Michael Fiedeldey

Analyse und Prognose elektrischer Lastgangzeitreihen

Cuvillier Verlag Göttingen

Analyse und Prognose elektrischer Lastgangzeitreihen

Von der Fakultät für Elektrotechnik und Informatik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades

Doktor-Ingenieur

genehmigte

Dissertation

von

Dipl.-Ing. Michael Fiedeldey

geboren am 1. Mai 1966 in Osnabrück

2010

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, 2010
Zugl.: Hannover, Univ. Diss., 2010

978-3-86955-363-4

1. Referent: Prof. Dr.- Ing. Ernst Gockenbach
2. Referent: Prof. Dr.- Ing. Gerd Balzer

Datum der mündlichen Prüfung: 21. Januar 2010

© CUVILLIER VERLAG, Göttingen 2010
Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie) zu vervielfältigen.
1. Auflage, 2010
Gedruckt auf säurefreiem Papier

978-3-86955-363-4

Vorwort

Diese Arbeit entstand in den Jahren 2003 bis 2009 und wurde durch das Institut für Energieversorgung und Hochspannungstechnik, Fachgebiet Hochspannungstechnik (Schering-Institut), der Leibniz Universität Hannover betreut.

Mein besonderer Dank gilt Herrn Prof. Dr.-Ing. Ernst Gockenbach, dem Institutsdirektor und Leiter des Fachgebietes Hochspannungstechnik, der die Arbeit ermöglichte und diese in jeder Hinsicht geduldig und mit hilfreichen Ratschlägen und Anregungen begleitet und unterstützt hat.

Herrn Prof. Dr.-Ing. Gerd Balzer, Leiter des Fachgebietes Elektrische Energieversorgung und Systemführung in Energieversorgungsnetzen an der Technischen Universität Darmstadt, danke ich für die bereitwillige Übernahme des Korreferates und das von ihm entgegengebrachte Interesse an dieser Arbeit.

Mein Dank geht auch an Herrn Prof. Dr.-Ing. habil. Lutz Hofmann, Leiter des Fachgebietes Elektrische Energieversorgung an der Leibniz Universität Hannover, für die Übernahme und Leitung des Prüfungsausschusses.

Des Weiteren möchte ich mich bei Herrn Prof. Dr. rer. nat. Theodor Gervens und Herrn Prof. Dr. rer. nat. Jürgen Biermann sowie Herrn Dipl.-Inf. Christian Karrasch von der Fachhochschule Osnabrück, Fakultät Ingenieurwissenschaften und Informatik, für ihre zahlreichen Impulse und eine Vielzahl von fruchtbaren Diskussionen bedanken, die wir im Rahmen des gemeinsamen Projektes E4T führten. Ebenso möchte ich dankend den Einsatz der vielen Diplom- und Studienarbeiter erwähnen, die nicht müde wurden, große Datenmengen auszuwerten und zu verarbeiten und ohne deren Leistungen diese Arbeit nicht zustande gekommen wäre.

Großer Dank gilt auch meinen Eltern, die mir meine Ausbildung und mein Studium ermöglichten und mich immer unterstützend begleitet haben.

Nicht zuletzt möchte ich mich bei meiner Frau Marita bedanken, die immer viel Geduld und Verständnis für meine Arbeit aufgebracht und mir „den Rücken frei gehalten" hat.

Kempten im Allgäu, im Januar 2010

Michael Fiedeldey

Inhaltsverzeichnis

Inhaltsverzeichnis I
Abbildungsverzeichnis IV
Tabellenverzeichnis VI
Verzeichnis der verwendeten Formelzeichen und Abkürzungen VIII
Kurzfassung 1
Abstract 2
1 Einleitung 3
1.1 Stand des Wissens 5
1.2 Zielsetzung und Gliederung der Arbeit 7
2 Energiewirtschaftliche Rahmenbedingungen 10
2.1 Prognosemodelle und -methoden 10
2.2 Planungshorizonte in der Energiewirtschaft 11
2.3 Energiewirtschaftliche Lastprognosen 13
2.4 Öffnung des Energiemarktes in Deutschland 14
2.5 Anforderungen an Lastprognosen 16
2.5.1 Anforderungen des Handels und Vertriebs 17
2.5.2 Anforderungen der Netzbetreiber 19
2.5.3 Lastprognosesysteme 23
2.6 Werkzeuge und IT- Systeme 28
3 Methodische Grundlagen der Zeitreihenanalyse 30
3.1 Einflüsse und Einflussgrößen auf die elektrische Last 30
3.2 Zeitreihenprognosen 35
3.3 Prozess der Prognoseerstellung 37
3.4 Beurteilungsmaß und Qualität von Prognosen 38
3.4.1 Ursachen von Lastprognosefehlern 39
3.4.2 Fehlerkenngrößen 39
3.5 Zeitreihenanalyse 41
3.5.1 Allgemein 41
3.5.2 Darstellung von Zeitreihen 43
3.5.3 Empirische Momente und Stationarität 45
3.6 Klassische Zeitreihenanalyse 46
3.6.1 Komponentenmodell 47
3.6.2 Trendbestimmung 49
3.6.3 Methode der kleinsten Quadrate 50
3.6.4 Filter 51
3.6.5 Zyklische Schwankungen 52
3.7 Stochastische Prozesse 53

4 Neuronale Netze 56
4.1 Grundlagen und historische Entwicklung 57
4.1.1 Das biologische Vorbild 60
4.1.2 Formal mathematische Beschreibung eines Neurons 62
4.2 Aufbau neuronaler Netze und Netztopologien 65
4.2.1 Feedforward-Netze 67
4.2.2 Jordan- und Elman-Netze 68
4.3 Lernalgorithmen 71
4.3.1 Das Lernproblem 71
4.3.2 Gradientenverfahren 72
4.3.3 Backpropagation 73
4.3.4 Feedforward-Algorithmus 77
4.3.5 Backpropagation für rekurrente Netze 79
4.4 Selbstorganisierende Karten (SOM) 82
4.4.1 Prinzip und Netzarchitektur der SOM 83
5 Entwicklung von Lastgangprognosemodellen 88
5.1 Modellbildung und Anpassung stochastischer Prozesse 88
5.1.1 Zerlegungsansatz und Modelle der Restkomponente 88
5.1.2 Modellbildung und Bestimmung der Parameter 91
5.2 Entwicklung eines entkoppelten Modells mit exogenen Größen 93
5.3 Prognosen von Zeitreihen mit neuronalen Netzen 95
5.3.1 Grundlagen bei der Prognoseerstellung mit neuronalen Netzen 95
5.3.2 Problemstellung und Entwicklungsmethodik 97
5.4 Mustererkennung in Lastgangdaten mit SOM 101
5.4.1 Entwicklung einer zweidimensionalen Netzstruktur zur Analyse von Lastgangdaten 101
5.4.2 Lernverfahren und Datenanalyse mit SOM 102
6 Fallstudien 107
6.1 Untersuchungsdaten und Vorgehensweise 107
6.2 Lastgang- und Wetterdatenanalyse 109
6.2.1 Lastgang 109
6.2.2 Empirische Momente und Stationarität 110
6.2.3 Trend- und Periodenanteil 111
6.2.4 Tagtypbestimmung 114
6.2.5 Wetterelemente 118
6.2.6 Zusammenhang zwischen Wettereinflussgrößen und Lastgang 119
6.3 Analyse von Einflüssen auf Tagtypen mit Hilfe SOM 124
6.4 Prognosen 128
6.4.1 Prognose mit klassischen Modellen 128
6.4.2 Prognose durch Kombination von deterministischen und stochastischen Ansätzen 132
6.4.3 Prognosen mit dem „entkoppelten“ autoregressiven Prozess 138
6.4.4 Zusammenfassung der Ergebnisse 139

6.5 Lastprognosen mit neuronalen Netzen 141
6.5.1 Einschichtiges Netz 141
6.5.2 Untersuchung verschiedener Netztopologien 143
6.5.3 Anwendungsstrategie mit kleinen Netzen 145
6.5.4 Zusammenfassung der Prognoseergebnisse mit neuronalen Netzen 147
7 Zusammenfassung und Ausblick 149
8 Literaturverzeichnis 154
9 Anhang 165
9.1 Lineare und Polynominale Trends 165
9.2 Periodogramm und Fourieranalyse 166
9.3 2-Stichproben T-Test 168
9.4 Chi-Quadrat-Test 169
9.5 Yule-Walker-Gleichung 170
9.6 Durbin-Methode 171
9.7 Ergebnisse des Mittelwertstationaritäts- sowie des Varianzstationaritätstests .. 172
9.8 Analyse ausgewählter Wetterelemente 172
9.8.1 Temperatur 173
9.8.2 Globalstrahlung 174
9.8.3 Windgeschwindigkeit 175
9.8.4 Luftdruck 176
Lebenslauf 178

Abbildungsverzeichnis

Abbildung 2-1: Planungszeithorizonte in der Energiewirtschaft 13
Abbildung 2-2: Strukturen der Energieversorgungsunternehmen 14
Abbildung 2-3: Kundenwechselquote 2007 18
Abbildung 2-4: Fahrplanmanagement zwischen Händler und Übertragungsnetzbetreiber 20
Abbildung 2-5: Windenergie-Einspeisung in Deutschland, ¼-h-Leistungswerte der Tagesmaxima und Tagesminima 21
Abbildung 2-6: CIGRE Umfrage: Verfahren zur kurzfristigen Lastprognose 23
Abbildung 2-7: Einführung von Lastprognosesystemen 24
Abbildung 2-8: Verteilung der eingesetzten Lastprofilverfahren 25
Abbildung 2-9: Einsatz von Prognosemodellen 26
Abbildung 2-10: Berücksichtigung exogener Größen zur Prognoseerstellung 27
Abbildung 2-11: Aufbau eines Energiedatenmanagementsystems 29
Abbildung 3-1: Elektrischer Lastgang einer Woche 30
Abbildung 3-2: Autokorrelation des Wochenlastganges 32
Abbildung 3-3: Scientific forecasting 37
Abbildung 3-4: Model construction 38
Abbildung 3-5: Übersicht von Zeitreihenanalyseverfahren 43
Abbildung 3-6: Periodogramm von Wochentagen der elektrischen Last 44
Abbildung 4-1: Schematische Darstellung eines biologischen Neurons 60
Abbildung 4-2: Schematische Darstellung eines künstlichen Neurons 63
Abbildung 4-3: Darstellung eines neuronalen Netzes als gewichteter Graph. 65
Abbildung 4-4: Schichtenmodell 65
Abbildung 4-5: Topologische Klassifikation 66
Abbildung 4-6: n-schichtige Multi-Layer-Perceptron (MLP) 67
Abbildung 4-7: Feedforward-Netz mit drei Neuronenschichten 68
Abbildung 4-8: Jordan-Netz 69
Abbildung 4-9: Elman-Netz 70
Abbildung 4-10: Probleme des Gradientenverfahrens 76
Abbildung 4-11: Struktur des Feedforward-Netzes 77
Abbildung 4-12: Rekurrentes Netz mit zwei Neuronen 80
Abbildung 4-13: Äquivalentes Feedforward-Netz 80
Abbildung 4-14: Netzstruktur einer Selbstorganisierenden Kette und Karte 83
Abbildung 4-15: Distanzfunktionen ($h_{cylinder}$,h_{gauss}, h_{cos}, h_{cone}) 86
Abbildung 5-1: Modell des Zerlegungsansatzes 90
Abbildung 5-2: Autokorrelation der Wochentage eines Jahres 91
Abbildung 5-3: Struktur des Prognosemodells 95
Abbildung 5-4: Lastprognose mit neuronalen Netzen 96
Abbildung 5-5: Entwicklungssystematik 98
Abbildung 5-6: Netzstruktur einer zweidimensionalen selbstorganisierenden Karte 102

