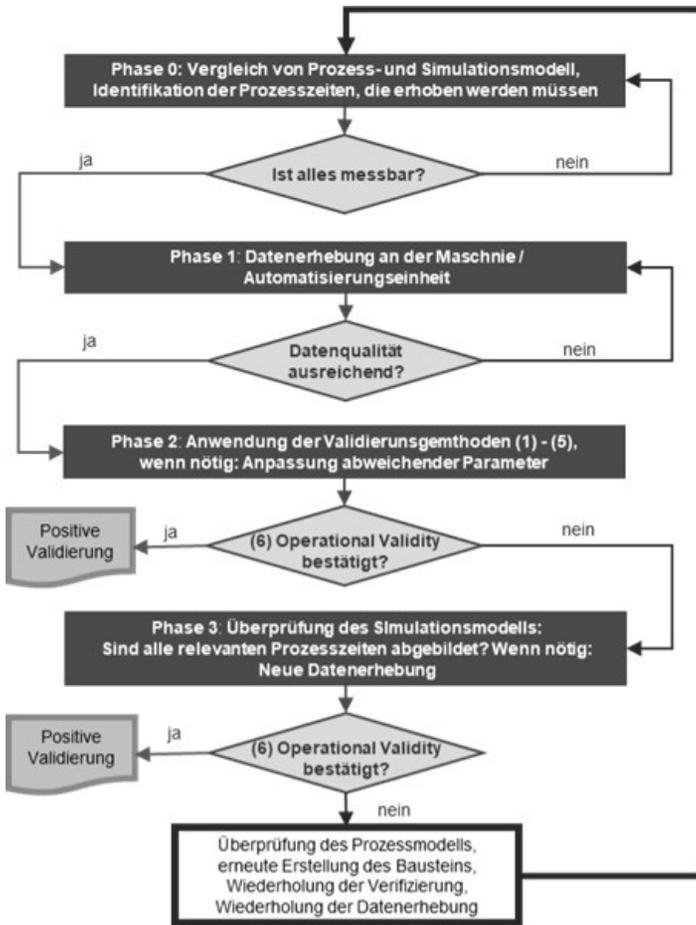


Abbildung 3: Das entwickelte Validierungsvorgehen.

(1) **Event Validity:** Ereignisse im Simulationsmodell werden mit denen der Realität verglichen. (2) **Historical Information Validation:** Vergleich mit historischen Daten, falls vorhanden. (3) **Predictive Validation:** Vorhersage des Systemverhaltens mit dem Simulationsmodell und anschließender Vergleich des realen Systemverhaltens mit der gemachten Vorhersage. (4) **Parameter Variability Sensitivity Analysis:** Änderung der Werte der Eingangs- und internen Parameter eines Modells, um die Auswirkungen auf das Verhalten oder die Ausgabe des Modells zu bestimmen. Im Modell sollten die gleichen Beziehungen auftreten wie im realen System (5) **Sensitivity Test:** Auswertung des Einflusses der Eingangsparameter auf die Ausgangsparameter mit besonderem Augenmerk auf die sogenannten sensiblen Parameter, bei denen schon kleine Änderungen große Auswirkungen auf die Ausgangsparameter haben. (6) **Operational Validity: (3 %)** Bestimmung, ob das Ausgangsverhalten des Simulationsmodells die für den beabsichtigten Zweck des Modells erforderliche Genauigkeit über den Bereich der

beabsichtigten Anwendbarkeit des Modells aufweist

Wenn die Qualität dieser Daten für die Validierung ausreicht, werden in Phase 2 die Validierungsmethoden durchgeführt. Für das entwickelte Validierungsvorgehen wurden die Methoden Event Validity, Historical Information Validation, Predictive Validation, Parameter Variability Sensitivity Analysis und Sensitivity Test ausgewählt, da diese Methoden datenbasiert sind und auch die Sensitivität der Parameter untersuchen. Die konkrete Auswahl aus diesen Methoden für das jeweilige Validierungsvorhaben ist abhängig von der zur Verfügung stehenden Datengrundlage und der Notwendigkeit weiterführender Untersuchungen der einzelnen Prozesszeiten. So sollen unterschiedliche Szenarien und falls nötig, mögliche Untersuchungen der Parameter berücksichtigt werden.

Eine interne Umfrage wurde zur vorläufigen Bestimmung der geforderten Genauigkeit für die unterschiedlichen Anwendungsmöglichkeiten des Fabriksimulationstools durchgeführt. Mittels Validierung soll nun bestimmt werden, ob die geforderte Genauigkeit erreicht wird. Dabei wird angestrebt, dass jeder Baustein für jeden Use Case eingesetzt werden kann und daher auch die höchste geforderte Genauigkeit von 3 % Abweichung zwischen Simulationsmodell und Realität bezogen auf die Zykluszeit und der Summe der Abweichungen der einzelnen Prozesse erreicht.

3 Anwendung der Validierungsmethode

Die Datenerhebung soll an zwei identischen Maschinen durchgeführt werden, um die Reproduzierbarkeit zu gewährleisten und systematische Fehler auszuschließen. Stimmen die erhobenen Daten von beiden Automatisierungseinheiten überein, so ist davon auszugehen, dass die erhobenen Daten von einer Automatisierungseinheit auf die andere übertragen werden können.

Um erste Aussagen über die Parametrierbarkeit der Bausteine treffen zu können, soll zudem das gleiche Experiment an Automatisierungseinheiten mit unterschiedlicher Maschinenkonfiguration durchgeführt werden. Dies bedeutet, dass die gleichen Automatisierungseinheiten mit einer größeren Maschine kombiniert werden können.

Nach der Durchführung der ausgewählten Validierungsmethoden wird die Operationale Validität berechnet, also die Genauigkeit des Simulationsmodells im Hinblick auf das reale System. Liegt diese im zuvor für die Simulationsbausteine definierten Rahmen, liegt ein erfolgreiches Validierungsergebnis vor. Überschreitet die berechnete Operationale Validität den definierten Grenzwert, wird Phase 3 des Validierungsvorgehens durchgeführt. Hier wird für das Simulationsmodell überprüft, ob alle relevanten Prozesszeiten abgebildet sind. Wenn nötig, wird eine erneute Datenerhebung durchgeführt. Anschließend wird die Validierung analog zu Phase 2 durchgeführt und die Operationale Validität erneut berechnet. Kann der Grenzwert nun eingehalten werden, wird die Validierung als erfolgreich betrachtet. Wird der Grenzwert trotz der Anpassungen weiterhin überschritten, so muss das Prozessmodell überprüft, der Simulationsbaustein neu entwickelt, die Verifikation und ggf. auch die Datenerhebung wiederholt werden.

Das Experiment für die Erhebung der Validierungsdaten wird an einer 2D-Laserschneidmaschine durchgeführt, die mit zwei Automatisierungskomponenten

zum Be- und Entladen zusammengeschlossen ist. In vorherigen Experimenten haben sich folgende Einstellungen bewährt:

- Schneiden mit geringer Laserleistung: Bei sehr geringer Laserleistung wird der Schneidprozess realitätsnah durchgeführt, ohne dass das Blech tatsächlich geschnitten wird. Dieses Vorgehen spart Material und ist damit nachhaltig und günstig.
- Tafelbelegung mit möglichst geringer Schneidzeit: Der zeitliche Aufwand der Experimente wird gering gehalten. Zudem lassen sich so einfacher die Wartezeiten der Automatisierungskomponenten ermitteln.
- Kleine und leichte Tafeln: Es werden Bleche genutzt, die eine Person manuell transportieren kann, um auch hier den Aufwand so gering wie möglich zu halten.
- Die maximal verfügbare Anzahl gleicher Bleche für die Experimente verwenden, die obige Anforderungen erfüllen (in diesem Fall 11 Bleche).
- Wiederholung der Experimente mit und ohne Nebenzeiten, wie z. B. der Blechlagenvermessung oder Düsenwechsel.

Um die Zeiten so genau wie möglich und ohne menschlichen Einfluss zu erheben, werden Messpunkte in das NC-Programm gesetzt und zusätzlich die verfügbaren Signale aus der Maschinensteuerung aufgezeichnet. Anhand dieser aufgezeichneten Daten werden die Parameter für die Validierung berechnet, z. B. die Schneidzeit, die Zeit für den Düsenwechsel oder die Blechlagenvermessung. Diese Zeiten dienen als Eingangparameter für das Simulationsmodell, in dem das Experiment für die Datenerhebung nachgestellt wird. Anschließend werden die Zeiten des Simulationsmodells und des realen Systems miteinander verglichen (s. Abb. 4).

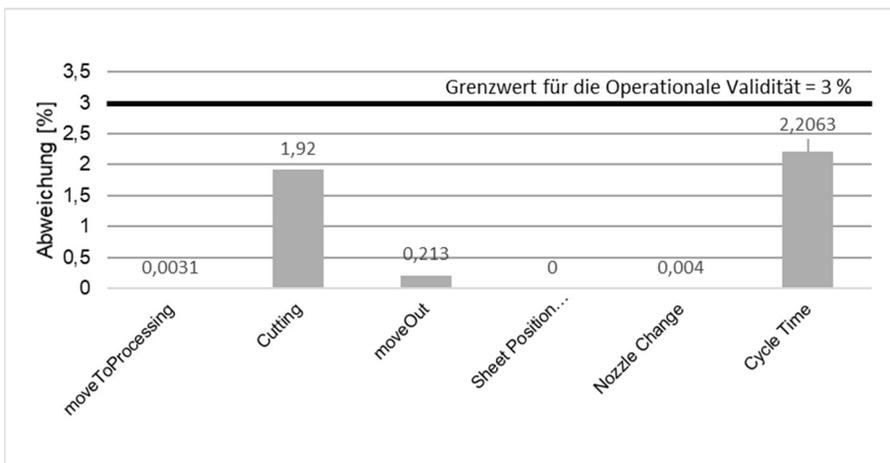


Abbildung 4: Vergleich der ermittelten Abweichungen für die Validierungsparameter zwischen Simulationsmodell und realem System (TruLaser 1030 fiber)

Es ist auf den ersten Blick erkennbar, dass keiner der Validierungsparameter den Grenzwert für die Operationale Validität von 3 % überschreitet. Der Simulationsbaustein kann dennoch nicht als positiv validiert eingestuft werden, weil

die Datengrundlage mit teilweise zwei ermittelten Werten noch zu klein ist, um statistische Sicherheit zu geben. Die Simulationsbausteine der Automatisierungskomponenten können nicht validiert werden, weil die Signale für die Datenerhebungen nicht zur Verfügung standen.

4 Einschränkungen der Ergebnisse

Es wurde gezeigt, dass ein Vorgehen zur Validierung der Simulationsbausteine entwickelt und erfolgreich angewendet werden konnte. Es gibt jedoch Fragen, die in diesen Zusammenhang noch nicht geklärt wurden.