Abbildung 5-7: Gewichtsvektoren des Kohonenlayers nach 5, 50, 100 und 500 Iterationsschritten 103
Abbildung 5-8: Entfaltung von Dienstag, Donnerstag, Samstage und Sonntage auf der Kohonenkarte 105
Abbildung 5-9: Schematischer Ablauf der Datenanalyse mit SOM 106
Abbildung 6-1: Lastgang in Wochenrasterdarstellung 110
Abbildung 6-2: Verlauf des linearen und des quadratischen Trends 112
Abbildung 6-3: Tageslastgangverläufe nach Wochentaggruppen 114
Abbildung 6-4: Autokorrelation der Wochentagtypen 116
Abbildung 6-5: Jahreseinfluss in den Werktagszeitreihen 121
Abbildung 6-6: Korrelationsfunktion des Vergleichs zwischen Wetterelementen und elektrischer Last 122
Abbildung 6-7: Korrelation zwischen Last und exogenen Größen an konstanten Zeitpunkten 123
Abbildung 6-8: Outputlayer eines ländlichen Versorgungskollektivs 125
Abbildung 6-9: Outputlayer in städtischem und ländlichem Gebiet 126
Abbildung 6-10: Outputlayer eines städtischen Versorgungskollektivs 126
Abbildung 6-11: Prognose eines Wochentages bereinigt um Feiertage 130
Abbildung 6-12: Prognose der Wochenenden 131
Abbildung 6-13: Ergebnis des AR-Prognosemodells für Wochentage 133
Abbildung 6-14: Vergleich zwischen ARMA- und ARIMA-Prognosemodell 137
Abbildung 6-15: Testergebnisse unterschiedliche Topologien 144
Abbildung 9-1: Temperaturverlauf 173
Abbildung 9-2: Periodogramm Temperaturverlauf 173
Abbildung 9-3: Globalstrahlungsverlauf 174
Abbildung 9-4: Periodogramm Globalstrahlung 174
Abbildung 9-5: Windgeschwindigkeitsverlauf 175
Abbildung 9-6: Periodogramm Windgeschwindigkeit 175
Abbildung 9-7: Luftdruckverlauf 176
Abbildung 9-8: Periodogramm Luftdruck 176

Tabellenverzeichnis

Tabelle 3-1: Klassifizierung von Einflussgrößen auf die elektrische Last 34
Tabelle 6-1: Tagesmerkmale 107
Tabelle 6-2: Übersicht der Wettermess- und Prognosewerte 108
Tabelle 6-3: Empirische Momente der Lastgangzeitreihe.....111
Tabelle 6-4: Koeffizienten des linearen und des quadratischen Trends 112
Tabelle 6-5: Identifizierung der Koeffizienten 113
Tabelle 6-6: MSE- Fehler bei der Rekonstruktion von zwei Jahren.....113
Tabelle 6-7: Korrelationsmatrix der Wochentage 115
Tabelle 6-8: Signifikanzgrenzen zu den Koeffizienten der Korrelationsmatrix Tabelle 6-7 116
Tabelle 6-9: Konfidenzintervallgrenzen zu den Koeffizienten der Korrelationsmatrix ohne Feiertage aus Tabelle 6-7 117
Tabelle 6-10: Mittelwert $(\bar{x})$, Standardabweichung (s) und Varianz (s^2) der Wochentagstypzeitreihen 118
Tabelle 6-11: Mittelwert $(\bar{x})$, Standardabweichung (s) und Varianz (s^2) der Feiertage 118
Tabelle 6-12: Korrelation zwischen den Wetterelementen über einen Zeitraum von drei Jahren.....119
Tabelle 6-13: Korrelationsbetrachtung zwischen Lastgang und Wetterelementen bei einer Signifikanzgrenze (α =0,05) 120
Tabelle 6-14: Korrelationsbetrachtung zwischen Lastgang und Wetterelementen nach Bereinigung der Trend- und Jahreseinflüsse (α =0,05) 121
Tabelle 6-15: Anzahl der Merkmalsvariationen in Abhängigkeit des Kalenderjahres 124
Tabelle 6-16: Klassifizierung der Wochentage in Abhängigkeit unterschiedlicher Netzgebiete 128
Tabelle 6-17: Parametrisierungen des klassischen Modells für Wochentage.....129
Tabelle 6-18: Parametrisierungen des klassischen Modells für Wochentage.....130
Tabelle 6-19: Prognose der Wochenenden 131
Tabelle 6-20: AR -Prognosemodell verschiedener Ordnung für Wochentage.....133
Tabelle 6-21: AR -Prognosemodell für Wochentage 134
Tabelle 6-22: AR -Prognosemodell für Wochenenden 134
Tabelle 6-23: ARMA -Prognosemodell für Wochentage 135
Tabelle 6-24: ARMA -Prognosemodell für Wochenenden 136
Tabelle 6-25: ARIMA -Prognosemodell für Wochentage 137
Tabelle 6-26: MRE - Fehler der entkoppelten Modelle 138
Tabelle 6-27: Gesamtübersicht des relativen mittleren Fehlers aller Modelle 140
Tabelle 6-28: Testergebnisse neuronaler Netze.....142
Tabelle 6-29: Ergebnis der Wetterdatenanalyse 143
Tabelle 6-30: Vergleich Feedforward- und Elman-Netz.....145

Tabelle 6-31: Prognosefehler neuronaler Netze 147
Tabelle 9-1: Ergebnisse des Mittelwertstationaritäts- und des Varianzstationaritätstests 172
Tabelle 9-2: Ergebnisse Temperatur 173
Tabelle 9-3: Ergebnisse Globalstrahlung 174
Tabelle 9-4: Ergebnisse Windgeschwindigkeit 175
Tabelle 9-5: Ergebnisse Luftdruck 176

Verzeichnis der verwendeten Formelzeichen und Abkürzungen

a_i	Ausgang des Neurons i
a_{vor}	Ausgang des Vorgängerneurons
$AE_{(t)}$	absoluter Prognosefehler
$APE_{(t)}$	absoluter prozentualer Fehler
AR(p)	AR-Prozess (Auto Regressive) der Ordnung p
ARIMA(p,d,q)	ARIMA-Prozess (Auto Regressive Integrated Moving Average)
ARMA(p,q)	ARMA-Prozess (Auto Regressive Moving Average)
AXARMA(p,q)	AXARMA-ARMA-Prozess mit exogenen Größen (Approximated eXogenous input AutoRegressive Moving Average)
BHKW	Blockheizkraftwerk
c	Kovarianz
c_τ	Autokovarianz
d	Anzahl der Prozessparameter
dll	Dynamic Link Library
DM	Data Mining
DWD	Deutscher Wetterdienst
E_{det}	deterministischer Fehleranteil
$e_{n,i}$	n-ter Eingang des Neurons i
E_{stoch}	stochastischer Fehleranteil
e_t	Rauschen (White Noise)
EVU	Energieversorgungsunternehmen
FFT	Fast-Fourier-Transformation
FTP	File-Transfer-Protocol
g_t	glatte Komponente
H_0, H_1	Hypothese
k_t	Konjunkturkomponente
KDD	Knowledge Discovery in Databases
KNN	künstlich neuronale Netze
m_t	Trendkomponente
MA(q)	MA-Prozess (Moving Average) der Ordnung q
MAE	mittlerer absoluter Fehler (mean absolute error)
MAPE	mittlerer absoluter prozentualer Fehler (mean absolute percentage error)
ME	mittlerer Prognosefehler (mean error)
MLP	Multilayer Perzeptron
MRE	mittlerer relativer Prognosefehler (mean relative error)
MSE	mittlerer quadratischer Fehler (mean square error)
net_i	Summe der Eingänge e_n und des Biaswertes des Neurons i
ODBC	Datenbanktreiber: Offene Datenbank-Verbindungsfähigkeit (Open Database Connectivity)
p	Anzahl der Prozessparameter
P_{prog}	prognostizierte Last
P_{tats}	tatsächliche Last
PAR(p)	Entkoppelter AR-Prozess (Parallel Auto Regressive)
PARX(p)	Entkoppelter AR-Prozess mit exogene Größen (Parallel Auto Regressive eXogenous input)

q	Anzahl der Prozessparameter
$Q(p)$	Güte der Anpassung
r	Korrelation
r_k	Autokorrelationskoeffizient
r_t	Restkomponente
$RAE_{(t)}$	relativer absoluter Prognosefehler
$RE_{(t)}$	relativer Prognosefehler
RMSE	Wurzel des mittleren quadratischen Fehlers (root mean square error)
s	Standardabweichung
s^2	empirische Varianz
s_t	Saisonkomponente
SOM	Self-Organizing-Feature-Maps
VBA	Visual Basic for Applications
w_{ij}	Gewicht zwischen Neuron i und Neuron j
$WE_{(t)}$	wahrer Prognosefehler
$\overline{x}$	arithmetischer Mittelwert
X_t	Zufallsvariable, Prozessvariable
z_t	zyklische Komponente
α	Koeffizient AR-Prozess
β	Koeffizient MA-Prozess
σ^2	Varianz
ω	Ereignis

Kurzfassung

Um die Kosten zur Deckung der Nachfrage nach leitungsgebundenen Energieträgern zu minimieren, wird in Energieversorgungs- und Handelsunternehmen die Energiebereitstellung detailliert geplant. Durch die Liberalisierung der Energiemärkte und die hieraus steigende Anzahl an Marktakteuren und Transaktionen ist es nun erforderlich, für Vertriebs- und Handelsunternehmen sowie für die Stromhandelsaufgaben des Netzbetreibers getrennte Prognosen mit zum Teil völlig unterschiedlichen Anforderungen an Methoden und Systemen zu erstellen. Diese Planungsaufgabe wird künftig zunehmend wichtiger und ist für den wirtschaftlichen Erfolg der Unternehmen maßgeblich. In der betrieblichen Praxis wird eine Vielzahl verschiedener Ansätze und Verfahren zur Lastprognose eingesetzt. Dabei ist derzeit das Verfahren der multiplen Regression am weitesten verbreitet.

Im Rahmen dieser Arbeit wurden auf der Basis von Lastgangdaten unterschiedlicher Versorgungskollektive Prognosemodelle erarbeitet. Es erfolgten zunächst statistische Untersuchungen zwischen den Lastgangdaten und den zugehörigen exogenen Größen. Mit den hieraus gewonnenen Erkenntnissen wurde die Entwicklung von Lastprognosemodellen mit Methoden der klassischen Zeitreihenanalysen, den Methoden der stochastischen Prozesse sowie einer Kombination von deterministischen und stochastischen Ansätzen durchgeführt. Weiterführend wurde die Technik der neuronalen Netze zur Lastprognose eingesetzt, da diese eine hohe Lernfähigkeit aufweisen und auch für andere Anwendungsgebiete leicht trainierbar sind. Die Art und Topologie sowie Anzahl der neuronalen Netze wurde in Abhängigkeit der Prognosezuverlässigkeit optimiert und anhand von Fallstudien erprobt.

Schlagworte - Energiemarkt, elektrische Lastprognose, neuronale Netze, stochastische Prozesse

Abstract

Energy supply companies and trading firms plan energy supply in detail in order to minimise the costs of meeting the demand for grid-bound energy carriers. The increasing number of market actors and transactions caused by the liberalisation of energy markets makes it necessary to establish demand forecasts separately for energy distribution and trading firms on one hand and grid operators with energy trading tasks on the other hand. The forecasts for both kinds of market actors require quite different methods and systems.

The task to forecast is getting more and more important and is determinant for the economic success of the companies. A broad range of different approaches and methods is used in the commercial practice. Among them, the approach of multiple regressions is presently the most widely used. In the framework of the work presented here, forecast models were elaborated on the basis of load curves of different supply communities. First, a statistical analysis of load curves and related exogenous parameters has been made. On the basis of the insights gained from that analysis, load forecast models have been developed with (1) the methods of classical time series analysis, (2) the methods of stochastic processes, and (3) a combination of deterministic and stochastic approaches. Furthermore, the technique of neural networks has been applied for establishing load forecasts because they show a high compliance and can be trained easily for new areas of application. The kind and topology as well as the number of neural networks have been optimised depending on the forecast reliability and have been tested in case studies.

Keywords - Energy market, electric load forecasting, artificial neuronal networks, stochastic process.

1 Einleitung

In allen Bereichen des Lebens haben Prognosen die Funktion, für Planungsfragen Handlungsanleitungen beziehungsweise Entscheidungsgrundlagen zu liefern. Die Prognose bildet damit eine Grundlage für künftige Entwicklungen und beeinflusst somit das Planungsergebnis und die daraus resultierende Entscheidung maßgeblich. Unter Berücksichtigung eines durch die Entscheidung möglichen Schadens werden Eintrittswahrscheinlichkeiten eines Ereignisses ermittelt und das aus einem Schaden resultierende Risiko minimiert. Damit fallen auch Risikoabschätzungen ebenfalls in den Bereich von Prognosen.

Die Energiewirtschaft nimmt in unserem Lande eine Schlüsselstellung ein. Die durch sie erbrachten Leistungen ermöglichen den Betrieb und Ausbau von Energieinfrastruktur, die für die wirtschaftliche Weiterentwicklung eines Landes unverzichtbar und zum Teil unersetzbar sind. Aufgrund der großen Investitionen und langen Einsatzzeiten von Betriebsmitteln stellt die Prognose der langfristigen Energienachfrage die Basis aller Planungsrechnungen dar. Um die Kosten zur Deckung der Nachfrage nach leitungsgebundenen Energieträgern zu minimieren, wird in Energieversorgungs- und Handelsunternehmen die Energiebereitstellung und auch -erzeugung detailliert mit kurzfristigen Prognosen geplant.

Da leitungsgebundene Energieträger in relevantem großtechnischem Umfang derzeit praktisch nicht speicherbar sind, muss der Leistungsbedarf aller Verbraucher zeitgleich unter Einhaltung aller technischen und wirtschaftlichen Rahmenbedingungen durch eine Erzeugung gedeckt werden. Die Prognose leitungsgebundener Energieträger wie elektrische Energie, Erdgas und Fernwärme ist somit eine wesentliche Voraussetzung für das Energiemanagement, d. h. für die Planung der Bedarfsdeckung nach technischen und wirtschaftlichen Gesichtspunkten. Qualitativ hochwertige Prognosen stehen damit in engem Zusammenhang mit der Güte der wirtschaftlichen Optimierung und folglich mit dem erzielten Geschäftsergebnis.

Der in der Vergangenheit monopolistische Markt stellte im Vergleich zu heute gänzlich andere Anforderungen an die Qualität von Lastprognosen, da sich diese im Regelfall auf die Vorhersage der Summenlast aller im angestammten Netzgebiet befindlicher Kunden beschränkte. Schwerpunkt dieser Planungsaufgaben war dabei die Kraftwerkseinsatzplanung, die eine Minimierung der Betriebskosten zum Erreichen des optimalen Einsatzes eines vorhandenen Kraftwerksparks zum Ziel hatte. Diese sehr komplexe Optimierungsaufgabe musste - von der momentanen Lastaufteilung im Minutenbereich auf die in Betrieb befindlichen Kraftwerksblöcke oder schnell aktivierbaren Gasturbinen- und Pumpspeicherkraftwerke über die Blockeinsatzplanung, die das An- und Abfahren von thermischen Kraftwerken berücksichtigt, bis hin zur Energieeinsatzplanung und Brennstoffeinsatzplanung - einen Zeitraum von einigen Minuten bis hin zu mehreren Jahren berücksichtigen und analysieren.

Mit der Einführung der Liberalisierung und Regulierung der Märkte in Europa haben sich die Rahmenbedingungen der Energieversorgung gänzlich verändert. Ausgelöst wurden sie durch die EU-Binnenmarktrichtlinie Elektrizität, die eine stufenweise Marktöffnung ermöglichte. Großabnehmer und Privatkunden können bereits seit Mitte der 90er Jahre den Stromlieferanten frei wählen. Stromhändler bieten bundesweit an, Strombörsen wurden aufgebaut und die Netze für freie Nutzung zunächst auf Basis der Entgeltberechnung der Verbändevereinbarung (VVII Plus) und dann auf Grundlage der Stromnetzzugangsverordnung geöffnet. In den traditionellen Versorgungsstrukturen waren Erzeugung, Übertragung und Verteilung, Handel und Vertrieb zusammengefasst. Das vom Gesetzgeber geforderte "Unbundling" (§ 6 EnWG), also die informatorische und gegebenenfalls sogar gesellschaftsrechtliche Trennung von Erzeugung, Übertragung, Verteilung und Handel, hat auch eine wesentliche Voraussetzung für freie Lieferbeziehungen geschaffen. Zusätzlich wurden mit der Einführung des Erneuerbare-Energien-Gesetzes (EEG) und des Kraft-Wärme-Kopplungs-Gesetzes (KWK-G) weite, gänzlich neue Anwendungsbereiche für Prognosen erforderlich. Aufgrund des ständig wachsenden Anteils regenerativ, dezentral erzeugter Energie, im Besonderen beim Zubau von Photovoltaik- und Windkraftanlagen, ist die Prognose dieser Energiemengen von zunehmender Bedeutung. Die räumliche Verteilung der Anlagen und die Fluktuation der erzeugten Energie stellen besondere Anforderungen an bestehende Energieversorgungssysteme. Eine Prognose der zu erwartenden Leistung ermöglicht eine effiziente und wirtschaftliche Steuerung konventioneller Kraftwerke.

Die Anforderungen an die bis dahin klassischen Prognoseverfahren sowie die Zuständigkeit und Verantwortlichkeit der Prognoseerstellung haben sich mit dem Fortschreiten der Liberalisierung verändert. Die im heutigen Markt agierenden Handels- und Vertriebsgesellschaften prognostizieren für ihre Kunden den gesamten Energiebedarf zur Beschaffungsoptimierung und Lastprognosen erfolgen zwecks Fahrplanmanagements für Kunden innerhalb einer Regelzone des Übertragungsnetzbetreibers. Die Einführung der Standardlastprofile stellt hierbei einen wichtigen Schritt auf dem Weg zum liberalisierten Strommarkt dar und ermöglicht einen einfachen Netzzugang für Haushalts- und kleine Gewerbekunden. Die Anwendungsgrenze dieser Normprofile liegt bei einem Jahresverbrauch von 100.000 kWh. Die Erstellung von Lastprognosen für diese Kundengruppe ist im klassischen Sinne damit überflüssig, gleichzeitig steigen die Anforderungen an die Prognosequalität für Kunden- und Kundengruppen ohne Normprofile und Großkunden. Durch die Stromnetzzugangsverordnung (StromNZV) wird auch der Netzbetreiber zur Führung von Bilanzkreisen verpflichtet. Eine Bewirtschaftung der Bilanzkreise erfordert Stromhandelsaufgaben für die Beschaffung von Mehr- und Mindermengen sowie Regel- und Verlustenergie. Im Hinblick auf die Prognose bedeutet dieses, dass bereits jetzt, und in verstärktem Maße künftig, eine Vielzahl unterschiedlicher Anforderungen an Netzbetreiber gestellt wird.

Mit fortschreitender Entwicklung der Strommärkte in Europa und der steigenden Anzahl an Transaktionen und Akteuren wird der Wettbewerbsdruck in der Energiebranche

zunehmend größer. Die Anforderungen an hochwertige Lastprognosemodelle sowie die hierfür eingesetzten Instrumente werden einen großen Stellenwert einnehmen und mit über den wirtschaftlichen Erfolg eines Unternehmens entscheiden.

1.1 Stand des Wissens

Eine aktuelle Übersicht zu den Lastprognosemethoden im deregulierten Markt wird im Artikel von Hahn, Meyer-Nieberg und Pickl, gegeben [Hah09]. Sie gliedern Lastprognosemethoden in klassische Zeitreihen- und Regressionsverfahren sowie Methoden der künstlichen Intelligenz.

Die Technik der neuronalen Netze ist innerhalb von Prognosesystemen für betriebs- und volkswirtschaftliche Größen bekannt. Es existieren effiziente Lernalgorithmen zur Bestimmung der Gewichtsparameter innerhalb der Systeme. Theoretische Untersuchungen belegen, dass schon ein einfach strukturiertes künstliches neuronales Netz (Multi-Layer-Perzeptron) beliebige, nicht lineare, stetige Zusammenhänge beliebig genau approximieren kann, wenn es nur hinreichend groß gewählt wird. Darüber hinaus wurde der Nachweis erbracht, dass insbesondere bei Verwendung sigmoidaler Funktionen als neuronale Transferfunktionen die künstlichen neuronalen Netze der klassischen multiplen Regression überlegen sind. All dies rechtfertigt den Einsatz der Technik neuronaler Netze auch in der Lastprognoseproblematik. Aus der zahlreichen Literatur zu diesem Thema neuronale Netze seien an dieser Stelle nur folgende genannt [Bis97], [Rip97], [Roj96], [Bär92].

Der Wissensstand bezüglich der Lastprognose mit neuronalen Netzen wird in dem Artikel von Hufendieck und Kaltschmitt vom Institut für Energiewirtschaft und rationelle Energieanwendung der Universität Stuttgart beschrieben [Huf99]. Sie weisen auf die Komplexität und die noch vorhandenen Probleme bei der Netzauswahl hin und darauf, dass die Entwicklung solcher Systeme bisher weitgehend auf Versuch und Irrtum beruht. Hufendieck stellt eine erste, aber noch sehr grobe und allgemeine Entwicklungsmethodik vor, die jedoch im Detail noch auszugestalten und an die spezielle Problematik anzupassen ist. Die Detaillierung dieser erfolgt im Rahmen der Veröffentlichung „Systematische Entwicklung von Lastprognosesystemen auf der Basis neuronaler Netze“ [Huf00].

AL-Sakaf empfiehlt in seiner Promotion [Als95], Daten der Wochentage Dienstag bis Donnerstag zusammenzufassen und für die anderen Wochentage einzelne weitere neuronale Netze zu entwickeln. Auf Grund von Konvergenzproblemen und langen Wartezeiten gibt er den selbstorganisierenden Kohonennetzen den Vorzug gegenüber dem überwachten Trainieren eines Multi-Layer-Perzeptron- Netzes, trainiert aber dieses nur mit einem reinen Backpropagation-Algorithmus ohne nennenswerte Modifikation. Der Backpropagation-Algorithmus ist ein Gradientenverfahren zur Fehlerminimierung

und besitzt dadurch das Problem lokaler Minima. Hier gibt es erhebliches Verbesserungspotenzial hinsichtlich Genauigkeit und Leistungsfähigkeit.

Eine allgemeinere Untersuchung zum Einsatz neuronaler Netze bei Zeitreihen liefert Thiesing [Thi98]. Er analysiert Zeitreihen von Verkaufszahlen und stellt den Vorzug neuronaler Netze gegenüber autoregressiven und Moving-Average Modellen der statistischen Zeitreihenanalyse dar. Zudem stellt er implementierte Erweiterungen des Stuttgart Neural Network Simulator (SNNS) sowie Parallelisierungen des Backpropagation-Algorithmus zur Beschleunigung vor.

Das Potenzial der neuronalen Netze hinsichtlich der Prognose des Gasverbrauchs zeigen auch Diplomarbeiten auf, in denen mit Hilfe neuronaler Netze ein erstes zufriedenstellendes Prognosemodell für den Gasverbrauch vorgestellt wird [Bra02]. Die in diesem Projekt zu prognostizierenden Stromverbrauchsdaten sind jedoch deutlich komplexer und verlangen ein tieferes methodisches Vorgehen.

Im Beitrag von Senjyu, Takara, Uezato und Funabashi werden sehr kurzfristige (1-6 h) Lastprognosen mit Hilfe neuronaler Netze erstellt [Sen02]. In üblichen Modellen erfolgt eine 24-stündliche Prognose der Last unter Verwendung prognostizierter exogener Größen wie beispielsweise der Temperatur des Prognosetages. Aufgrund von Abweichungen vom tatsächlichen zum prognostizierten Temperaturverlauf resultieren hieraus Lastprognosefehler. Im Lösungsansatz der Autoren werden Netztopologien mit einer Zwischenschicht verwendet, denen zu Trainingszwecken mit dem Backpropagation-Algorithmus dreißig Tage als Vortageswert und sechzig Tageswerte des Vorjahres präsentiert werden. Des Weiteren erfolgt in diesem Ansatz eine Rückkopplung des Prognosewertes auf den Eingang des Netzes und eine Korrektur des Ausgangsprognosewertes um Vergleichstage. Mit diesem Ansatz wird ausschließlich auf Basis historischer Daten ein mittlerer absoluter prozentualer Fehler des folgenden stündlichen Lastwertes von < 2 % erzielt. Eine Erweiterung dieses Modellansatzes um eine Preisvorhersage erfolgt in der Veröffentlichung von Mandal, Senjyu, Uezato und Funabashi [Man05]. Mit der Preisvorhersagemethode wird ein um < 10% geringerer mittlerer absoluter prozentualer Fehler in Abhängigkeit der Jahreszeit ermittelt im Vergleich zu Standard-Preisprognosen.

Im Beitrag von Al-Shareef, Mohamed und Al-Judaibi wird die Entwicklung eines KNN-basierten Modells für sehr kurzfristige Lastprognosen beschrieben. Im Rahmen der Untersuchungen werden einfache Back-Propagation-Netzwerke unter Berücksichtigung einer exogenen Größe verwendet. Im Ergebnis empfehlen die Autoren zur Verbesserung der Modelle eine Einbeziehung weiterer Wettergrößen sowie den Einfluss der Merkmale von Sondertagen zu berücksichtigen [Als08].

Die Veröffentlichung von Hayati beschreibt den Aufbau eines Prognosesystems mit unterschiedlichen Netztopologien [Hay07]. Im Rahmen der Arbeit werden Unter-

suchungen mit Mulit-Layer-Perceptron (MLP), rekurrenten Elman-Netzen und Radiale-Basis-Funktionen (RBF)-Netzen durchgeführt. Neben historischen Lastgangdaten wird den Netzen zu Trainings- und Prognosezwecken eine Vielzahl exogener Einflussgrößen präsentiert. Hayati verwendet grundsätzlich Netztopologien mit einer Zwischenschicht, die mit elf bis neunzehn Neuronen ausgestattet ist. Als Aktivierungsfunktionen kommen sigmiode Funktionen zum Einsatz. Minimale Prognosefehler werden hierbei mit Netztypen der Radial-Basis-Funktion erzielt.