Es wurde bisher nicht untersucht, ob die Simulationsbausteine nach der Validierung in allen Kombinationen anwendbar sind, oder ob jede Kombination aus Automatisierungskomponente und Maschine einzeln validiert werden muss. Daher können keine Aussagen über das Verhalten in anderen Kombinationen gemacht werden. Zudem ist die Parametrisierung der einzelnen Bausteine noch nicht abschließend geklärt. Fest steht, dass die Simulationsbausteine parametrierbar sind, dass also ein Simulationsbaustein unterschiedliche Größen eines Maschinentyps abbilden kann. Die Parametrisierung der einzelnen Größen muss jedoch noch erfolgen.

Um die Reproduzierbarkeit des Validierungsvorgehens zu gewährleisten, muss die Messung noch an einer identischen Kombination von Automatisierungskomponente und Maschine wiederholt werden. Aufgrund eines Softwarefehlers, der eingespielt wurde, um einen Kunden-Fehlerfall zu reproduzieren, danach jedoch nicht sofort entfernt wurde, ist die damals durchgeführte Messung an einer identischen Maschinenkombination nicht verwendbar gewesen. Dass der Softwarefehler jedoch durch die Validierung entdeckt wurde, zeigt, dass nicht nur Simulationsmodelle, sondern auch Maschinen mit der Methode validiert werden können.

Zur Ermittlung des Grenzwerts der Operationalen Validität der Simulationsmodelle wurde eine interne, vorläufige Umfrage durchgeführt, die jedoch nicht repräsentativ für den Maschinenbau ist. Außerdem muss verifiziert bzw. falsifiziert werden, ob die geforderte Genauigkeit der Simulationsmodelle für die Bausteine übernommen werden kann, oder ob die Bausteine selbst eine andere Genauigkeit erreichen müssen, damit die aus ihnen erstellten Simulationsmodelle den festgelegten Grenzwert der Operationalen Validität nicht überschreiten.

Einige Validierungsparameter wie z. B. die Blechlagenmessung sind abhängig von der Blechgröße. Dieser Einfluss ist noch nicht untersucht worden und muss zu einem späteren Zeitpunkt nachgeholt werden. Die Analyse der verschiedenen Maschinengrößen ist noch nicht abgeschlossen, da nur zwei Maschinengrößen analysiert wurden.

Schließlich konnten im Rahmen dieser Arbeit aufgrund der Tafelverfügbarkeit nicht mindestens 30 Durchläufe (Meran et. al., 2013. S. 94) pro Experiment durchgeführt werden. Um die Datenbasis so weit wie möglich zu erweitern, wurden die Daten für die Parameter, die bei den einzelnen Versuchen ermittelt wurden, zusammengefasst. Dennoch ist nicht gewährleistet, dass die stochastische Natur der Produktionsprozesse hinreichend erfasst worden ist.

5 Fazit

Es wurde eine umfangreiche Literaturrecherche durchgeführt, um Methoden für die V&V von Simulationsmodellen und -bausteinen zu identifizieren. Dabei wurden sowohl Erhebungen als auch Anwendungen dieser Methoden berücksichtigt. Aufbauend auf diesen Ergebnissen wurde ein iteratives Vorgehen zur Validierung der Simulationsbausteine entwickelt. Dieses Vorgehen ist vorwiegend quantitativer Natur und berücksichtigt Validierungsmethoden unterschiedlicher Zeithorizonte sowie die Evaluierung von Parametern. Dabei werden sowohl der Abgleich von Prozess- und Simulationsmodell sowie die Planung, Durchführung und Auswertung der Datenerhebung berücksichtigt, bevor die eigentliche Validierung beginnt. Um eine hohe Qualität der Validierungsergebnisse zu gewährleisten, wurde die Methode von Simulationsexpert*innen und nach in der Literatur identifizierten Qualitätskriterien für V&V evaluiert. Die Durchführbarkeit der Validierungsmethode wurde an einem Beispiel demonstriert. Voraussetzung für das entwickelte Validierungsvorgehen ist die Verfügbarkeit von Daten. Die Bausteine, aus denen das Simulationsmodell besteht, konnten validiert werden.

In zukünftigen Arbeiten wird die Übertragbarkeit der Bausteine in andere Simulationsmodelle, also in anderen Kombinationen von Maschinen und Automatisierungskomponenten, überprüft. Sobald die entsprechenden Signale zu jeder Zeit im Feld verfügbar sind, soll die Validierung automatisch erfolgen. Dies würde die Verfügbarkeit der Daten erhöhen und die Kontinuität der Validierungsmethode verbessern. Zukünftig soll mithilfe von repräsentativen Produktionsdaten die Validierung von Simulationsmodellen vereinheitlicht werden. Diese Referenzdatensätze enthalten Produktionsdaten, -schritte, Auftragsstrukturen und weitere Kundenmerkmale.

Literatur

- Aboud, S. J., Al Fayoumi, M., & Alnuaimi, M. (2009). Verification and Validation of Simulation Models, 58–74.
- Al-Weshashi, M. A., Anderson, A., Tian, G., & Makhdoum, B. M. A. (2013). Validation of Simulation Model for Cogeneration Power and Water Desalination Plant. *International Journal of Modeling and Optimization*.
- Balci, O. (1998). Verification, Validation, and Accreditation.
- Cano-Moreno, J. D., & Cabanellas, J. M. (2019). Experimental Validation of an Escalator Simulation Model. *Latin American applied research* *Pesquisa aplicada latino americana = Investigación aplicada latinoamericana*.
- Detering, S., & Schnieder, L. (2011). Calibration and Validation of Simulation Models for Investigation of Traffic Assistance Systems, 217–225.
- Dozortsev, V. M., Kneller, D. V., & Levit, M. J. (2002). On the Validity of Simulation Models in Process Engineering and Operator Training.
- Haffke, B., Möller, R., Melz, T., & Strackeljan, J. (2015). Validation of Simulation Models without Knowledge of Parameters Using Differential Algebra. *Mathematical Problems in Engineering*, 2015, 1–9.
- Henderson, S. G., Biller, B., Hsieh, M.-H., Shortle, J., Tew, J. D., & Barton, R. R. (1970). Landmark Paper Reprise - A tutorial on verification and validation of simulation models, 33–39.

- Jeon, H.-H., Jung, Y.-J., Siddique, M. A. A., Nam, K.-C., Kim, T.-B., Choi, C.-H., & Kim, Y.-J. (2019). Development and Validation of simulation model for three point-hitch during agricultural operation.
- Kleijnen, J. P.C., Bettonvil, B. W.M., & van Groenendaal, W. J. H. (1996). Validation of Simulation Models: Regression Analysis Revisited. SSRN Electronic Journal.
- Martens, J., Put, F., & Kerre, E. E. (2003). A fuzzy set and resemblance relation approach to the validation of simulation models.
- Meran, R., John, A., Roenpage, O., Staudter, C., & Lunau, S. (Eds.) (2013). Six Sigma+Lean Toolset: Mindset for Successful Implementation of Improvement Projects. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Moreno-Díaz, R., Pichler, F., & Quesada Arencibia, A. (2009). Computer Aided Systems Theory -- EUROCAST 2009: 12th International Conference, Las Palmas de Gran Canaria, Spain, February 15-20, 2009 revised selected papers, 5717.
- Ni, D., Leonard, J. D., Guin, A., & Williams, B. M. (2004). Systematic Approach for Validating Traffic Simulation Models. Transportation Research Record: Journal of the Transportation Research Board. (1876), 20–31.
- Olsen, M., & Raunak, M. (2018). Increasing Validity of Simulation Models Through Metamorphic Testing. IEEE Transactions on Reliability, 68(1), 91–108.
- Pei, Y., Page, J., & Pearce, G. (2015). Validating Airframe Simulation Model of R22 Helicopter by Using Crashworthiness Standards.
- Rabe, M., Spieckermann, S., & Wenzel, S. (2008). A new procedure model for verification and validation in production and logistics simulation, 1717–1726. Retrieved September 11, 2020.
- Rabe, M., Spieckermann, S., & Wenzel, S. (2008). Verifikation und Validierung für die Simulation in Produktion und Logistik. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Rebba, R., Huang, S., Liu, Y., & Mahadevan, S. (2006). Statistical validation of simulation models. International Journal of Materials and Product Technology, 25(1/2/3), 164.
- Sargent, R. G. (1984). A tutorial on verification and validation of simulation models, 77–87.
- Sargent, R. G. (2011). Verification and validation of simulation models, 166–183.
- umati (2018). connecting the world of machinery. https://umati.org/wp-content/uploads/umati_Folder_HM21.pdf, letzter Zugriff 05.07.2021.
- Verein Deutscher Ingenieure (VDI) (2014). Simulation von Logistik-, Materialfluss- und Produktionssystemen Grundlagen. (Nr. 3633).





*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Simulation der Einplanung der terminkritischen Montage von Abfüllanlagen am Beispiel der Krones AG

Simulation of planning of time-critical assembly of bottling machines based on example of Krones AG

Frank Herrmann, Innovationszentrum für Produktionslogistik und Fabrikplanung,
Ostbayerische Technische Hochschule Regensburg, Regensburg (Germany),
frank.herrmann@oth-regensburg.de.

Faruk Savasci, Krones AG, Neutraubling (Germany), faruk.savasci@krones.com.

Abstract: Many companies need to make the best possible use of their production areas. For others, meeting predetermined customer due dates is critical. Few companies have to meet both requirements simultaneously. One example is the final assembly at Krones AG. An extensive simulation study identified potential for improvement. In some planning problems, preferring orders due to buffer time proves to be advantageous. In others, avoiding unused space is better. This project is to be continued with the development of a simulation-based planning program.

1 Einleitung

In einigen Unternehmen sind Aufträge mit sehr großen Platzbedarfen zu produzieren. Trotz sehr hohen Investitionen in geeignete Hallen stellt der verfügbare Platz einen wesentlichen Engpass dar. Ein Beispiel ist die in diesem Beitrag betrachtete Endmontage von Abfüllanlagen bei der Krones AG in Neutraubling, dem Weltmarktführer in dem Bereich der Getränkeindustrie.

Der generelle Ansatz, sicher auch bei anderen Unternehmen mit vergleichbaren technologischen Restriktionen, ist die Einplanung der Halle durch Planer. So wie bei der Krones AG auch, dürften Unternehmensleitungen zentrale Kennzahlen wie vor allem Durchlaufzeiten und Termineinhaltung als stark verbesserungswürdig ansehen. In den letzten Jahren waren, bei Krones, stets wenigstens 21% der Aufträge verspätet. Verspätungen führen zu signifikanten Konventionalstrafen. Außerdem unterminieren ihr Auftreten das Ziel von Krones, jede Anlage termingerecht in Betrieb zu setzen. Bei vielen Aufträgen wurden nur deswegen Verspätungen vermieden, da in den meistens langen Produktionsprozessen durch Maßnahmen wie Überstunden Prozessbeschleunigungen vorgenommen worden sind. Damit ist das Produktionsergebnis im Hinblick auf die Termineinhaltung nicht repräsentativ für die

Planungsgüte. Im Fall von Krones wurde festgestellt, dass ein Planer in erster Linie technologische Randbedingungen, wie vor allem die Platzbeschränkung berücksichtigt.