In einem Aufsatz von Sideratos und Hatziargyriou [Sid06] wird der Ansatz eines kombinierten Fuzzy-Logik-Modells mit neuronalen Netzen aufgenommen und zur Prognose der Einspeisung eines Windparks eingesetzt. Dabei wird zunächst mit dem Fuzzy-Logik-Modell die Zuverlässigkeit und Qualität der Wettervorhersage geprüft und in einem zweiten Schritt mit einem neuronalen Netz (RBF) korrigiert.

Mit dem Ziel, eine verbesserte Leistungsvorhersage photovoltaischer Einspeiseanlagen durchzuführen, wurden in den zurückliegenden Jahren Anstrengungen unternommen, auf Basis von Globalstrahlungs- und Temperaturvorhersagen die Regelenergieaufwendungen zu minimieren. In den Veröffentlichungen von Bofinger [Bof06], [Bof07] und Lorenz [Lor07] wird der Vorhersagefehler minimiert, indem die räumlich verteilten Anlagenensembles zum Prognosemodell der Temperatur und Strahlung optimiert werden.

Im Beitrag von Amjady erfolgt die Modellierung der kurzfristigen Lastprognose von Zeitreihen mit einem modifizierten ARIMA-Prozess (Auto-Regressive-Moving-Average) [Amj01]. Dabei wird im ersten Schritt die Expertenschätzung verwendet, um eine erste Prognose zu erstellen. In einem zweiten Schritt wird diese Prognose mit der Temperatur und den aktuellen Lastgangdaten mit Hilfe eines multivariaten Regressionsverfahrens verbessert. Mit der von Amjady beschriebenen Methode werden genauere Ergebnisse als mit den herkömmlichen Techniken, wie künstliche neuronale Netze oder dem Box-Jenkins-Modell, erzielt.

1.2 Zielsetzung und Gliederung der Arbeit

In der betrieblichen Praxis werden heute meist manuelle oder empirische Prognoseverfahren eingesetzt. Dabei macht man sich eine Einteilung der auftretenden Zeitreihen in Trendkomponente, zyklische Komponente und Rest- bzw. Zufallskomponente zu Nutze. In den letzten Jahren hat sich die Technik neuronaler Netze als robustes, multivariates Prognoseverfahren auch bei Zeitreihenproblemen etabliert. Neuronale Netze sind mehr oder weniger inspiriert durch biologische Vorbilder wie menschliche Problemlösungsstrategien oder die neuronale Struktur des Gehirns. Erste vielversprechende Ansätze zur Anwendung neuronaler Netze auch im Rahmen von Lastprognosen sind bereits vorhanden. Doch existiert bei der Vielzahl der neuronalen

Strukturen und Datenrelationen noch kein prinzipielles Gesamtkonzept für die Vorgehensweise zur Lösung dieser Prognoseproblematik.

Ziel dieser Arbeit ist es, auf der Basis vorliegender Lastgangdaten unterschiedlichster Abnahmekollektive ein verallgemeinerungsfähiges innovatives Prognosemodell zu erarbeiten. Die hierfür zu entwickelnden Prognosealgorithmen sollen neben den klassischen Zeitreihenanalysen auch Methoden der stochastischen Prozesse beinhalten. Der Technik der neuronalen Netze soll besondere Bedeutung zukommen, da diese eine hohe Lernfähigkeit aufweisen und für andere Anwendungsgebiete auch leicht trainierbar sind. Die Art und Topologie sowie Anzahl der neuronalen Netze soll in Abhängigkeit der Prognosezuverlässigkeit optimiert werden. Ziel ist die Erarbeitung eines Prognosemodelles mit einer Zuverlässigkeit von > 95 %.

Ausgehend von der Bereitstellung von Energiedienstleistungen ist es erforderlich, den Anwendungsbereich elektrischer Energie näher zu analysieren. Dabei ist festzustellen, dass diese das breiteste Anwendungsspektrum aufweist. Die im Rahmen dieser Arbeit angestellten Überlegungen sollen dann für andere Energieträger analog abgeleitet werden können.

Die innerhalb des AGIP-Projektes (Arbeitsgemeinschaft Innovative Projekte beim Ministerium für Wissenschaft und Kultur des Landes Niedersachsen) E4T in den Jahren 2002 bis 2005 gewonnenen Erkenntnisse und betreuten Diplomarbeiten sind in diese Arbeit eingeflossen. Die Arbeit gliedert sich in sechs Abschnitte. Im folgenden zweiten Kapitel erfolgt eine Darstellung der veränderten energiewirtschaftlichen Rahmenbedingungen und den hieraus resultierenden Anforderungen - in Abhängigkeit der jeweiligen Marktrolle - an die Prognoseerstellung. Um eine aktualisierte Übersicht der zurzeit am Markt eingesetzten Methoden und Modelle sowie anderer für die Prognoseerstellung relevanter Einflussgrößen zu erhalten, wurde im Rahmen dieser Arbeit eine Befragung deutscher Energieversorgungsunternehmen durchgeführt. Die Ergebnisse der Befragung werden in diesem Kapitel dargestellt und interpretiert.

Im dritten Kapitel werden die methodischen Grundlagen der Zeitreihenanalyse und das Problem der Lastprognose vorgestellt. Dabei gliedert sich dieses Kapitel in die Themenbereiche der klassischen Zeitreihenanalyse und den Abschnitt der stochastischen Prozesse.

Den Schwerpunkt des vierten Kapitels bilden die methodischen Grundlagen neuronaler Netze. Die Funktionen dieser selbstlernenden Algorithmen sowie deren Arbeitsweise werden dargestellt. Dabei wird neben unterschiedlichen Netzarchitekturen und Lernalgorithmen vorwärtsgerichteter und rekurrenter Netzstrukturen auch ausgiebig auf die Netztypen der selbstorganisierenden Karten eingegangen.
Kapitel fünf beschreibt die Entwicklungsmethodik von Lastprognosemodellen. Dabei erfolgt zunächst eine Modellierung und Anpassung eines stochastischen Prozesses zur

Lastprognose. Des Weiteren werden Prognosemodelle auf Basis neuronaler Netze unterschiedlicher Netztypen und deren systematischer Einbindung in Prognosesysteme entwickelt. Zusätzlich wird ein entkoppeltes Modell mit der Einbindung exogener Größen vorgestellt. Abschließend wird mit Hilfe selbstorganisierender Karten (SOM - Self-Organizing-Maps) ein Verfahren entwickelt, aus Lastdatensätzen vergangener Jahre Unregelmäßigkeiten im Lastgang selbst zu detektieren und Datenclusterungen innerhalb der Lastgangdaten zu bilden, um so eine Reduktion diskontinuierlicher Werte (Merkmale eines Tages) zu erhalten.

Im sechsten Kapitel erfolgen anhand von Fallstudien verschiedener Abnahmekollektive und zugehöriger exogener Größen grundlegende statistische Analysen und es werden die im vorherigen Kapitel erarbeiteten Modelle auf ihre Anwendbarkeit überprüft. Hierbei werden Prognosemodelle durch Kombination von deterministischen und stochastischen Ansätzen getestet und diese mit Prognosen der Methode neuronaler Netze unterschiedlicher Netz- und Lerntypen verglichen. Das letzte Kapitel fasst abschließend die wesentlichen Ergebnisse der gesamten Arbeit zusammen und gibt im zweiten Abschnitt einen Ausblick auf denkbare Ansätze für weiterführende wissenschaftliche Untersuchungen und Weiterentwicklungen. Die Arbeit endet mit dem Literatur-verzeichnis und einem Anhang.

2 Energiewirtschaftliche Rahmenbedingungen

2.1 Prognosemodelle und -methoden

Ein Energieverteilungssystem kann als eine Gruppe von unterschiedlichen Objekten oder Teilen, die mit Beziehungen zueinander zusammenwirken, beschrieben werden. Die Bildung eines Modells zur Beschreibung eines solchen Systems stellt in allen Bereichen der wissenschaftlichen Forschung die Grundlage einer systematischen Untersuchung von Zusammenhängen zwischen Einflussgrößen und interessierenden Größen dar. Modelle werden somit gebildet, um bestimmte Merkmale der Realität aufzuzeigen und sind eine idealisierte Darstellung der Wirklichkeit. Ziel der Modellbildung ist, eine vereinfachte Abbildung der wesentlichen Zusammenhänge zwischen Einflussgrößen und zu untersuchenden Größen durch Reduzierung der Einflussgrößen auf eine endliche Anzahl darzustellen.

Bei der Untersuchung von Phänomenen durch Beobachtung und Messung werden empirische Daten gesammelt mit dem Ziel, in einem ersten Schritt in den Daten verborgene Ordnungen, Regeln und Gesetze zu entdecken. Die dann folgende Beschreibung der ermittelten Gesetzmäßigkeit kann in Form von Algorithmen erfolgen und als deskriptive Modellierung bezeichnet werden. Ist ein deskriptives Modell ermittelt, so gilt es in einem zweiten Schritt, die empirisch und algorithmisch ermittelte Regularität mit den Gesetzmäßigkeiten von Mathematik, Physik, Chemie und Biologie zu verknüpfen und sie im traditionellen Sinn durch ein kausales Modell diskutierbar oder "erklärbar" zu machen. Die Kausalität zwischen Einflussgröße und zu untersuchender Größe ist Voraussetzung einer Prognose, die Aussagen über eine künftige Entwicklung ermöglicht. In Abhängigkeit vom Ziel der Modellierung umfasst der Weg der Modellbildung im Wesentlichen drei Schritte:

- Phänomenologie: Beobachtungen und Sammlung von Daten in der Realität und Suche nach Ordnungsprinzipien, Regeln und Gesetzen.
- Deskriptives Modell: Umsetzung der Ordnungsprinzipien in einen Kompaktalgorithmus zur Beschreibung der Phänomene (Daten).
- Kausales Modell: Verknüpfung der Algorithmen bzw. Ordnungsprinzipien mit naturwissenschaftlichen Ursachen zur Erklärung sowie Deutung und Diskussion der Phänomene (Daten).
- Prognosemodell: Ermöglichen von Vorhersagen des zukünftigen Verhaltens einer interessanten Größe eines Systems in Abhängigkeit Ihrer Einflussgrößen.

Im Rahmen der vorliegenden Arbeit wird der Schwerpunkt der Untersuchungen in der Entwicklung von Prognosemodellen zur Beschreibung der elektrischen Last liegen. Die Methoden zur Prognose und speziell der elektrischen Last werden in der Literatur [Cha82], [Gro87], [Bun85], [Als95] in vier Hauptkategorien klassifiziert. Die subjektiven oder manuell empirischen Prognoseverfahren, die aufgrund subjektiver Eindrücke er-

stellt werden. Hierbei fließen zur Prognoseerstellung Urteile, Intuition und andere reale Informationen ein. Die Prognosen der elektrischen Last erfolgen unter Zuhilfenahme von Ähnlichkeiten in vergangenen Jahres- und Tageswerten sowie durch Kombination von typischen Wetterbedingungen an diesen Tagen. Die Prognose erfolgt hierbei empirisch [Cha82].

Univariate Verfahren basieren auf Daten der Vergangenheit sowie Anpassung eines Modells und der Methode der Extrapolation. Vorhersagen werden aus der historischen Lastganglinie selbst erzeugt, sodass nur die Zeitabhängigkeiten Berücksichtigung finden. Dieser Ansatz wird bei Systemen mit geringer Abhängigkeit von exogenen Einflussgrößen der Last angewendet oder dann, wenn keine oder unzureichende Einflussgrößen aus der Historie bekannt sind oder Prognosen für sehr kurze Zeithorizonte zu erstellen sind [Pet86].

Mulitvariate Prognoseverfahren berücksichtigen bei einer Modellierung die kausalen Zusammenhänge exogener Einflussgrößen auf den zu prognostizierenden Wert. Voraussetzung für die Anwendung der Verfahren ist die Vorhersagbarkeit exogener Einflussgrößen. Regressionsmodelle und auch Mustererkennungsmodelle fallen in diese Verfahrensgruppe.

Die Verfahren, bei denen uni- und multivariate wie auch subjektive Verfahren innerhalb eines Prognosemodells Verwendung finden, bezeichnet man als Kombinationsverfahren. Hierbei können Vorhersagen, uni- oder multivariate Prognosen errechnet und dann subjektiv korrigiert werden.

Die Auswahl eines Vorhersageverfahrens wird durch die Anforderung der zu erledigenden Aufgabe und deren Randbedingungen bestimmt. Es steht die Fragestellung der Qualität an eine Prognose im Vordergrund, die jedoch in einem ausgewogenen Verhältnis zwischen Aufwand und Nutzen stehen muss. Zusätzlich nehmen Prognosehorizont und -intervall sowie das Vorhandensein historischer Daten und exogener Einflussfaktoren maßgeblichen Einfluss. Der Verfahrenseinsatz sollte individuell der Aufgabe und dem Zweck entsprechend abgewogen werden, da eindeutige allgemeingültige Regeln sich nicht ableiten lassen.

2.2 Planungshorizonte in der Energiewirtschaft

Die Erfordernisse der Praxis führen zu unterschiedlichen Planungsaufgaben in der Energiewirtschaft. Diese Planungsaufgaben können dabei eine Optimierung in verschiedene Zeitbereiche zum Ziel haben. Dabei können unterschiedliche Zeithorizonte klassifiziert werden, die von einem Betrachtungszeitraum von einigen Tagen bis hin zu Jahrzehnten reichen. Da Lastprognosen die Grundlage aller Planungsaufgaben darstellen, ist daher die Einteilung der Prognosehorizonte entsprechend der sich aus

Planungsaufgabenstellungen ergebenden Zeitbereiche sinnvoll. Um eine geeignete Strukturierung herbeizuführen, werden fünf Aufgabenblöcke definiert [Lin96], [Huf00] und [Als95].

- Die prägenden Kriterien der strategischen Langzeitplanung leiten sich aus den Unternehmenszielen und der Strategie ab und liefern Ergebnisse zum Primärenergieeinsatz, zur Finanz- und Kostenplanung, der Vertragsgestaltung und zur Standort- und Stilllegungsplanung. Das Optimierungsintervall dieser Planung erstreckt sich über einen Zeitraum von fünf bis hin zu zwanzig Jahren. Die Unsicherheiten der Prognosedaten rechtfertigen hierbei die Wahl vereinfachter Verfahren und die Reduzierung auf eine geringe Anzahl von Eingangsdaten. Die hierfür benötigten Lastprognosen beschränken sich daher auf die Jahresdauerlinien in Prognoseintervallen von einem bis hin zu zehn Jahren [Edv90].

- Langfristige Planungen erstrecken sich üblicherweise über einen Zeitbereich zwischen einem und fünf Jahren, mit dem Ziel, Bezugs- und Lieferverträge sowie die zulässigen Regelbereiche der Kraftwerksblöcke einzuhalten. Die Planungen sollen dabei die Jahreshöchstlast und Jahresenergiemengen sowie deren Aufteilungen auf die jeweiligen Monate in Form von Lastganglinien eines Jahres liefern. Die Ergebnisse dieser Optimierung sind Kostendarstellungen, Kraftwerksfahrpläne mit Energie- und Leistungsmengen, Emissionsbelastungen und Reststoffmengenprognosen sowie die Bewertung des statistischen Versorgungsrisikos [Edv90].

- Die mittelfristigen Planungen dienen der Revisionsplanung, der Brennstoffeinsatzplanung sowie damit verbundener Bevorratung und Speicherbewirtschaftung. Zusätzlich nehmen diese maßgeblichen Einfluss auf den Abschluss energie- und stromwirtschaftlicher Vereinbarungen. Die Prognosehorizonte erstrecken sich von einer Woche bis hin zum einem Jahr. Teile der benötigten Eingangsparameter der Planungsaufgabe werden durch die strategischen und langfristigen Planungsaufgaben vorgegeben. Das Planungsergebnis sind Lastganglinien typischer Tage [Edv90], [Huf00].

- Die kurzfristigen Planungsaufgaben der Energiewirtschaft erstrecken sich über einen Zeitraum von einem Tag bis hin zu einer Woche. Die Planung liegt hierbei im Schwerpunkt auf der Erstellung von Lastganglinien, die in Abhängigkeit der Bezugs- und Liefervereinbarungen erfolgen mit einer Auflösung in Zeitintervallen, die zwischen einer viertel und einer Stunde liegen können. Die kurzfristige Planung dient zur Festlegung der Regelreserven des Übertragungsnetzes und zur Bezugs- und Erzeugungsplanung und ist abhängig von der Rolle des Unternehmens im Markt. Die Anforderung an die Güte und Genauigkeit der Prognose sind aufgrund des hohen finanziellen Risikos mit größter Genauigkeit festzulegen.

- Sehr kurzfristige Planungsaufgaben erstrecken sich über einen Betrachtungszeitraum von Sekunden bis hin zu Stunden. Die Aufgaben dieser Planung bestehen

primär in der Bereitstellung von Regelenergie wie z. B. der Einsatzplanung von Stunden-, Minuten- oder Sekundenreserven sowie dem Lastaustausch zwischen Übertragungsnetzbetreibern. Die erstellten Prognosen finden zusätzlich Anwendung bei der Früherkennung von Grenzwertüberschreitungen und der Planung kurzfristiger Schaltvorgänge.

Zur Erzielung eines optimalen Ergebnisses kann die zeitliche Einteilung der Planungsaufgaben in Funktionsblöcke nicht isoliert, sondern nur zusammenhängend betrachtet werden, da im Regelfall die Ergebnisse des vorhergehenden Funktionsblocks Einfluss auf den folgenden nehmen. Der Zusammenhang der Funktionsblöcke ist in Abbildung 2-1 dargestellt.

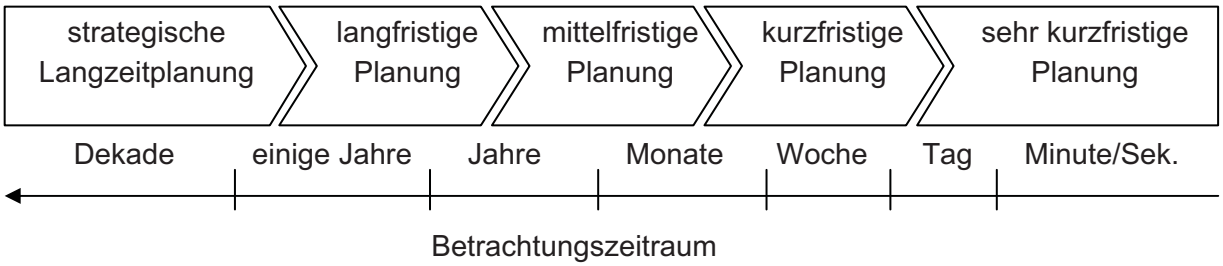


Abbildung 2-1: Planungszeithorizonte in der Energiewirtschaft

Aufgrund der fortschreitenden Veränderung des Marktes und den daraus resultierenden Bedingungen wie auch der energiepolitischen Rahmenbedingungen, beispielsweise der Förderung regenerativ erzeugter Energie oder der Förderung von Kraft-Wärme-Kopplungsanlagen, ergeben sich eine Vielzahl von Planungsunsicherheiten, die zu einer Verringerung des Planungshorizontes führen. Der Fokus operativer Unternehmensbereiche liegt derzeit auf der Erstellung kurzfristiger Lastprognosen, die im Hinblick auf die Modellierung der elektrischen Last die höchsten Anforderungen aufgrund der Vielzahl unterschiedlicher Einflussfaktoren stellen und im Rahmen der vorliegenden Arbeit detaillierter untersucht werden.

2.3 Energiewirtschaftliche Lastprognosen

Die zentrale Aufgabe der Energiewirtschaft ist die Erzeugung und Verteilung elektrischer Energie, die mit Hilfe der in Abbildung 2-2 dargestellten Komponenten erfolgt. Da Energieerzeugung und -wandlung beim Anwender innerhalb eines elektrischen Versorgungssystems zeitgleich erfolgen muss, werden Differenzen zwischen diesen beiden Größen durch Regelenergie oder Energieaustausch mit den umliegenden Verbundnetzbetreibern ausgeglichen, die damit oftmals zu ungewollter zusätzlicher Belastung des Transportnetzes führen können. Vor der Liberalisierung des Energiemarktes existierte in den bestehenden Gebietsmonopolen, in denen Versorgungsunternehmen für Erzeugung, Transport, Verteilung und Vertrieb zuständig

waren, kein Wechsel von Kunden zu anderen Versorgern, was eine quasistationäre Verbrauchs- und Erzeugungsstruktur zur Folge hatte. Den Unternehmen waren die wichtigsten Kunden im Netzgebiet und deren voraussichtlicher Leistungsbedarf auf Grund einer weitreichenden, konsistenten, historischen Datenbasis weitgehend bekannt. Somit konnten mit relativ einfachen statistischen Methoden sehr gute Last- und Regelenergieprognosen erstellt werden.

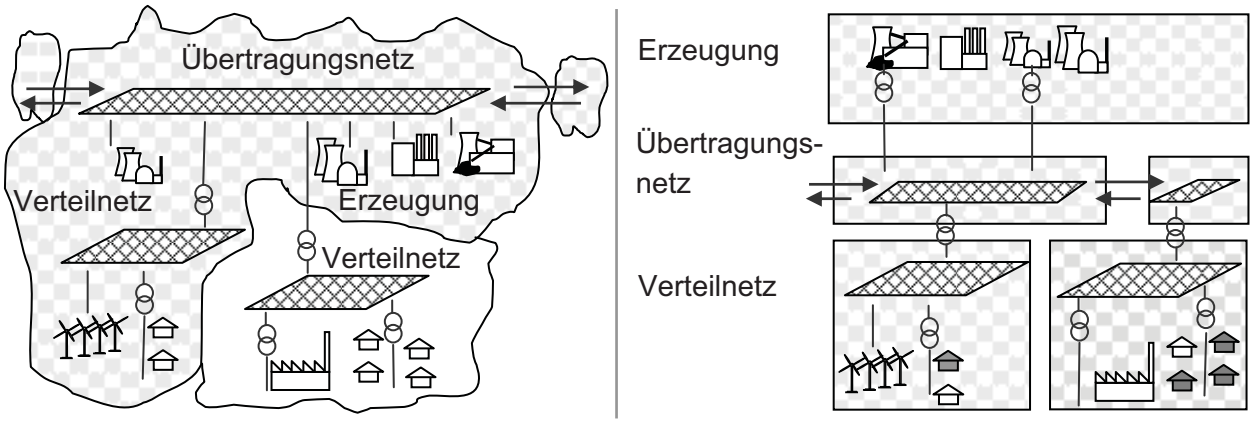


Abbildung 2-2: Strukturen der Energieversorgungsunternehmen

Diese Prognosen bildeten gleichzeitig sowohl die Last im Netzgebiet als auch den Fahrplan für die Erzeugung in den eigenen Kraftwerken ab. Die Anforderung an die Güte und Genauigkeit dieser Prognoseklasse wurde aufgrund des hohen finanziellen Risikos mit größter Genauigkeit festgelegt. Die Aufgabe der Prognoseerstellung erfolgte in den bis dahin vertikal integrierten Unternehmen von Dispatchern innerhalb der Kraftwerke, die die Höhe und den zeitlichen Verlauf der vorzuhaltenden Reserven mit Hilfe einer Lastprognose bestimmten und die Verteilung möglicher Reserven innerhalb der zur Verfügung stehenden Kraftwerke nach wirtschaftlichen Gesichtspunkten optimierten [Str02]. Diese Planungsaufgaben der Energiebereitstellung erfolgten unter Berücksichtigung verschiedener Zeithorizonte.

2.4 Öffnung des Energiemarktes in Deutschland

Mit der Umsetzung der ersten Elektrizitätsbinnenmarktrichtlinie der Europäischen Union in nationales Recht im April 1998 zur Neuregelung des Energiewirtschaftsrechts wurde der rechtliche Ordnungsrahmen der Energieversorgung bereits in drei wesentlichen Punkten grundlegend verändert. Wo früher Demarkations- und Konzessionsverträge mit Ausschließlichkeitsbindung den Stromabsatzmarkt in einzelne Monopole einteilten, wurde der Gebietsschutz der bisherigen geschlossenen und flächendeckenden Versorgungssysteme aufgehoben. Des Weiteren wurden die Netzbetreiber zu einer diskriminierungsfreien Öffnung ihrer Netze gegenüber Dritten verpflichtet. Und schließlich wurde das organisatorische, eigentumsrechtliche oder buchhalterische Unbundling gefordert, was eine Aufteilung der klassischen, integrierten Unternehmens-

strukturen in separate Unternehmen der Erzeugung, Übertragung, Verteilung, des Handels und Vertriebs zur Folge hatte [Ewg98]. Besonders der Unbundling-Forderung wurde durch den Erlass der EU-Beschleunigungsrichtlinie Strom im Juni 2003 Nachdruck verliehen und diese auch im Referentenentwurf des Gesetzes zur Neufassung des Energiewirtschaftsrechts integriert [RL 03]. Die Abbildung 2-2 verdeutlicht die unterschiedlichen Unternehmensstrukturen der monopolistischen und liberalisierten Energiemärkte Deutschlands.