Krones wollte nun genauer wissen:

- Wie gut ist die Planungsgüte tatsächlich?
- Wie gut ist eine Planung, bei der (vor allem) nach Terminen gesteuert wird?

Die angesprochenen Abläufe verhindern eine Analyse der Planung. Daher sollten diese Fragen durch eine Simulationsstudie beantwortet werden. Eine geeignete wurde aufgesetzt und deren Ergebnisse werden im Folgenden vorgestellt. Diese werfen weitere Fragen auf, die in einer weiteren Untersuchung in naher Zukunft beantwortet werden sollen.

Dieser Artikel ist wie folgt strukturiert. Zunächst werden komprimiert einige Literaturquellen zur Simulation im Allgemeinen sowie zur Einplanung von Aufträgen in einer Werkstatt, ebenfalls im Allgemeinen, genannt. In den folgenden beiden Abschnitten wird die Produktion und das bisherige Vorgehen erläutert. Dem schließt sich die Darstellung der Simulationsuntersuchung an. Entscheidend ist das bei Krones entwickelte Simulationssystem. Die erzielten Ergebnisse werden danach vorgestellt. Das Projekt als Ganzes wird zusammengefasst und die geplanten weiteren Arbeiten werden genannt.

2 Literaturüberblick

Solche Unternehmen müssen folglich planen, wo eine Anlage montiert werden soll und in welcher Reihenfolge diese in die Werkshalle eingelastet werden. In der Literatur finden sich zahlreiche Arbeiten zur Reihenfolgeplanung wie in (Herrmann, 2011) oder (Pinedo, 2016) und zur Layout-Planung wie in (Drira, et al., 2007). Bei Kombinationen von beiden Problemen dominiert entweder die Reihenfolgeplanung oder die Layout-Planung. Ein Beispiel ist die Projektplanung bei Schiffen in (Ge & Wang, 2020), bei denen ein gutes Layout am wichtigsten ist. Bei der hier betrachteten Klasse von Problem handelt es sich nicht um Projekte, sondern um voneinander unabhängige Einzelmontageaufträge.

Im Hinblick auf die Literatur zur Simulation für diese Arbeit sei auf die Tagungsbände der alle zwei Jahre stattfindenden Fachtagung Simulation in Produktion und Logistik der Arbeitsgemeinschaft Simulation (ASIM) hingewiesen. Besonders relevant für diese Untersuchung waren (März, Krug, Rose, & Weigert, 2010) und die Beschreibung einer Ablaufsimulation in (Mayer, et al., 2020).

3 Montageprozess

Die Krones AG mit Hauptsitz in Neutraubling, Deutschland beschäftigt weltweit circa 15000 Mitarbeiter. Sie produziert Komponenten, Linien und Anlagen für Getränke und Liquid Food seit 1951, ist der Branche des Maschinen- und Anlagenbaus zugeordnet und ist führender Hersteller in der Verpackungs- und Abfülltechnik. Hierbei deckt das Unternehmen jeden Prozessschritt der Produktion ab, angefangen bei der Produkt- und Behälterherstellung über die Abfüllung und Verpackung hin zu Materialfluss und Recycling der Behälter. Eines der letzten Schritte bildet die Endmontage von Abfüllmaschinen. Hierfür nutzt Krones die Werkshalle 5 die

ausschnittsweise in Abbildung 1 dargestellt ist, und 5 Füller in der finalen Endmontageposition zeigt. Die nutzbare Fläche ist schematisch in der Abbildung 2 angegeben.



Abbildung 1: Werkshalle 5 Endmontage von Abfüllmaschinen mit aktuell zu montierenden Anlagen aus der Fülltechnik

Wegen den großen Abmessungen der hergestellten Maschinen mit Durchmessern von 1,0 m bis 7,2 m hat die Werkshalle 5 eine gesamte Länge von 105 m und eine Breite von 30 m, also eine Gesamtfläche von 3150 m². Wie in der Abbildung 2 dargestellt, ist die Halle durch einen Transportweg längs mittig mit 4 m und einem Transportweg quer außermittig von 3 m Breite geteilt, wodurch vier Einzelflächen zur Einplanung von Anlagen zur Verfügung stehen. Zwei Nutzflächen sind 1008 m² groß - mit 42 m Länge und 12 m Breite - und die anderen sind 1200 m² groß - mit 50 m Länge und 12 m Breite. Damit ergibt sich insgesamt eine für die Montage nutzbare Fläche von 2208 m². Die restlichen Abschnitte der Werkshalle werden für Bearbeitungsmaschinen reserviert und sind folglich für die Montage nicht nutzbar.

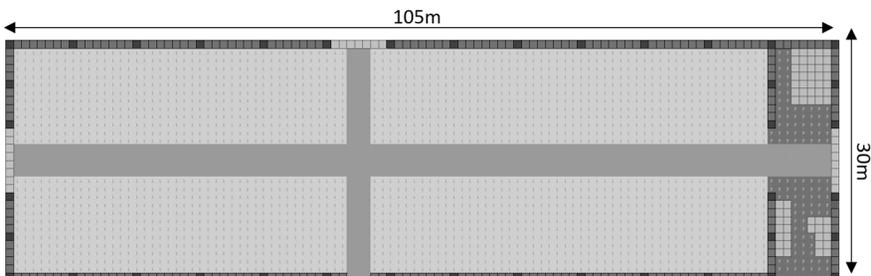


Abbildung 2: Werkshalle 5 in voller Komplexität

Die folgende Beschreibung der Montage diene zur Illustration der Anforderungen an den Platzbedarf, der im Zeitablauf zu und abnehmen kann. Zugleich zeigt es

technische Restriktionen auf, die belegen, dass die Bearbeitungszeiten nicht im Vorhinein bekannt und konstant sind.

Durch die Werkstore kann jeweils immer nur eine Anlage in die Werkshalle ein- bzw. austreten. Bevor mit der Endmontage einer Maschine begonnen wird, wird je nach Maschinengröße ein Endmontageplatz mit ausreichendem Platz festgelegt. Beispielsweise sind in der folgenden Abbildung 3 drei Füller zu montieren. Es handelt sich um ein Fabriklayoutsegment. Die Werkshalle 5 besteht aus fünf solchen Fabriklayoutsegmenten. Auf den Anlieferungsplätzen (i.e. „Anlieferungsspot“ in Abbildung 3) werden alle Unterbaugruppenteile und Materialien aller Maschinen angeliefert. Hier können Materialien von verschiedenen Maschinen und Baugruppen gemischt werden; durch „Baugruppe“ in Abbildung 3 gekennzeichnet.

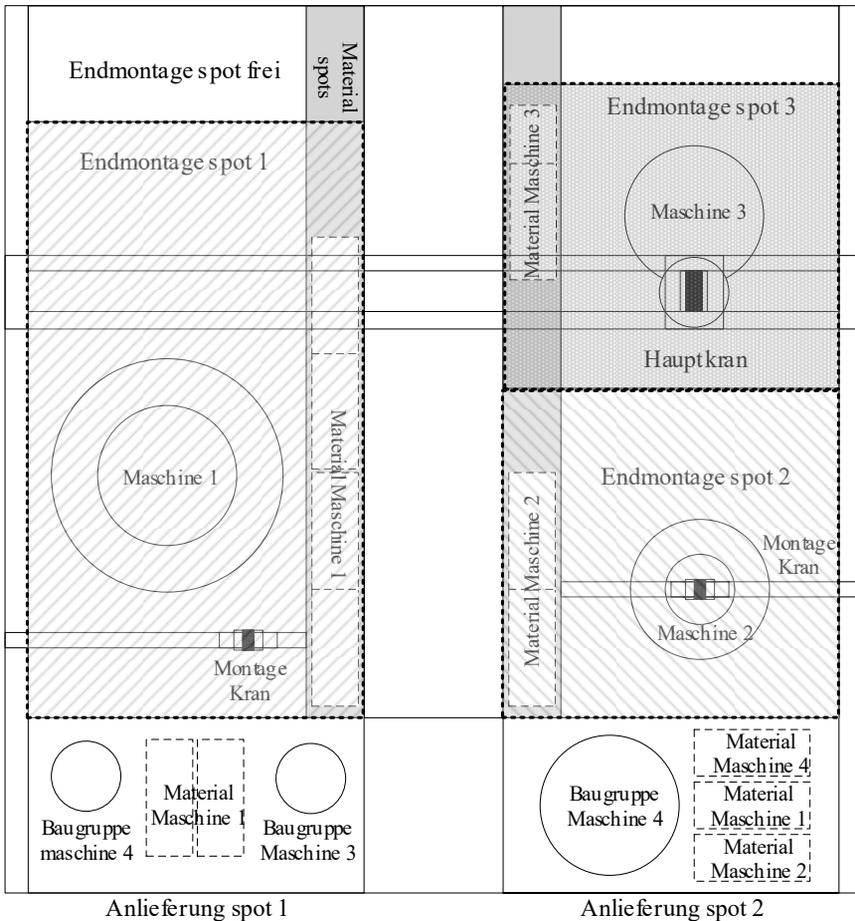


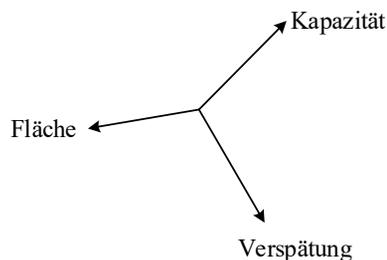
Abbildung 3: Ausschnitt Fabriklayout der Endmontage der Fülltechnik

Diese Außenanlieferungsplätze sind begrenzt, können aber von jedem Maschinen- und Materialtyp genutzt werden, bis die Platzkapazität der Plätze erreicht ist.

Die Unterbaugruppen und Materialien werden in zwei Gruppen von Stammmaterialtypen aufgeteilt. Die erste Gruppe von Baugruppen kann aufgrund von Konstruktion und / oder Gewicht nur mit dem Hauptkran in die Fabrikanlage gehoben werden (s. Abbildung 3 Materialien mit durchgezogenen Randlinien). Meistens handelt es sich bei diesen Teilen um Hauptmaschinenteile. Die zweite Gruppe von Materialien kann von Bedienern an ihren Ziel-Materialplatz gebracht werden (s. Abbildung 3 Materialien mit gestrichelten Umrandungslinien). Hier handelt es sich dabei um Montageteile, die in Wagen oder Behältern sortiert sind. Diese Teile können mit Hilfskränen an ihren endgültigen Platz in der Maschine gehoben werden und werden einer einzelnen Maschine im Innenmaterialplatz zugewiesen. Diese Indoor-Materialspots sind entsprechend dem Endmontageplatz räumlich fixiert und sind individuell für jede Maschine. Nach Beendigung des letzten Montagevorgangs kann die komplette Maschine nur noch mit dem Hauptkran aus der Werksanlage transportiert werden. Der Hauptkran und die Hilfskräne arbeiten auf unterschiedlichen Ebenen, so dass eine Blockierung nicht möglich ist. Hilfskräne befinden sich entweder auf der rechten oder auf der linken Seite der Werkshalle und können an jeder Seite nicht aneinander vorbeifahren.

4 Bisheriges Vorgehen durch einen Planer

Nach der jahrelangen Erfahrung eines Planers besteht ein Zielkonflikt zwischen der Minimierung der ungenutzten Fläche in der Werkshalle 5, der Minimierung der Anzahl an eingesetzten Mitarbeitern und der Minimierung der Verspätung. Dies führt zu dem nebenstehenden Trilemma der Ablaufplanung; für Weiteres sei auf (Herrmann, 2011) verwiesen.



In 2020 wurden 145 Maschinen montiert und in den nächsten Jahren ist mit einer eher höheren Anzahl zu rechnen. Die Bearbeitungsdauern der Montageaufträge schwanken zwischen 6 Wochen und 19 Wochen mit einem Mittelwert von 10.24 Wochen und einer Streuung von 2.5 Wochen. Wochen sind auch die generellen Betrachtungsperioden; zum Ende einer Woche verlassen die Anlagen die Halle. Geplant wird stets zu Beginn einer Woche. Der Arbeitsvorrat umfasst 12 Wochen.

Im Kern erfolgt die Planung bei Krones derzeit durch das folgende Vorgehen:

Zu Beginn jeder Woche kann ein Auftrag in die Werkshalle eingeplant werden, sofern die benötigte Hallenfläche, ohne Umstellung von bereits in der Halle sich befindenden Anlagen, zur Verfügung steht. Als Datengrundlage verwendet ein Planer ein ausgedrucktes Layout der Werkshalle, in dem die bisher eingeplanten Anlagen eingezeichnet sind. Die Anlagendaten wie Länge, Breite, Kundentermin befinden sich in dem SAP System von Krones. Ist ein geeigneter Platz für eine Anlage gefunden worden, so wird diese auf dem allgemein verfügbaren Layout der Werkshalle für die Montagedauer sichtbar markiert. Je nach Größe der Anlagen werden diese zuerst eingeplant gefolgt von Kleineren. Auch werden lange Montagedauern kürzeren gegenüber bevorzugt.

5 Simulationsstudie

Sowohl verschiedene Reihenfolgen der Belegung der Werkshalle 5 von einem der beiden Planer bei der Krones AG wie auch eine durch die Kürzeste Pufferzeitregel (KPZ-Regel) bestimmte Reihenfolge sollen simulativ analysiert werden. Ein nach der Literatur etabliertes Mittel für eine solche Analyse ist eine Ablaufsimulation, s. beispielsweise (Mayer, et al., 2020), für die die Werksleitung sich entschied. Konkret soll die Positionierung der Anlagen nachgestellt werden. Ferner werden – im simulierten Ablauf – die Anlagen für die Dauer der Montage eingeplant und verlassen dann die Werkshalle wieder. Das Projekt folgt dem Vorgehen in (Mayer, et al., 2020) und gliederte sich in die Phasen Analyse, Daten, Modellierung, Verifikation und Validierung (V&V) des Modells, Experimentplanung, -durchführung, -bewertung und Änderungen sowie Lösung aufbereiten.

Für die Simulationsstudie wurden zunächst gängige kommerziell verfügbare Simulationstools wie Plant Simulation von Siemens gegenüber einer Eigenentwicklung evaluiert. Die Tools erlauben umfangreiche Visualisierungen der Abläufe. Mit Ihnen lässt sich jedoch nicht der Entwicklungsaufwand zur Steuerung der Einplanung zufällig eintreffender Kundenaufträge in die Werkshalle gegenüber einer Eigenentwicklung sehr signifikant reduzieren. Außerdem sind hohe Investitionskosten für die Anschaffung eines solchen System, für Schulungsmaßnahmen usw. erforderlich.

Konkret wurde die Ablaufsimulation in Excel unter Nutzung der Programmiersprache VBA implementiert. Dabei repräsentiert ein Excel-Blatt eine Belegung der Werkshalle 5. Wie bereits in Abbildung 2 dargestellt wurde, repräsentiert jede dort angegebene Zelle eine tatsächliche Fläche über 1 m² in der Werkshalle. Eine 0 bedeutet ein ungenutzter Platz und eine 1 eine Belegung durch eine Anlage. Bestandteile wie Büroflächen, blockierte Flächen, Säulenflächen, Wandflächen, Türflächen, Torflächen, Fläche von Durchgangswegen oder Auftragsflächen lassen sich beispielsweise durch zweistellige Ziffern größer als 80 darstellen.

Eine Abfüllanlage mit allen im Zeitablauf (maximal) benötigten Flächen wird durch das kleinste Rechteck in einem Excel-Blatt dargestellt, welches die benötigte Fläche enthält; wie bereits genannt, werden die Länge und Breite einer Anlage aus dem SAP System von Krones entnommen. Eine Repräsentation einer konkreten Hallenbelegung ist in Abbildung 4 visualisiert; jede Graustufe steht für eine spezifische Anlage.

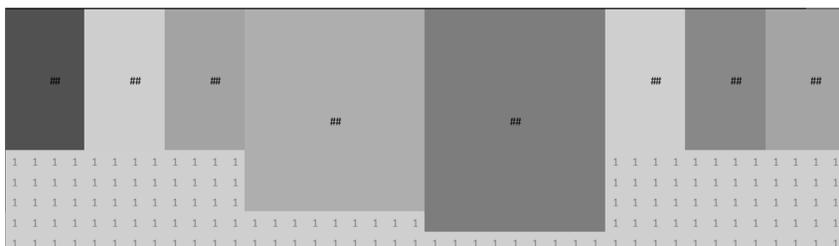


Abbildung 4: Abstraktion einer eingeplanten Werkshalle 5 mit farbigen Flächen für Anlagen.

Für die Simulation wird für jede Woche (Periode) des Simulationszeitraums ein solches Blatt erzeugt. Ausgangspunkt ist ein Workload für den kompletten Simulationszeitraum. Ein Beispiel enthält Abbildung 5. Zu jedem Auftrag ist seine Nummer, die Bearbeitungszeit, sein Produktname, die betroffene Halle (für eine Erweiterung der Simulation auf alle Hallen von Krones), seine Länge und Breite, eine etwaige Rotation sowie sein frühester möglicher Starttermin bzw. Freigabetermin (i.e. release date) und sein Soll-Endtermin (i.e. due date) angegeben.

Order ID [#]	Processing time [period]	Product [#]	Production hall [#]	Product width [unit]	Product length [unit]	Product rotation R09	Release date [period]	Due date [period]
1001	14	1	1	12	12	N	1	5
1002	15	2	1	13	12	N	1	6
1003	13	3	1	10	12	N	1	5
1004	12	4	1	7	9	N	1	4
1005	12	5	1	7	9	N	1	4
1006	13	6	1	11	12	N	1	5
1007	12	7	1	7	9	N	1	5
1008	10	8	1	8	11	N	1	3

Abbildung 5: Beispielhafter Workload

Gemäß einer Einplanungs-Reihenfolge, die von einem Planer oder durch die KPZ-Regel erzeugt wurde, werden die Anlagen bzw. Aufträge nacheinander wie folgt in die Werkshalle 5 eingelastet. Zu einem einzulastenden Auftrag wird die Produktionsfläche von links nach rechts und danach von oben nach unten solange durchsucht, bis eine hinreichend große freie Fläche gefunden wurde. Diese muss auch in den kommenden Wochen frei sein, da eine Einplanungs-Reihenfolge nicht zwingend nach den Freigabeterminen sortiert ist. Im positiven Fall erfolgt die Belegung, und zwar so weit links-oben wie möglich.

Row ID [#]	Order ID [#]	Worklist [sequence]	Processing time [period]	Product [#]	Production hall [#]	Product width [unit]	Product length [unit]	Product rotation R090	Release date [period]	Due date [period]	Starting date [period]	End date [period]	Hall distance length [unit]	Hall distance width [unit]	Simulation time [second]
1001	1001	1001	14	1	1	12	12	N	1	5	1	14	2	2	1
1002	1002	1002	15	2	1	13	12	N	1	6	1	15	2	6	3
1003	1003	1003	13	3	1	10	12	N	1	5	1	13	2	10	1
1004	1004	1004	12	4	1	7	9	N	1	4	1	12	2	14	2
1005	1005	1005	12	5	1	7	9	N	1	4	1	12	2	23	2
1006	1006	1006	13	6	1	11	12	N	1	5	1	13	2	32	2
1007	1007	1007	12	7	1	7	9	N	1	5	1	12	2	36	2
1008	1008	1008	10	8	1	8	11	N	1	3	1	10	2	40	2

Abbildung 6: Beispiel für ein Ergebnis eines simulierten Ablaufs

Durch die X- und Y-Koordinaten der linken oberen Ecke der Anlagenfläche und der Information über eine etwaige Drehung ist die Belegung einer Anlage eindeutig festgelegt. Vermerkt wird ferner als Ist-Fertigstellungstermin (später) die Woche, in der die Anlage die Halle verlässt. Dies führt beispielsweise zu dem in Abbildung 6 angegebenen Ergebnis.

Nach einem erfolgreichen Simulationslauf werden die folgenden Kennzahlen für eine Analyse (im Abschnitt Ergebnisse) berechnet:

- Für jeden Auftrag i seine Verspätung (V_i) als Differenz aus seinem Ist-Fertigstellungstermin minus seinem Soll-Endtermin, sofern eine Verspätung vorliegt. Anderenfalls ist sie 0.
- Für jede Woche t wird die belegte Fläche berechnet und kumuliert (A_t). Die Differenz aus der Hallenfläche (R_t) und der belegten Fläche (A_t) ist die freie Fläche (S_t); in der Woche t .
- Die Gesamtbearbeitungszeit (Z) zu einem Workload ist der späteste Ist-Fertigstellungstermin.