Die Umsetzung der Vorgaben des Energiewirtschaftsgesetzes basiert in Deutschland erstmalig auf dem Konzept des verhandelten Netzzugangs. Dieses gewählte Durchleitungsprinzip baut auf einer freiwilligen Verbändevereinbarung auf, die zwischen der Vereinigung Deutscher Elektrizitätswerke e. V., dem Verband der industriellen Energie- und Kraftwirtschaft e. V. und dem Bundesverband der Deutschen Industrie (BDI) vereinbart wurde. Sie regelt die Kriterien zur Bestimmung von Durchleitungsentgelten, Netznutzung sowie organisatorische Rahmenbedingungen und wurde mehrfach überarbeitet.

Mit der zweiten Novelle des Energiewirtschaftsgesetzes (EnWG) setzte die Bundesregierung am 13.07.2007 das EU-Gemeinschaftsrecht für die leitungsgebundene Energieversorgung in nationales Recht um. Zweck des EnWG ist die "möglichst sichere, preisgünstige, verbraucherfreundliche, effiziente und umweltverträgliche leitungsgebundene Versorgung der Allgemeinheit mit Elektrizität und Gas" (§ 1 Abs. 1 EnWG). Ferner dient die Regulierung der Elektrizitäts- und Gasversorgungsnetze den Zielen der Sicherstellung eines wirksamen und unverfälschten Wettbewerbs [EnWG07] und der Sicherung eines langfristig angelegten leistungsfähigen und zuverlässigen Betriebs von Energieversorgungsnetzen (§ 1 Abs. 2 EnWG), [Eick06], [Jan96], [PWC07]. Die entscheidenden Inhalte des Gesetzes lassen sich wie folgt zusammenfassen:

1. Das neue Energiewirtschaftsgesetz verpflichtet alle Netzbetreiber, ihre Netze (natürliche Monopole) diskriminierungsfrei allen Kunden gegen ein angemessenes Entgelt zur Verfügung zu stellen.
2. Das System des regulierten Netzzugangs tritt an die Stelle des bisher geltenden Prinzips des verhandelten Netzzugangs (auf Basis der Verbändevereinbarungen).
3. Der Netzbetreiber darf dem Kunden nur genehmigte Netzentgelte in Rechnung stellen. Basis für die Netzentgelte ist die NetzentgeltVO Strom/Gas.
4. Die Regulierungsbehörden überwachen die Netzbetreiber.
5. Größere Energieversorger (mit mehr als 100.000 angeschlossenen Kunden) müssen ihren Netzbereich von allen anderen wirtschaftlichen Aktivitäten innerhalb des Unternehmens trennen (Unbundling). Das Gleiche gilt für Energieversorger, die im Sinne der EG-Fusionskontrollverordnung verbunden sind. Damit werden Kunden unterschiedliche Ansprechpartner für Lieferverträge bzw. für Netznutzungs-/Netzanschlussverträge im selben Versorgungsunternehmen haben.

6. Völlig neu ist der Zugang zu Gasversorgungsnetzen geregelt und hierbei speziell die Fragestellung der Einspeise- bzw. Ausspeiseverträge. Damit wird der Zugang zum gesamten deutschen Gasnetz ermöglicht.

Damit werden folgende wesentlichen Änderungen der Novellierung zu den bis dato erfolgten Liberalisierungsbemühungen umgesetzt:

- Der Informationsaustausch zwischen den Marktpartnern wird einheitlicher geregelt.
- Das Unbundling wird gesetzlich geregelt.
- Die Regelungen und Aufgaben für Übertragungsnetzbetreiber werden konkretisiert.
- Die Befugnisse und Aufgaben der Bundesnetzagentur werden festgelegt.
- Die Grund- und Notversorgungsverpflichtungen werden dem liberalisierten Energiemarkt angepasst.
- Eine Kennzeichnungspflicht für Stromrechnungen wird eingeführt.
- Festlegungen zum Berechnungs- und Genehmigungsverfahren für die Netznutzungsentgelte, zur Informationspflicht im Internet (Entgelte) und zum Reporting an die Regulierungsbehörde werden getroffen.
- Der Messstellenbetrieb wird liberalisiert, d. h. der Anschlussnehmer kann sich den Betreiber seines Strom- oder Gaszählers frei aussuchen.
- Die Zählerablesung kann auf Grund einer Verordnung der Bundesregierung liberalisiert werden, hier kann dann der Anschlussnutzer entscheiden.

Die in den zurückliegenden Jahren veränderten Rahmenbedingungen - und hier speziell auch das geforderte Unbundling - führten somit zu Handelsbeziehungen zwischen den einzelnen Marktteilnehmern. Ein am Markt agierendes Vertriebsunternehmen kann nun Kunden in unterschiedlichen Verteil- oder Übertragungsnetzen mit Energie beliefern, die bei Erzeugungsgesellschaften produziert und durch eine Netzgesellschaft dem Kunden geliefert wird. Durch das Stromhandelsunternehmen wird in der Regel eine Prognose aller Kunden in Form eines Erzeuger- und Verbrauchsfahrplans dem zuständigen Übertragungsnetzbetreiber zur Verfügung gestellt. Übertragungsnetzbetreiber sind gesetzlich verpflichtet, für ein frequenz- und spannungsstabiles Übertragungsnetz zu sorgen, verfügen aber aufgrund des Unbundlings nicht über eigene Kraftwerke, die zur Ausregelung der Abweichung von Ist-Last zu den angemeldeten Fahrplänen erforderlich sind. Als Konsequenz hieraus erfolgt die Beschaffung der notwendigen Regelenergie in Form von Reserveleistung am Markt. Die Festlegung der Höhe und des zeitlichen Verlaufs der benötigten Regelenergie erfolgte somit nicht mehr durch Dispatcher innerhalb der Kraftwerke sondern durch den Übertragungsnetzbetreiber.

2.5 Anforderungen an Lastprognosen

Es ist somit aufgrund der gesetzlichen Rahmenbedingungen und der veränderten Marktsituation erforderlich, für Vertriebs- und Handelsunternehmen sowie für die Stromhandelsaufgaben des Netzbetreibers getrennte Prognosen mit zum Teil völlig

unterschiedlichen Anforderungen an Methoden und Systemen zu erstellen. Im Folgenden werden vor dem Hintergrund der Marktrollen die Anforderungen an Lastprognosen definiert.

2.5.1 Anforderungen des Handels und Vertriebs

Gemäß den heutigen Vorgaben wird Deutschland als eine und die Netze der europäischen Nachbarstaaten als weitere Energiehandelszone betrachtet. Diese sind wiederum innerhalb Deutschlands in vier Regelzonen der jeweiligen Übertragungsnetzbetreiber amprion, E.ON, EnBW und 50Hertz Transmission GmbH unterteilt. Jeder der vier deutschen Übertragungsnetzbetreiber führt Bilanz über alle Lieferungen und Bezüge innerhalb seiner Regelzone sowie über die regelzonenübergreifenden Energielieferungen. Die von einem Großkunden, Händler oder Vertrieb belieferten Einspeise- und Entnahmestellen innerhalb einer Regelzone sowie Fahrplanlieferungen an andere Marktteilnehmer werden in einem Bilanzkreis zusammengefasst und über diesen mit dem Übertragungsnetzbetreiber unter Berücksichtigung eines Toleranzbandes oder Naturalkontos abgerechnet. Dabei übernimmt der Verantwortliche des Bilanzkreises die Verpflichtung, für eine Ausgeglichenheit zwischen Beschaffung (Erzeugung oder Import) und Abgabe (Verbrauch oder Export) innerhalb eines 15-minütigen Zeitrasters zu sorgen [VDN02]. Weitere Möglichkeiten zur Abwicklung von Handelsgeschäften können über Subbilanzkreise als Teilmengen von Bilanzkreisen - oder Aggregationskreise - als Zusammenschluss mehrerer Händler innerhalb eines Bilanzkreises erfolgen. Der prognostizierte Fahrplan eines Händlers muss am Tag vor der Lieferung in Form eines 15-minütigen Zeitrasters bis 14:30 Uhr bei dem betroffenen Übertragungsnetzbetreiber eingehen. Dieser prüft die formale Richtigkeit und die Übereinstimmung der Fahrpläne mit anderen Handelspartnern und weist gegebenenfalls den Verantwortlichen auf Fehler hin. Zusätzlich erhalten Übertragungs- und Verteilnetzbetreiber die Fahrpläne von ebenfalls einspeisenden Händlerkraftwerken mit einer Leistung größer 100 MW. Bei all diesen Verfahren werden Bilanzabweichungen zwischen Einspeisung und Entnahme, wie oben beschrieben, durch den Übertragungsnetzbetreiber mit Hilfe von Regelenergie ausgeglichen und mit dem Bilanzkreisverantwortlichen abgerechnet.

Der Stromhandelsmarkt, der im Gegensatz zum Erzeugermarkt erheblich niedrigere Marktzutrittsschranken hat, reagierte in den vergangenen Jahren deutlich auf die Liberalisierung des Elektrizitätssektors. Seit dem Beginn der Liberalisierung haben sich mehr als 200 überwiegend Vertriebs- und Handelsunternehmen als neue Teilnehmer im Markt etabliert [Mell03]. Die Handelsunternehmen sind den Bereichen Großhandel, z. B. Handelsabteilungen der Verbundunternehmen und Einzelhandel, zuzuordnen. Die Struktur der Kundenwechsel ist in Abbildung 2-3 dargestellt. Im Jahr 2007 war der anhaltende Trend zu beobachten, dass der größte Kundenwechsel im Marktsegment der Groß- und Industriekunden stattgefunden hat. Ein im Verhältnis hierzu eher geringer

Teil der Wechselkunden bildeten die Segmente der Gewerbe- und der Haushaltskunden. Um den Marktzugang für Haushalts- und Gewerbekunden ohne registrierende ¼-h-Leistungsmessung zu ermöglichen und dabei die vorhandenen Messeinrichtungen weiterhin zu nutzen, wurden Anfang 2000 die synthetischen und analytischen Lastprofile für diese Kundengruppen entwickelt. Der Netzbetreiber legt fest, welches Verfahren zur Leistungsermittlung der Netznutzung im Rahmen einer Belieferung Anwendung findet. Zur Ermittlung der synthetischen Lastprofile werden für verschiedene Kundengruppen aus der Vergangenheit repräsentativ gemessene und normierte Lastprofile verwendet. Die Festlegung des Lastverlaufes in einem ¼-stündlichen Zeitraster erfolgt unter Berücksichtigung verschiedener Jahreszeiten und Tagtypen sowie der prognostizierten Jahresarbeit des jeweiligen Kunden exakt.

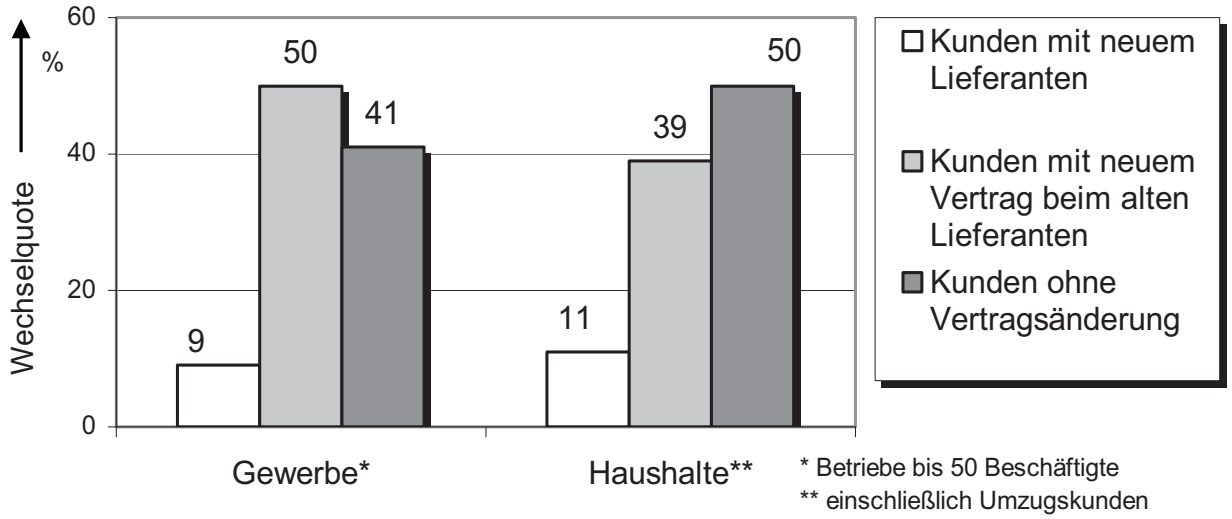


Abbildung 2-3: Kundenwechselquote 2007 (Quelle: VDE Kundenfokus 2007)

Der Lastverlauf eines Händlers wird entsprechend dem aggregierten Verbrauch seiner Kunden in einem Netzgebiet ermittelt. Das Risiko der Profilabweichungen wird dabei vom Netzbetreiber getragen und durch einen finanziellen Zuschlag bei der Netznutzung berücksichtigt. Die Möglichkeit, das synthetische Lastprofilverfahren anzuwenden, findet für Kunden mit einer Anschlussleistung von mindestens 30 kW und einem Jahresverbrauch größer 30.000 kWh/a Anwendung, wobei die Anhebung dieser Grenzen zurzeit lebhaft diskutiert wird [Mell03].