6 Ergebnisse

Die Grundlage für die Simulationsexperimente sind die Auftragsdaten aus dem Jahr 2020. Die direkte Verwendung dieser Daten führte zu einer sehr hohen Streuung der oben angegebenen Kennzahlen. Daher war es nicht möglich, statistisch signifikante Ergebnisse zu erkennen. Um zu aussagekräftigen Ergebnissen zu kommen, wurde die Folge an eingehenden Aufträgen über die einzelnen Wochen genauer untersucht. Es zeigte sich, dass es Wochen gibt, bei denen die Endtermine leicht eingehalten werden können, weil die verfügbare Zeit zur Bearbeitung – also Endtermin - möglicher Startzeitpunkt – hoch gegenüber der Nettobearbeitungszeit ist. Demgegenüber ist in anderen Wochen dieses Verhältnis aus verfügbarer Zeit und Nettobearbeitungszeit näher an eins oder sogar kleiner als eins; im zuvor betrachteten Fall ist der Quotient deutlich höher als 1.

Tabelle 1: Kennzahlen zu den beiden Klassen von Arbeitsvorräte (Workloads).

Arbeitsvorrat (Workload)	WL1	WL2
Anzahl an Aufträgen	50	50
Bearbeitungszeit: Minimum in [Wochen]	6	8
Bearbeitungszeit: Maximum in [Wochen]	17	19
Bearbeitungszeit: Mittelwert in [Wochen]	9,1	11,4
Bearbeitungszeit: Standardabweichung in [Wochen]	2,0	2,5
Produktbreite: Minimum in [Meter]	3	3
Produktbreite: Maximum in [Meter]	15	18
Produktbreite: Mittelwert in [Meter]	7,2	8,5
Produktbreite: Standardabweichung in [Meter]	2,7	2,9
Produktlänge: Minimum in [Meter]	3,0	3,0
Produktlänge: Maximum in [Meter]	12	12
Produktlänge: Mittelwert in [Meter]	8,9	10,2
Produktlänge: Standardabweichung in [Meter]	2,1	2,0

Soll-Fertigstellungstermin: Minimum in [Wochen]	11	11
Soll-Fertigstellungstermin: Maximum in [Wochen]	23	23
Soll-Fertigstellungstermin: Mittelwert in [Wochen]	14,3	14,3
Soll-Fertigstellungstermin: Standardabweichung in [Wochen]	2,8	2,8

Basierend auf dieser Analyse wurden zwei Klassen von Arbeitsvorräten gebildet. Jede besteht aus einer höheren Anzahl an Mengen an Aufträgen. Jede dieser einzelnen Mengen an Aufträgen wird von einem Planer und der KPZ-Regel geplant. Über alle Mengen an Aufträgen liegen die in der Tabelle 1 angegebenen Kennzahlen vor. Beide haben die gleichen Kennzahlen für die Endtermine. Deswegen bewirken höhere Bearbeitungszeiten bei der durch WL2 bezeichneten Klasse eine – im Sinne der obigen Bemerkung – kritischere Terminalsituation. Zudem sind die benötigten Flächen für die Aufträge in WL2 etwas höher.

Die durch die Planer und die KPZ-Regel erzielten Ergebnisse führen zu den in Tabelle 2 angegebenen Kennzahlen.

Tabelle 2: Kennzahlen zu den beiden simulierten Planungen.

	WL1		WL2	
	Planer	KPZ	Planer	KPZ
Kumulierte Verspätung in [Wochen]				
Minimum	71	76	287	266
Mittelwert	74,3	77,6	290,5	270
Maximum	78	75	294	304
Standardabweichung	2,37	2,76	3,50	2,70
Kumulierte freie Fläche in [Meter ²]				
Minimum	18920	21128	26989	24781
Mittelwert	20392	24808	29197	25885
Maximum	21128	27752	31405	26989
Standardabweichung	1040,86	2753,86	2208	1104
Gesamtbearbeitungszeit in [Wochen]				
Minimum	23	24	35	34
Mittelwert	23,6	25,6	36	34,5
Maximum	24	27	37	35
Standardabweichung	0,47	1,25	1,00	0,50
Durchschnittliche freier Fläche je Woche in [Meter ²]				
Minimum	822,61	880,33	771,11	728,85
Mittelwert	861,09	963,55	809,95	749,98
Maximum	880,33	1027,85	848,78	771,11
Standardabweichung	27,21	61,69	38,83	21,13

Verantwortlich für die Unterschiede ist, dass ein Planer Flächenanforderungen gegenüber Termineinhaltungsmöglichkeiten bevorzugt. Dies ist bei unübersichtlichen Mengen von Aufträgen eher der Fall. Umgekehrt gibt es Situationen, in denen der Planer deswegen bessere Ergebnisse erzielt, weil er durch eine bessere Nutzung der Werkshalle 5 eine schnellere Bearbeitung der Aufträge erreicht. Hinweis: die Angabe der durchschnittlichen Fläche bezieht sich auf die Gesamtbearbeitungszeit.

7 Zusammenfassung

Die Ressourcenbelegungsplanung in der Literatur berücksichtigt vor allem eine dominante Randbedingung wie die begrenzte Kapazität der Maschinen. Anforderungen an die Nutzung der begrenzten Fläche für die Montage von Anlagen wird in anderen Publikationen behandelt. Verantwortlich dafür sind unterschiedliche Vorgehensweisen zu diesen beiden Problemklassen.

Für die Endmontage bei der Krones AG sind beide Planungsprobleme simultan zu lösen. Dazu werden zwei erfahrene Planer eingesetzt.

Zur Analyse der Verbesserungsmöglichkeiten wurde ein Simulationsprojekt durchgeführt. Die bisherigen Simulationsexperimente zeigen, dass die Planer oft einen guten Ausgleich zwischen der Nutzung der Fläche und der Vermeidung von Verspätungen finden. Ist die Situation unübersichtlich, so liefert eine Bevorzugung der Termineinhaltung bereits durch die kürzeste Pufferzeitregel (KPZ-Regel) bessere Ergebnisse. Die beiden Planer sind daher angehalten, eine Zuteilungsentscheidung durch die KPZ-Regel zu berücksichtigen.

Die bisherigen Messungen belegen einen signifikanten Unterschied in der Nutzung der freien Flächen. Daher soll in der nächsten Projektphase das dadurch mögliche Potential an besseren Einlastungsentscheidungen untersucht werden. Dazu sollen Kombinationen aus Regeln zur Termineinhaltung mit Regeln zur Vermeidung von ungenutzter Fläche konzipiert und simulativ untersucht werden.

Literatur

- Drira, A., Pierreval, H., & Hajri-Gabouj, S. (2007). Facility layout problems: A survey. *Annual Reviews in Control* Volume 31 Issue 2, pp. 255-267.
- Ge, Y., & Wang, A. (2020). Spatial scheduling for irregularly shaped blocks in shipbuilding. *Computers & Industrial Engineering* Volume 152 Issue November 2020, pp. 1–14.
- Herrmann, F. (2011). *Operative Planung in IT-Systemen für die Produktionsplanung und -steuerung*. Regensburg, Germany: Vieweg+Teubner.
- März, L., Krug, W., Rose, O., & Weigert, G. (2010). *Simulation und Optimierung in Produktion und Logistik: Praxisorientierter Leitfaden mit Fallbeispielen*. Springer.
- Mayer, G., Pöge, C., Spieckermann, S., & Wenzel, S. (2020). *Ablaufsimulation in der Automobilindustrie*. Springer Verlag.
- Pinedo, M. (2016). *Scheduling: Theory, Algorithms and Systems*, Fifth Edition. New York, USA: Springer Science+Business Media.



*Simulation in Produktion
und Logistik 2021*
Jörg Franke & Peter Schuderer (Hrsg.)
Cuvillier Verlag, Göttingen 2021

Notations in Simulation Development: A state-of-the-art Literature Research

Notationen in der Simulationentwicklung: Eine state-of-the-art Literaturrecherche

Maik Trott, Marvin Auf der Landwehr, Christoph von Viebahn, Hochschule Hannover, Hannover (Germany), maik.trott@hs-hannover.de; marvin.auf-der-landwehr@hs-hannover.de; christoph-von.viebahn@hs-hannover.de

Abstract: With this systematic literature review we investigate the use of graphical tools and standard notations, such as the Business Process Modelling Notation (BPMN) and the Unified Modelling Language (UML), in the development of simulations. In our review, we focus on different simulation methods, such as agent-based simulation, discrete event simulation, system dynamic simulation and hybrid simulation approaches. We examine more than 1.000 scientific articles, which cover simulation approaches in the area of operations management and business. Our results provide insights into the frequency of the use of notations in simulation development, the relationship between notations and simulation methods and secondarily the frequency of used simulation methods in the business environment.

1 Problem Statement

In recent decades, the complexity of corporate decision-making for executives as well as managers has increased significantly. In addition to accretive requirements for information systems (IS), increasingly intertwined business processes, and intricate business relationships, the implications of modern trends such as the Internet of Things and Industry 4.0 have led to unprecedented complexities in corporate decision making (Xu et al. 2016). As a result, companies and decision-makers face significant challenges when designing or adapting business processes or operational systems (Auf der Landwehr et al., 2020). Computer simulation methodologies represent one possibility for decision support and evaluation in the managerial environment since they have proven to be an effective tool for description, exploration, and theory development (Davis et al. 2007; Auf der Landwehr et al. 2020). Dynamic simulation can be described as a “method for using computer software to model the operation of ‘real-world’ processes, systems, or events” (Davis et al. 2007, p.481). The burgeoning need for support, higher availability of computer resources, and the improved applicability and suitability of simulation environments could explain the increased number of simulation projects, publications, and combinations of simulation methods

in recent years (Xu et al. 2016; Brailsford et al. 2019). Here, the increased complexity of real-world problems, lead to a need for greater coordination and a higher intricacy in developing simulations. The information interface between the business perspective and the technological development of the conceptual model as an abstraction of the real world and the transition between theory development as well as simulation modelling is perhaps one of the most important aspects of simulation development and often the key challenge in the communication between functional experts and developers (Davis et al. 2007; Robinson 2013; García et al. 2014; Wagner 2018; Brailsford et al. 2019).

This type of interface challenge has been widely researched, and there are many studies available in the literature on IS and software engineering. Regarding the field of classic software development, standard notations for conceptual modelling are widespread and used extensively. However, in simulation development, such notations are not yet been formally established (Wagner 2018; Wilsdorf et al. 2020). Nevertheless, standard notations seem to be highly useful in simulation-related research and application contexts (Rosenthal et al. 2018).

The research gap concerning the use of notations in simulation was identified by Rosenthal et al. (2018), who conducted a literature analysis to investigate the usefulness of notations in simulation development. While their work provides a reasonable overview, their contribution is limited to business processes and does not consider different simulation methods. Jahangirian et al. (2010) and Brailsford et al. (2019) did consider different simulation methods, but did not address notations.

To acquire a holistic and up-to-date view of currently used simulation methods and potentially used notations and graphical tools, it is important to discover which notations are used and the extent to which they are used in various disciplines of the simulation community. In a first unstructured review of current literature considering 183 simulation approaches, we noted that there are publications that use standardized notations in their simulation approach, but the spread of standard notations in simulations has not been explored (Jahangirian et al. 2010; Rosenthal et al. 2018; Brailsford et al. 2019). To rectify this, we undertake a *systematic literature review* that analyzes the prevalence and utilization of notations in simulations. We thus propose the following research questions.

RQ1: How many scientific articles use standard notations or graphical tools for simulation development in the business domain?

RQ2: For which simulation methods are notations used and how often?

Based on the subject's topicality, the previous analysis of Rosenthal et al. (2018), and to focus our research framework on current events as well as changes, we propose a review period from 2015 to 2020. We thus rely on the findings of Rosenthal et al. (2018), although they addressed a different research focus. A five-year scope is also useful to set a limit for the literature search, in order to allow for a widespread conceptual assessment level (i.e. different simulation methods and business applications). Due to the relevance of simulation-based approaches for the field of production and logistics (P&L), we specifically focus on research approaches in this particular business area. Nevertheless, to ensure a comprehensive overview, we also take into account more general *business applications* (i.e. Business and workflow).

2 Theoretical Background

We focus on simulation applications from the business domain, using the 26 *business applications* proposed by Shafer and Smunt (2004), Brailsford et al. (2019), and the “Applications of simulation techniques in manufacturing and business” (Jahangirian et al. 2010, p.5-7). From this *business applications*, we derived six clusters to group the articles for our analysis, as per Jahangirian et al. (2010) (to be seen in Table 1).

Table 1: Business simulation applications (26); Business simulation clusters (6) [Jahangirian et al. 2010 (a); Shafer & Smunt 2004 (b); Brailsford et al. 2019 (c)]

Production and manufacturing	Logistic & Supply Chain Management
Manufacturing (c) & Purchasing (a, b)	Distribution (b) & Just-in-time (a, b)
Cellular manufacturing (a, b)	Inventory management (a, b)
Production planning & inventory (a, b)	Supply chain management (a, b, c)
Workforce planning (a, b)	Transportation management (a, c)
Business and workflow	Management
Assembly line balancing (a, b)	Financial management (a)
Capacity planning (a, b)	Maintenance management (a)
Facility location (a, b)	Knowledge management (a)
Forecasting (a, b)	Organizational design (a)
Resource allocation (a, b)	Project management (a, c)
Scheduling (a, b)	Quality management (a) & Strategy (a, b)
Training	Process
Management training & education (a)	Process engineer-manufacturing (a, b)
	Process engineer-service (a, b)

The following list provides an overview of the simulation methods used in the area of *business applications*. The methods that cannot be assigned to the presented methods, e.g. traffic simulation, are termed as other simulation (OS) (Jahangirian et al. (2010).

System Dynamics (SD) is a time-continuous method characterized by positive and negative feedback loops. SD aims to create a macroscopic image of systems focused on the whole system rather than short-term events (Morgan et al. 2017).

Discrete Event Simulation (DES) is mostly used on the micro- and mesoscopic level to model events in time-discrete systems. Entities pass through a series of events so that the flow, variability, and random factors in a system can be observed. DES does not aim to obtain feedback from the system (Morgan et al. 2017).

Agent-Based Simulation (ABS) systems consist of a network with agents that follow predefined rules and conditions in a time-continuous system. The agents then interact with each other and the environment, allowing inferences to be made about the system or individual agents (Macal 2016; Brailsford et al. 2019).

Monte Carlo Simulation (MC) is one of the oldest simulation methods, especially in the *fields* of production or processes. MC solves static or numerical problems, such as evaluations or risk deviations (Jahangirian et al. 2010).

Hybrid simulation (HS) is a combination of two or more simulation methods that are combined to solve problems simultaneously (Mustafee et al. 2018). HS can represent both time-continuous and time-discrete systems (Jacob et al. 2010). Modern HSs can be developed in a single development environment and are mostly meshed internally without external interfaces (Brailsford et al. 2019).

In addition to many conceptual models for the development of simulations from the simulation community - see the overview by Wilsdorf et al. (2020) - there are notations in other areas of computer science that have long been accepted as standard (Rosenthal et al. 2018; Wilsdorf et al. 2020). These notations from software development are also used in the development of simulation models. The best-known current notations for development or business process modelling are:

Business Process Modelling Notation (BPMN) is a widely used notation and official standard for process modelling. All information about BPMN can be found on the official OMG site (OMG 2010). Due to its widespread use, there are many approaches and attempts to use BPMN for simulation, especially in the area of DES (Rosenthal et al. 2018). To use BPMN for modelling and describing simulation projects, many authors have suggested extending BPMN notation by adding input, output, and other environmental parameters. A further step is the use of extensions that transform BPMN into an executable program, such as eBPMN, domain-specific simulation languages that softens the boundaries between notation and simulation. Thus, for example, eBPMN can be considered as DES which are implemented with the help of graphic notations (Bocciarelli et al. 2019).

Unified Modelling Language (UML) is an internationally standardized software modelling language and is likely the most widely used modelling technique in software engineering, especially in object-oriented development (Xie 2008). All information about UML can be found on the official OMG (OMG 2017). Beyond software development, UML is also used in business process modelling due to the large number of possible modelling tools provided (Allweyer 2016). UML can be used for simulation development (i.e., UML activity diagrams), but there are also attempts to modify UML for simulation-compliant notations (i.e., Systems Modelling Language, SysML) (Weyprecht & Rose 2011; Rosenthal et al. 2018).

Stock and Flow Diagrams and **Causal Loop Diagrams** (hereafter together **Stock&Flow**) do not represent a IS notation per se; rather, the diagram types are used to represent dynamic systems. With Stock&Flow, system relationships can be represented in a continuous space with feedback loops (Bala et al. 2016).

We have included the graphical formalization of the Stock&Flow, for the purpose of completeness, as this is frequently used in SD. There are also other notations that can help to plan and develop simulation projects, such as PetiNets, or (extended) event-driven process chain, which we abbreviate collectively as **EPC**.

3 Methodology

Our procedure chosen can be seen in Figure 1. The *first step* was "scope definition," or the definition of the problem and formulation of the research questions. For this step, we - the authors - used an unstructured literature review that we performed together in a 1-day workshop. In the *second step*, we elaborated the methodology, the related, theoretical foundations, and the most important terms of our project. This theoretical foundation is found in the previous *Chapter 2* and starts with a definition of *business applications* for simulation projects and publications (Jahangirian et al. 2010; Brailsford et al. 2019). In addition, we discuss notations for simulation development (Robinson 2013; Rosenthal et al. 2018). There is a special focus on the BPMN and UML. But *Stock&Flow* and other notations (e.g. EPC) are also under

consideration (Allweyer 2016; Rosenthal et al. 2018; Robinson 2020). Furthermore, various simulation methods are described, including *SD*, *DES*, *MC*, and *ABS*. *HS* and *OS* are also discussed (Jacob et al. 2010; Brailsford et al. 2019).

After laying out the theoretical foundation, we begin the systematic literature review in *Chapter 4*, which is divided into a database search (*step three*) and deep search (*step four*) following the two-step model of Jahangirian et al. (2010). The database search included journal and conference selection, a literature database definition, and the main database search, which included a continuous evaluation of sources and exclusion of irrelevant publications through the allocation of business and simulation applications as well as reduction of duplicate or non-English publications. (Vom Brocke et al. 2009). Our abstract filtering included title review, abstract review, a keyword review of the full text, and a consideration of the full research approach to classify the simulation method and the use of notations. Through the deep search, we summarized the literature search process of Vom Brocke et al. (2009), which suggested an additional backward and a forward search of the most relevant articles for quality assurance. For the *fifth step* of the literature analysis, we deductively developed a matrix that summarizes relevant information like the simulation method, the notation used, or the associated business application. The analysis of the findings merged in the matrix, are presented in the discussion in *Chapter 5*. We then summarize our results to answer the research questions in *step 6* (*Chapter 6*).

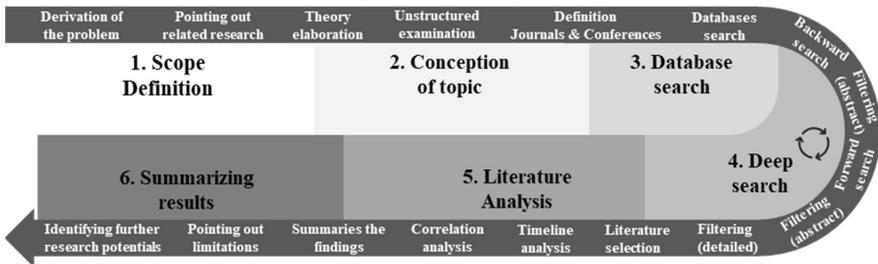


Figure 1: Procedure of our literature analysis (based on Vom Brocke et al. (2009))

4 Systematic Literature Review

As suggested by Jahangirian et al. (2010), we identified the journals and conferences relevant to our research environment to derive associated literature database for our review. To restrict the journal and conference selection, we examined the relevance in the field of simulation, impact factors of the journals, and rely on the selection of Jahangirian et al. 2010, Rosenthal et al. (2018), and Robinson (2020). In total, we focused on 12 of the most popular journals and conferences in the field of simulation. Based on the journal selection, we chose the Web of Science database for the review, which includes all the primary sources listed in *Table 2*.

Like Jahangirian et al. (2010), we searched for simulation approaches in the area of the *business applications*. We searched using the Boolean keyword combination “(simula* OR 'system dynamic*)”. To narrow the large number results, we extended the search with our business clusters, complementing the search grid of Jahangirian et al. (2010) to “logistic”, “production”, “manufacturing”, “process”, “management”, “business”, “workflow”, “supply chain”, and “training”.

Table 2: Primary journals and conferences

Business Process Management Journal	J. of Operational Research Society
Decision Sciences	Journal of Business Research
Euro. J. of Operational Research	Journal of Cleaner Production
Int. J. of Operations & Production Management	Journal of Simulation
Int. J. of Production Research	Transportation Science
International Conference on IS	Winter Simulation Conference

After applying the main keywords and the application keywords in the title, abstract and keyword search for the selected journals, we found 1,068 articles. Whereby we have only relied on English publications. We filtered the results to 383 relevant articles by initial coding of the papers, for which we manually searched the title and abstract for business simulation approaches (Jahangirian et al. 2010). In this step we have also eliminated duplicate articles through filtering and a final manual search for duplicates. In the course of our filtering process, we used backward and forward searching in 40 particularly relevant articles and identified 11 additional articles. The final 396 articles with 400 simulation approaches were examined in detail according to the simulation methods and notation used. Here, some authors used several simulation approaches, for example for validation reason. *Table 3* offers an overview of the review. A description of the identified *business applications*, notations used, and a combined consideration of methods and notations follows in *Chapter 5*.