Im Unterschied zum synthetischen basiert das analytische Verfahren auf der ¼-stündlichen Messung der Netz- und Großkundenlast, d. h. es handelt sich um gemessene Werte. Damit können die tatsächlich verbrauchten Energiemengen und Leistungen jedem Händler viertelstundenscharf zugeordnet werden. Das Verfahren sieht im einfachsten Fall vor, dass der Restlastgang, gebildet als Differenz von Netzlast und der Summe von Netzverlusten sowie des Betriebsverbrauches und der Last der gemessenen Sondervertragskunden, als analytisches Lastprofil zur Belieferung und Ab-

rechnung der Klein- und Gewerbekunden genutzt wird. Um dem unterschiedlichen Abnahmeverhalten dieser Kundengruppen Rechnung zu tragen, wird der ermittelte Restlastgang über synthetische Profile für die einzelnen Kundengruppen aufgeteilt. Die für diese Aufteilung benötigten Zerlegungsfaktoren (Z-Faktoren) lassen sich auf Grundlage der Mess- und Abrechnungswerte des Vorjahres bilden. Äquivalent zu den synthetischen Lastprofilen wird die Endabrechnung nach Ablesung des Stromverbrauchs beim Endkunden, aber unter Nutzung des zeitlich dazugehörigen analytischen Lastgangs der jeweiligen Kundengruppe, durchgeführt. Der Netzbetreiber entscheidet, welches Verfahren zur Anwendung kommt. Derzeit verwendet eine Vielzahl Verteilnetzbetreiber das synthetische Verfahren und übernimmt das damit verbundene wirtschaftliche Risiko. Umgekehrt liegt beim analytischen Verfahren die Verantwortung einer exakten Prognose ausschließlich beim Händler. Unabhängig von diesen Verfahren ist für die Belieferung von Sondervertragskunden mit registrierter ¼-h-Leistungsmessung eine Prognose des Bedarfs zu erstellen, die ebenfalls in den Fahrplan integriert wird.

Aufgrund der beschriebenen Rahmenbedingungen müssen Händler unter Berücksichtigung von zum Teil fehlenden und/oder fehlerhaften Archivdaten sowie eines verringerten Zuganges zu Informationen über die signifikanten Einflussgrößen auf das Verbrauchsverhalten ihrer Kunden Prognosen erstellen. Die Anforderungen im liberalisierten Markt lassen sich wie folgt beschreiben:

- Bestimmung der 15-Minuten-Mittelwerte, der innerhalb der nächsten 24 Stunden zu erwartenden Last, für die einem Handel und Vertrieb zugeordneten Groß- und Industriekunden eines Bilanzkreises unter Berücksichtigung exogener Einflussfaktoren.
- Prognose der Jahresarbeit für Klein- und Gewerbekunden.
- Prognose der 15-minütigen Mittelwerte für Klein- und Gewerbekunden sowie Sondervertragskunden in Netzgebieten mit einem analytischen Lastprofilverfahren unter Berücksichtigung exogener Einflussfaktoren auf das Verbraucherverhalten.
- Prognose der möglichen Abweichung der Ist- zur tatsächlichen Last sowie Eintrittswahrscheinlichkeitsbestimmung und Risikobewertung.

2.5.2 Anforderungen der Netzbetreiber

Die Handelszone Deutschland unterteilt sich in die Regelzonen der vier Übertragungsnetzbetreiber. Innerhalb ihrer Regelzonen führen diese Bilanz über alle Lieferungen und Energieentnahmen sowie den regelzonenübergreifenden Energieaustausch. Die Abwicklung der Handelsgeschäfte zwischen den Marktteilnehmern erfolgt über das Fahrplanmanagement. Dabei werden von den Bilanzkreisverantwortlichen der zugelassenen Bilanzkreise den Übertragungsnetzbetreibern am Vortag bis 14:30 Uhr für den folgenden Tag die Bezugs- und Entnahmeenergiemengen in einem 15-minütigen Zeitraster übermittelt. Bilanzkreise können dabei Stromhändler, Vertriebsgesellschaften,

aber auch große Industriekunden sein, die ihre Strombeschaffung in eigener Verantwortung durchführen. Aus der Addition aller Fahrpläne einer Regelzone errechnet sich der prognostizierte Lastgang des folgenden Tages. Der Zusammenhang ist in Abbildung 2-4 graphisch dargestellt.

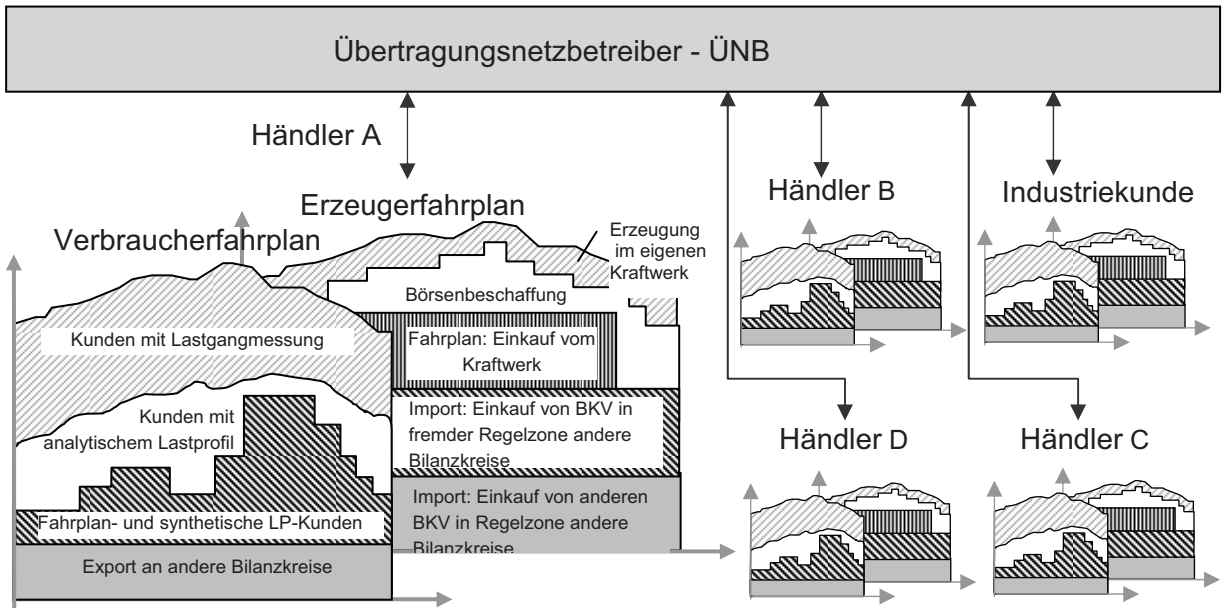


Abbildung 2-4: Fahrplanmanagement zwischen Händler und Übertragungsnetzbetreiber

Zu einer zuverlässigen Versorgung und zur Vorhaltung von Leistungsreserven zum Ausgleich der Differenz zwischen tatsächlicher Netzlast und Vertriebs- und Händlerprognose ist der Netzbetreiber laut den Anforderungen des Energiewirtschaftsgesetzes verpflichtet. Nach Vorgabe des Unbundlings dürfen Netzbetreiber jedoch nicht über eigene Kraftwerke verfügen, womit die Beschaffung des Ausgleichsenergiebedarfs am freien Markt erforderlich ist. Die Etablierung dieser Märkte gewinnt für eine kostenoptimierte Regelenergiebeschaffung für den Netzbetreiber künftig sicherlich eine stärkere Bedeutung.

Im Regelungsrahmen des zurzeit geltenden Energiewirtschaftsrechtes fallen den Übertragungs- und Verteilnetzbetreibern somit neue Aufgaben des Stromhandels zu, die wie folgt klassifiziert werden können:

- Beschaffung von Regelenergie durch den Übertragungsnetzbetreiber zum Ausgleich der Ist-Netzlast zur Handelsprognose
- Beschaffung von Netzverlustenergie durch den Netzbetreiber
- Übernahme von Kunden insolventer Bilanzkreise in eine Notversorgung
- In Netzen mit synthetischen Lastprofilverfahren ist der Netzbetreiber für den Ausgleich von Leistungs- und Arbeitsabweichungen zwischen Lastprofilwerten und tatsächlichen Abnahmen der Kleinkunden verantwortlich

- In Netzen, mit analytischen Lastprofilverfahren, ist der Netzbetreiber für den Ausgleich von Arbeitsmengenabweichungen, die im Rahmen der Mehr- und Mindermengen entgolten werden, verantwortlich.

Gerade vor dem Hintergrund der Zunahme regenerativer Energie erfordert diese künftig den optimierten Einsatz von Regelenergie. Exemplarisch verdeutlicht dieses die Abbildung 2-5, in der die Tagesmaxima und Tagesminima der ¼-h-Leistungsprofile von EEG-Windenergie-Einspeisungen in Deutschland eines Monats dargestellt sind.

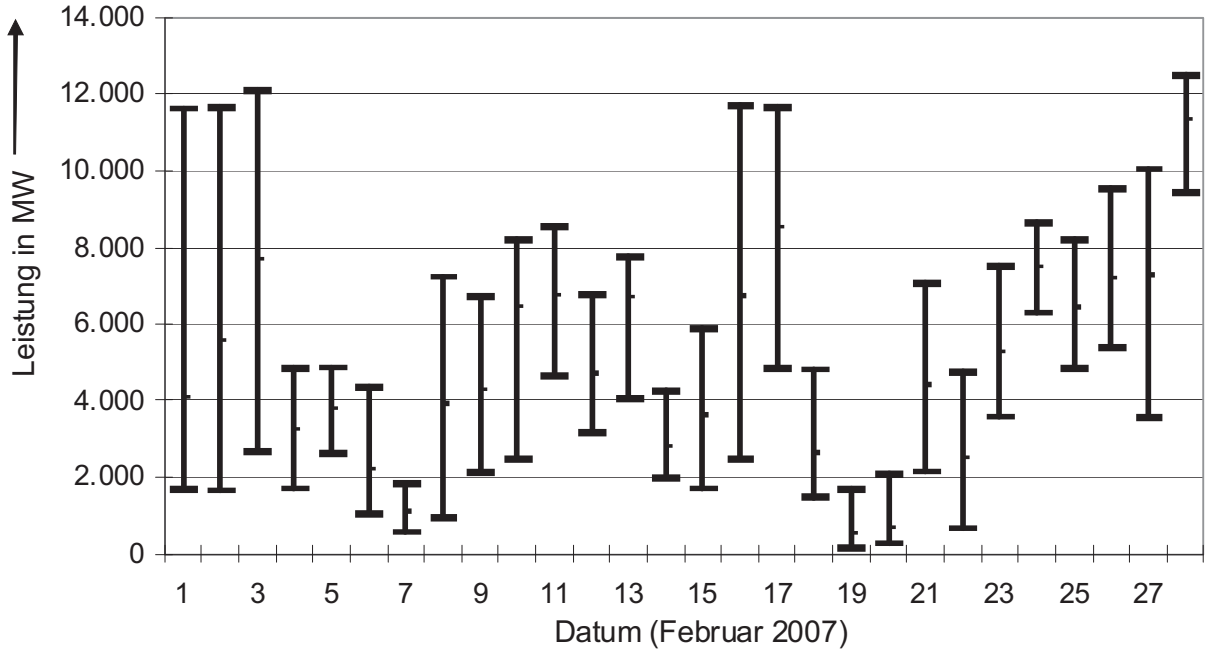


Abbildung 2-5: Windenergie-Einspeisung in Deutschland, ¼-h-Leistungswerte der Tagesmaxima und Tagesminima (Quelle: VDN)

Die Beschaffung der Energie zum Ausgleich dieser Differenz lässt sich grundsätzlich in die Regelenergiearten Primärregelleistung, Sekundärregelleistung und Minutenreserve (Tertiärregelreserve) unterteilen. Während die Primärleistungsregelung automatisch in Abhängigkeit der lokal gemessenen Frequenzabweichungen durch die Kraftwerksblöcke innerhalb der Regelzonen erbracht wird, obliegt die Aktivierung von Sekundär- und Minutenreserve der Systemführung des zuständigen Übertragungsnetzbetreibers der betroffenen Regelzone.