Table 3: Databases and relevant articles (several methods are possible)

Databases	Articles	Relevant articles	Simulations per method / relevant articles					
			ABS	DES	SD	MC	HS	OS
Web of Science	1068	385	43	111	62	9	90	73
Back&Forward	40	11	1	9	0	0	0	2
Total	1108	396	44	120	62	9	90	75

5 Analysis of the Findings

In the course of our literature analysis, we examined 396 scientific articles presenting 400 simulation approaches and have found 83 articles with notations. We subjected the applications, methods, and notations to perform a correlation analysis. For this purpose, we used the Pearson correlation coefficient (r) with 2-sided significance (p).

Between the *business applications* and the simulation methods as well as notations, only one weak correlation could be found, namely between the applications and ABS ($r = -0.122 \mid p < 0.05$). Closer inspection shows that many ABSs are assigned to "business and workflow" (34%) and P&L (39%); whereby 40% of all articles belong to P&L. *Fig 2* shows the distribution of the simulation methods used. The most commonly applied method is DES (30%), followed by HS (23%). In P&L the most used method remains DES (29%) followed by HS (26%). The use of HS increased in the five years considered. It should be emphasized that 28% of the articles from the logistics sector used HS, representing the highest usage of HS. Since HS is a combination of several simulation methods, we break down the HS types in *Fig 4*.

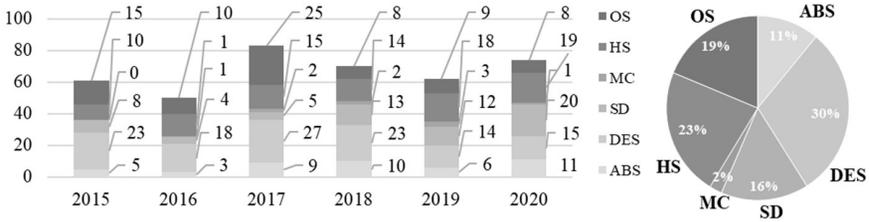


Figure 2: Methods used (400 simulations); proportions of methods (400 simulations)

The most used tool out of the 83 identified notation articles, was Stock&Flow (67%), followed by UML (17%) and BPMN (12%). The preference for Stock&Flow can be explained by the complementary attributes of Stock&Flow and SD. As seen in Fig 3, notations were used in 83 of the 396 articles examined, or 20.96%. In detail, the following notation uses were identified: 14 UML (3.54%), 10 BPMN (2.53%), 3 EPC (0.76%) and 56 Stock&Flow (14.14%). In addition to the reviewed notations, we found 220 articles that used flowcharts or similar schematic representations and no other notation, representing 56% of our relevant articles. Looking at P&L alone, UML was used in 3%, BPMN in 3%, EPC in 1%, and Stock&Flow in 9% of 160 P&L articles. Excluding articles that used SD or Stock&Flow, the share of notations (UML, BPMN, and EPC) is 8.5% (out of 319 articles). While an increase of the Stock&Flow can be seen, the use of UML and BPMN is constant over the studied period.

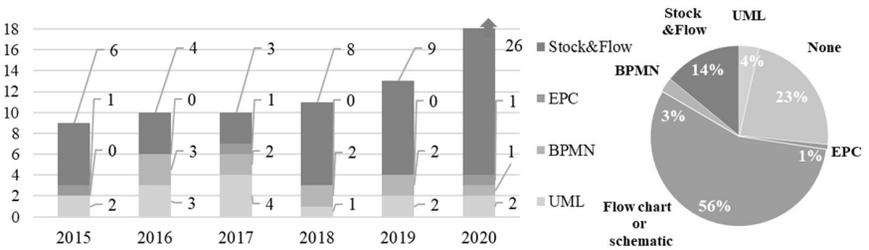


Figure 3: Notations used (multi. notations per article possible); Proportions of notations (rounded; 396 articles total)

Next, we looked at the correlations between simulation methods and the used notations, see Fig 4. Only one strong correlation between SD and Stock&Flow was identified ($r = 0.643 \mid p < 0.001$), which is due to the methodological relationships, explained in Chapter 2. Furthermore, there is a moderate negative correlation between DES ($r = -0.236 \mid p < 0.001$) and Stock&Flow, which is related to the previous correlation. However, one moderate correlation between DES and BPMN ($r = 0.209 \mid p < 0.01$) and a weak correlation between ABS and UML ($r = 0.106 \mid p < 0.05$) were also identified, indicating another methodological correlation. Thus, BPMN is often used to model processes within DES paradigms. In addition, many tools offer DES on the basis of BPMN, which also leads to a conclusive connection especially when using simulation environments that allow graphical simulation with

BPMN; like Di Leva et al. (2017) who used a framework called BP-M*. The use of UML for ABS can also be explained on the basis of the methodical relationships, because the object orientation of UML allows good representation of ABS models (e.g., Liu et al. 2019, who developed a use-case diagram to demonstrate an autonomous system in their ABS). As can be seen in Fig 5, UML was used in nine percent of the ABS articles, BPMN in eight percent of the DES articles, and Stock&Flow in 66% of the SD articles. The high use of Stock&Flow in the 90 HS can be explained by the share of HS with combined SD (see Fig 4).

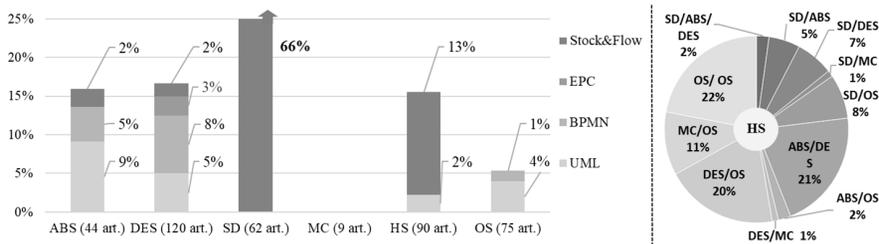


Figure 4: Notations in Methods (400 simul.); Combined methods in the HS (90 simul.)

6 Conclusion

This study conducted a wide-ranging literature review using publications from 2015–2020 based on six *business applications* and identified 83 simulation articles with notations from a set of 396 relevant articles. A total dataset of 1,068 articles was screened, including a forward-backward search for 40 articles.

In relation to **RQ1**, our results show that 20.96% of the considered simulations use a graphical tool (e.g. Stock&Flow) or a standard notation (such as UML or BPMN). The percentage of UML, EPC, and BPMN usage (excluding Stock&Flow and SD) was 8.5%. Among the articles dealing with P&L, the share of UML, EPC and BPMN usage was 7%. In comparison, Rosenthal et al. (2018) identified 38 notations in 300 articles in their review of process models (12.6%). If we look in vain only at the DES articles, then 15% of this article uses notations. Given the different observation areas and times, our result and those of Rosenthal et al. (2018) are comparable, particularly because our results assume a constant use of IS notation in simulations approaches since 2015. To answer **RQ2** we could show the clear connection between SD and Stock&Flow and also the connections between the simulation method ABS and UML as well as between DES and BPMN. A connection with HS and notations was not found; this is likely because there are too many different expressions of HS.

Our results are transferable to a wide range of simulation approaches due to the broad scope of our study. Moreover, they reflect current discourse on notations in simulation, which can provide future systematic guidance for scholars and developers opting to employ notations for increasing conceptual simulation credibility. However, our review is methodologically limited due to its restriction to *business applications*. Moreover, the classification of *business applications* and notations can only be carried out to a limited extent by means of a literature analysis; which offers limited insight

into the underlying research and simulation development. For this reason, we could not determine how the authors proceeded in the development of the simulation and whether notation of the IS was used in the background.

Finally, more research on the use of notations in the simulation environment would be useful. Future studies should investigate why notations are used in some cases and not in others to clarify how standardization in the representation of simulation projects can be achieved; this will increase the quality and comparability of studies (Auf der Landwehr et al. 2020).

The analyzed, articles are available at: <https://doi.org/10.25625/XFWDKB>

References

- Allweyer, T.: BPMN 2.0: Introduction to the Standard for Business Process Modelling (2nd ed.). Norderstadt, Germany: Books on Demand 2016.
- Auf der Landwehr, M.; Trott, M.; von Viebahn, C.: Waste of Time and Money? Constructing an Applicability Framework for Organizational Use of Simulation Studies and Digital Twins. In: Proceedings of the International Conference on Information Systems, Hyderabad (India), December 2020.
- Auf der Landwehr, M.; Trott, M.; von Viebahn, C.: Computer Simulation as Evaluation Tool of Information Systems: Identifying Quality Factors of Simulation Modeling. In: Proceedings of the 2020 IEEE 22nd Conference on Business Informatics, Antwerp (Belgium), 22.–24. June 2020, pp. 211-220.
- Bala, B. Bala, B. K.; Arshad, F. M.; Noh, K. M.: System dynamics: Modelling and simulation. Singapore: Springer 2016.
- Bocciarelli, P.; D'Ambrogio, A.; Giglio, A.; Paglia, E.: Bpmn-based business process modelling and simulation. In: Proceedings of the 2019 Winter Simulation Conference, Maryland (USA). December 2019, pp. 1439-1453.
- Brailsford, S. C.; Eldabi, T.; Kunc, M.; Mustafee, N.; Osorio, A. F.: Hybrid simulation modelling in operational research: A state-of-the-art review. *European Journal of Operational Research* 278 (2019) 3, pp. 721-737.
- Davis, J. P.; Eisenhardt, K. M.; & Bingham, C. B.: Developing theory through simulation methods. *Academy of Management Review* 32 (2007) 2, pp. 480-499.
- Di Leva, A.; Sulis, E.; Vinai, M.: Business Process Analysis and Simulation: The Contact Center of a Public Health and Social Information Office. *Intelligent Information Management* 9 (2017) 5, pp. 189-205.
- García, M. T.; Barcelona, M. A.; Ruiz, M.; García-Borgoñón, L.; Ramos, I.: A discrete-event simulation metamodel for obtaining simulation models from business process models. In: Escalona, M.; Aragon, G.; Linger, H.; Lang, M.; Barry, C.; Schneider, C. (Eds.): *Information System Development*. Springer 2014, pp. 307-317.
- Jacob, M.; Suchan, C.; Ferstl, O. K.: Modelling of Business Systems using Hybrid Simulation: A New Approach. In: Proceedings of the European Conference on Information Systems, Pretoria (South Africa). September 2010, pp. 1-12.
- Jahangirian, M.; Eldabi, T.; Naseer, A.; Stergioulas, L. K.; & Young, T.: Simulation in manufacturing and business: A review. *European Journal of Operational Research*, 203, 2010 1, pp. 1-13.
- Liu, Y.; Wang, T.; Zhang, H.; Cheutet, V.; Shen, G.: The design and simulation of an autonomous system for aircraft maintenance scheduling. *Computers & Industrial*

- Engineering 137 (2019).
- Macal, C. M.: Everything you need to know about agent-based modelling and simulation. *Journal of Simulation* 10 (2016) 2, pp. 144-156.
- Morgan, J. S.; Howick, S.; Belton, V.: A toolkit of designs for mixing discrete event simulation and system dynamics. *European Journal of Operational Research* 257 (2017) 3, pp. 907-918.
- Mustafee, N.; Powell, J. H.: From hybrid simulation to hybrid systems modelling. In: *Proceedings of the 2018 Winter Simulation Conference (WSC)*, Gothenburg (Sweden), December (2018). pp. 1430-1439.
- OMG.: About the Business Process Model and Notation Specification version 2.0, (2010). Retrieved March 8, 2021, from <https://www.omg.org/spec/BPMN/2.0/About-BPMN/>.
- OMG.: About the unified modelling language specification version 2.5.1, (2017). Retrieved March 8, 2021, from <https://www.omg.org/spec/UML/About-UML/>.
- Robinson, S.: Conceptual modelling for simulation. In: *Proceedings of the 2013 Winter Simulations Conference*, Washington (USA). December 2013, pp. 377-388.
- Robinson, S.: Conceptual modelling for simulation: Progress and grand challenges. *Journal of Simulation* 14 (2020) 1, pp. 1-20.
- Rosenthal, K.; Ternes, B.; Strecker, S.: Business process simulation: A systematic literature review. In: *Proceedings of the 26th European Conference on Information Systems*, Portsmouth (UK). June 2018, pp. 1-17.
- Shafer, S. M.; Smunt, T. L.: Empirical simulation studies in operations management: context, trends, and research opportunities. *Journal of Operations Management* 22 (2004) 4, pp. 345-354.
- Vom Brocke, J.; Simons, A.; Niehaves, B.; Niehaves, B.; Reimer, K.; Plattfaut, R.; & Cleven, A.: Reconstructing the giant: On the importance of rigour in documenting the literature search process. In: *Proceedings of the European Conference on Information Systems*, Verona (Italy). 2009, pp.1-14.
- Wagner, G.: Information and Process Modelling for Simulation–Part I. *Journal of Simulation Engineering* 1 (2018), pp. 1-26.
- Weyprecht, P.; Rose, O.: Model-driven development of simulation solution based on SysML starting with the simulation core. In: *Proceedings of the 2021 Spring Simulation Conference*, Virginia (USA). April 2011, pp. 189-192.
- Wilsdorf, P.; Haack, F.; Uhrmacher, A. M.: Conceptual models in simulation studies: Making it explicit. In: *Proceedings of the 2020 Winter Simulation Conference*. December 2020, pp. 2353-2364.
- Xie, Y.: Process modelling and simulation based on extended UML activity and GPSS. In: *Proceedings of the 2008 IEEE International Conference on Automation and Logistics*, Qingdao (China). September 2008, pp. 2931-2935.
- Xu, J.; Huang, E.; Hsieh, L.; Lee, L. H.; Jia, Q. S.; Chen, C. H.: Simulation optimization in the era of Industrial 4.0 and the Industrial Internet. *Journal of Simulation* 10 (2016) 4, pp. 310-320.



Autorenindex

Keller, Volker	1
Knapp, Marc	11
Abraham, Tim	21
Langer, Adrian	21
Henningsen, Nadja	21
Ortmeier, Christian	21
Martin, Niels L.	21
Sjarov, Martin	31
Distler, Johannes	31
Fürst, Jens	31
Weigold, Matthias	51
Niehus, Christian	61
Götz, Matthias	75
Abraham, Tim	103
Thomitzek, Matthias	103
Wincheringer, Walter	123
Mühlmann, Felix	133
Fürst, Jens	133
Lechler, Tobias	143
Necil, Jan	157
Herzog, Alexander	157
Busch, Heike	157
Schallow, Julian	167
Deuse, Jochen	167
Wucherer, Stefanie	177
Klarmann, Steffen	177
Kerber, Florian	177
Ulrich, Jan Hendrik	187
Zadek, Hartmut	187
Reyes Rubiano, Lorena Silvana	187
Scherbarth, Christoph	197
Prehm, Johann	207
Szczerbicka, Helena	217



Autorenindex

Hinckeldeyn, Johannes	257
Genc, Emin	267
Natterer, Niklas.....	267
Hörauf, Leenhard.....	275
Müller, Rainer.....	275
Verbeet, Richard.....	285
Scholl, Wolfgang.....	295
Leissau, Madlene.....	295
Palmer, Gregory	305
Fohlmeister, Silas	305
Stark, Rainer	315
Meisen, Tobias	325
Gützlaff, Andreas	325
Lütticke, Daniel	325
Samsonov, Vladimir	325
Schmidhuber, Matthias.....	325
Sekar, Sughanthan	337
Abdine, Mazhar Zein El	337
Deuse, Jochen	347
Hoffmann, Florian	347
Hernandez Moreno, Victor	347
Schulte, Lukas	347
Kuhnle, Andreas	357
Altenmüller, Thomas	357
Hoffmann, Clara	357
Salloum, Yara	387
Gellrich, Sebastian.....	409
Hürkamp, André.....	409
Czarski, Marvin	409
Aschersleben, Franziska	409
Rett, Angelika.....	421
Widera, Adam.....	433
Reuter-Oppermann, Melanie	433
Neuhäuser, Thomas	443
Ullmann, Felix	443
Hohmann, Andrea.....	443



Schilp, Johannes	443
Komenda, Titanilla	453
Schelle, Maximilian.....	453
Kamhuber, Felix.....	453
Schlund, Sebastian.....	453
Schilp, Johannes	463
Sobottka, Thomas	475
Lielacher, Magdalena	475
Sihn, Wilfried	475
Müller, Rainer.....	485
Adelsbach, Jan	485
Bashir, Attique.....	485
Kuhn, Dominik	485
Hörauf, Leenhard.....	485
Wenk, Matthias.....	495
Davidsson, Paul	515
Johansson, Emil.....	515
Persson, Jan	515
Pache, Hannah	545
Franzkeit, Janna.....	545
Munkelt, Torsten	565
Matthes, Marvin	565
Krockert, Martin.....	565
Klotz, Matthias	575
Hausberger, Thomas.....	575
Geiger, Dominique Mathäus.....	575
Verl, Alexander	585
Büttner, Thomas	595
Wenzel, Ken	595
Friedemann, Marco.....	595
Le Louarn, Arthur.....	609
Brützel, Oliver	609
Stricker, Nicole.....	609
Grünhöfer, Lukas.....	619
Schiessel, Svenja	619
Tjaden, Greta	641



Autorenindex

Savasci, Faruk.....	653
Altendorfer, Klaus.....	399
Ammouriova, Majsja.....	227
Auf der Landwehr, Marvin.....	525, 663
Bergmann, Sören.....	367, 377
Betker, Vincent.....	555
Billiet, Harold.....	315
Borst, Fabian.....	51
Claus, Thorsten.....	75
Denkena, Berend.....	305
Dér, Antal Zoltán.....	21
Dettelbacher, Johannes.....	85
Dittrich, Marc-André.....	305
Donhauser, Toni.....	267
Dröder, Klaus.....	409
Eichenwald, Max.....	275
Felberbauer, Thomas.....	399
Feldkamp, Niclas.....	367, 377
Fottner, Johannes.....	421
Franke, Jörg.....	31, 133, 143, 495, 619
Freitag, Michael.....	217
Galka, Stefan.....	197
Genath, Jonas.....	367, 377
Gustafsson-Ende, Linda.....	505
Heger, Jens.....	337, 545
Hellingrath, Bernd.....	433
Herbert, Meike.....	31, 133
Herrmann, Christoph.....	21, 103, 409
Herrmann, Frank.....	75, 653
Hilmer, Frank.....	505
Horstkemper, Dennis.....	433
Jahn, Carlos.....	535, 545
Janke, Tim.....	325
Jodehl, Annika.....	387
Kauke, Dirk.....	421
Kemmerling, Marco.....	325



Kemp, Daniel.....	305
Kienzlen, Annika.....	585
König, Markus.....	387
Lang, Sebastian.....	207
Lange, Ann-Kathrin.....	535
Lanza, Gisela.....	357, 609
Laroque, Christoph.....	295
Lenze, David.....	167
Lichtenstern, Isabella.....	177
Lorig, Fabian.....	515
Lütjen, Michael.....	217
Massow, Benjamin Bastian.....	575
May, Marvin Carl.....	357
Metzner, Maximilian.....	619
Meyer, Torben.....	247
Middelhoff, Michael.....	433
Mieth, Carina.....	641
Moog, Daniel.....	51
Moroff, Nikolas.....	61
Müller, Marcel.....	187, 207
Nellen, Nicole.....	535
Overbeck, Leonard.....	609
Pöge, Carsten.....	247
Pollet, Martin.....	157
Prell, Bastian.....	93
Purgander, Jana.....	257
Putz, Matthias.....	575
Rabe, Markus.....	93, 123, 227
Rank, Sebastian.....	113
Reggelin, Tobias.....	187, 207
Reith, Karl-Benedikt.....	113
Rippel, Daniel.....	217
Rolf, Benjamin.....	207
Roßmann, Jürgen.....	41, 629
Rückgauer, Andreas.....	237
Russwurm, Eva.....	495



Autorenindex

Scheer, Richard.....	11
Schlegel, Andreas	575
Schluse, Michael.....	41
Schlüter, Wolfgang.....	85
Schmid, Alexander	475
Schmidl, Elisabeth.....	495
Schmidt, Thorsten.....	113, 555
Schneider, Germar	295
Schreiber, Lucas	61
Schuh, Günther	325
Seiringer, Wolfgang	399
Selmaier, Andreas.....	31, 133
Selmair, Maximilian	75
Sitz, Sarah.....	143
Sohny, Tobias	123
Sokoll, Kristina.....	505
Spieckermann, Sven	377
Staritz, Johannes	525
Stauber, Stephan	377
Stoldt, Johannes	93
Straßburger, Steffen.....	11, 367
Terbrack, Hajo.....	75
Thewes, Markus	387
Thiede, Sebastian.....	93
Thieling, Jörn.....	629
Trott, Maik.....	525, 663
Ventura Silva, Gabriela	103
Völker, Michael.....	555
Völker, Sven.....	285, 565
von Braunschweig, Philipp.....	505
von Viebahn, Christoph.....	525, 663
Voss, Thomas	337, 545
Wenzel, Sigrid	93
West, Nikolai.....	347
Xie, Xinyi	595
Zeiser, Reinhard	443



Zerreis, Maximilian	143
Zimbrod, Patrick.....	463
Zülch, Gert.....	1