Bei einer Abweichung der momentanen Netzlast vom erwarteten Wert sind zwei Anteile, die Abweichung des Leistungsmittelwertes der Abrechnungsperiode von der geplanten mittleren Leistung (der eigentliche Lastprognosefehler), und die Abweichung der momentanen Leistung vom Leistungsmittelwert der Abrechnungsperiode (Einspeise- und Entnahmeschwankungen) zu unterscheiden [Gra01]. Die unterschiedlichen Arten von Regelenergien werden seit Beginn des Jahres 2001 über eine Ausschreibung unter Wettbewerbsbedingungen im Strommarkt beschafft. Anbieter für die verschiedenen Regelenergiearten können sich an einem Präqualifikationsverfahren beteiligen. Nach

erfolgter Präqualifikation ist der Erzeuger berechtigt, sich an den sich wiederholenden Ausschreibungsverfahren zur Erbringung von Regelenergie zu beteiligen.

Künftig fallen somit den Übertragungs- und Verteilnetzbetreibern neue Aufgaben und Verantwortung der Strombeschaffung zu, die bisher in integrierten Versorgungsunternehmen beim verbundenen Vertrieb lagen. Gerade die zurzeit noch integrierten Versorgungsunternehmen, die von den Unbundlingvorschriften betroffen sein werden, stehen somit im Spannungsfeld zwischen den Mindestanforderungen der Richtlinienerfüllung und dem damit verbundenen Maß an Transparenz entsprechend den Vorgaben der Regulierungsbehörde. Es ist zu erwarten, dass aufgrund der zunehmenden Trennung zwischen Netz und Vertrieb zurzeit noch verbundener Unternehmen die eigenverantwortliche Führung von Differenzbilanzkreisen für Netzbetreiber, insbesondere vor dem Hintergrund der Beschaffung von Energie als Ausgleichsmengen der Netzverluste und dem damit gegenüber der Regelenergie verbundenen Preisvorteil, künftig an Bedeutung gewinnen wird [StromNZV05], [StromNEV05], [Hai04], [EnWG07], [EnWG98]. Maßgeblich für das Risiko bei der Übernahme einer Bilanzkreisverantwortung ist hierbei das im Netzgebiet angewendete Lastprofilverfahren für Kleinkunden (bis 120.000 kWh/a). Dieses setzt sich aus zwei Komponenten zusammen, der Abweichung zwischen angenommenem und tatsächlichem Profil (Profilrisiko) und der Mengenabweichung, die jedoch standardisiert mit der Mehr- und Mindermengenabrechnung nach Ablauf eines Abrechnungsjahres finanziell ausgeglichen wird. Um künftig Regelenergiekosten für solche Differenzbilanzkreise zu minimieren, wird es neben der bisher praktizierten Meldung von Ist-Werten an den Bilanzkreiskoordinator auch erforderlich sein, Netzverluste optimal zu prognostizieren. Wird nämlich aus einer solchen Prognose resultierender Verlustlastgang am Markt beschafft, so sollte dies erheblich günstiger als die Beschaffung von Regelenergie sein. Dieses bedeutet aber einen für die laufende Prognose sowie das Bilanzkreismanagement entsprechenden Mehraufwand. Zielsetzung der Netzbetreiberprognose ist die Minimierung des Risikos und damit die Reduzierung der Kosten für die Beschaffung der vorzuhaltenden Regel- und Ausgleichsenergiemengen der Netzverluste, woraus sich folgende Anforderungen an Optimierungssysteme ergeben:

- Bestimmung der 15-minütigen Mittelwerte der innerhalb der nächsten 24 Stunden zu erwartenden Netzlast unter Berücksichtigung dezentraler Erzeugung, z. B. durch Windkraftanlagen oder Blockheizkraftwerke und exogener Einflussfaktoren.
- Erstellung einer Prognose des Erwartungswertes der Last sowie Prognose der Prognosefehler bzw. Konfidenzintervalle.
- Quantifizierung der Unsicherheiten und Risikobewertung einer kostenoptimierten Regelenergiebeschaffung.
- In Netzen mit synthetischen Lastprofilverfahren ist die Prognose des Lastprofils erforderlich, um das Profilrisiko für den Netzbetreiber zu minimieren.
- Prognose der Netzverluste.

2.5.3 Lastprognosesysteme

Untersuchungen über den Einsatz der verwendeten Prognoseverfahren wurden erstmalig im Jahr 1992 im Rahmen einer CIGRE-Umfrage [CIG92] in 44 internationalen Versorgungsunternehmen durchgeführt. Als Ergebnis hat sich dabei herausgestellt, dass in der Praxis zu diesem Zeitpunkt neben den bis heute noch angewendeten manuell empirischen Prognoseverfahren überwiegend Regressionsverfahren und hybride Verfahren Anwendung fanden. Abbildung 2-6 zeigt den prozentualen Anteil der verwendeten Verfahren in Abhängigkeit unterschiedlicher Zeithorizonte.

Um eine aktualisierte Übersicht der zurzeit am Markt eingesetzten Methoden und Modelle sowie anderer für die Prognoseerstellung relevanter Einflussgrößen zu erhalten, wurde im Rahmen dieser Arbeit eine Befragung deutscher Energieversorgungsunternehmen durchgeführt. Dabei wurden Gruppen gebildet, die sich entsprechend der Rollen im liberalisierten Markt (Händler/Vertrieb/Verteil- und Übertragungsnetzbetreiber) gliedern.

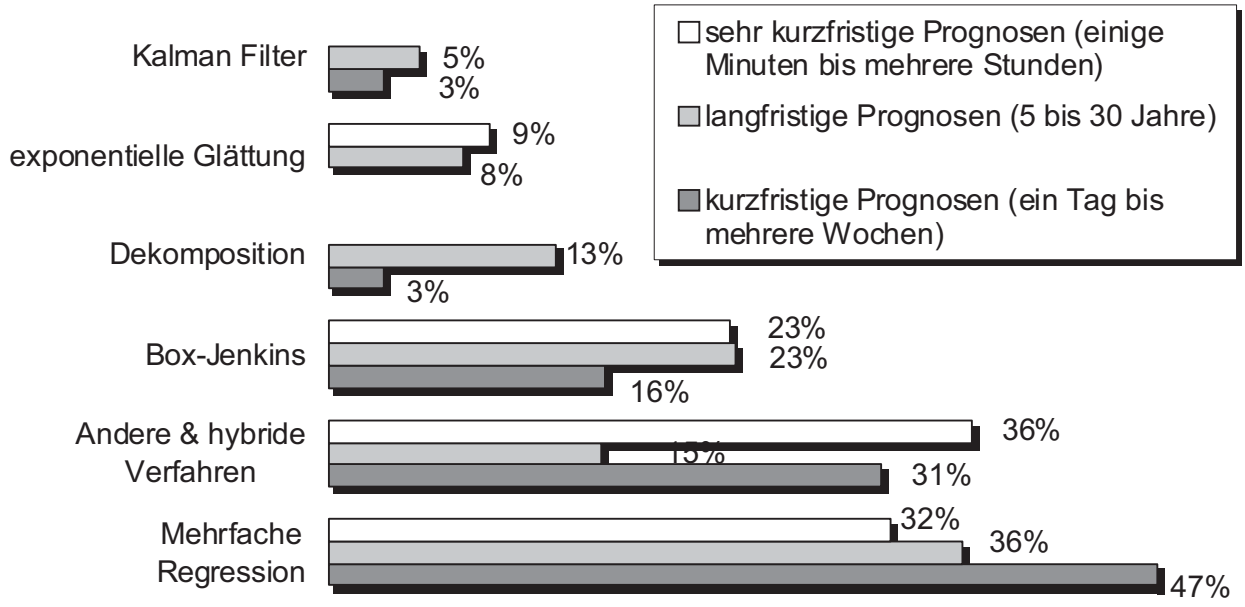


Abbildung 2-6: CIGRE Umfrage: Verfahren zur kurzfristigen Lastprognose

Die Befragung erfolgte mit Hilfe eines im Internet zur Verfügung gestellten Fragebogens. Die Beteiligung lag bei den 620 befragten Netzbetreibern bei 12 %. Der Rücklauf von Antworten der rund 100 Vertriebs- und Handelsunternehmen kann mit 45 % beziffert werden, damit können die Ergebnisse als durchaus repräsentativ bewertet werden.
Einhergehend mit der Marktöffnung erfolgte bei vielen der befragten Unternehmen die Einführung von Systemen zur Lastprognose. Bei der Befragung konnte festgestellt werden, dass bis zum Jahr 2001 rund 68 % der teilnehmenden Handels- und Vertriebs-

unternehmen DV-gestützte Systeme zur Erstellung von Lastprognosen in ihren Unternehmen eingeführt hatten; die meisten Netzbetreiber jedoch erst seit 2000 begonnen haben, Lastprognosesysteme zu nutzen. Der zeitliche Verlauf der Einführung von Prognosesystemen seit 1990 bei Netzbetreibern sowie Handels- und Vertriebsunternehmen ist in Abbildung 2-7 dargestellt.

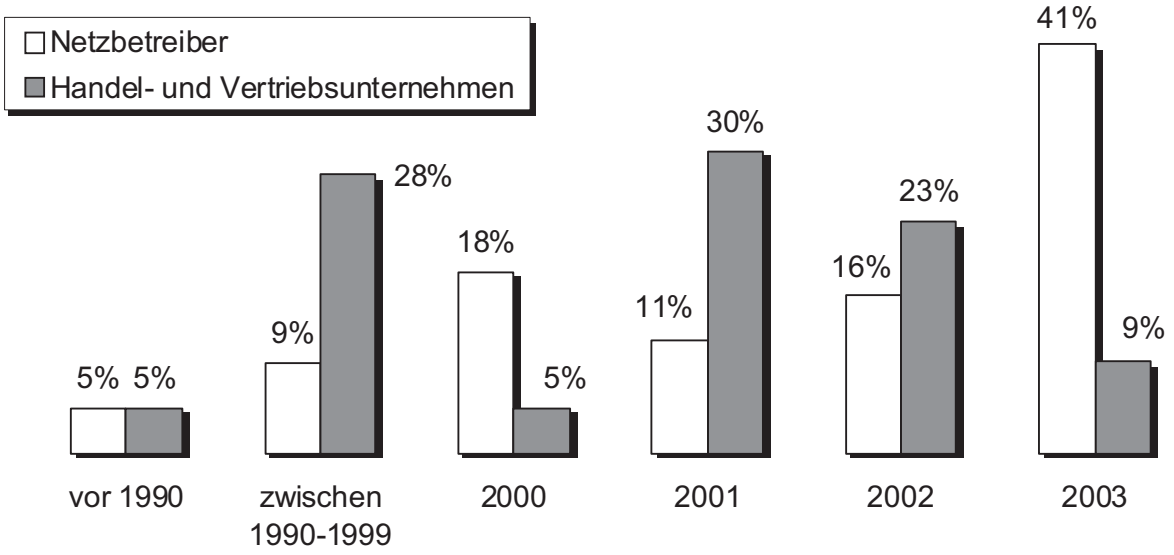


Abbildung 2-7: Einführung von Lastprognosesystemen

Rund 60 % der teilnehmenden Netzbetreiber führen Prognosen ihrer Netzlast durch. Der Aufwand hierfür wurde durchschnittlich mit 1,2 Personentagen pro Woche beziffert, die laufenden jährlichen Kosten für den Betrieb solcher Systeme wurden mit durchschnittlich 21.000 €/a angegeben. Das angewendete Lastprofilverfahren für Klein- und Gewerbekunden und die Verwendung der Lastprofilarten beim Einsatz des synthetischen Verfahrens sind in Abbildung 2-8 dargestellt.

In der Gruppe der Vertriebs- und Handelsunternehmen waren 16 % reine Handelsunternehmen, die über keinen eigenen Netzbetrieb verfügen. Der Personalaufwand für die Erstellung von Lastprognosen innerhalb dieser Gruppen kann mit 1,6 Personentagen pro Woche beziffert werden, die jährlichen Betriebskosten für Lastprognosesysteme wurden durchschnittlich mit 12.800 € angegeben. Rund 75 % der befragten Unternehmen führen Tageslastprognosen durch, wobei diese zu 90 % in einem Zeitintervall von ¼-h erfolgen. Knapp die Hälfte aller Unternehmen erstellt zusätzlich Wochen- und Monatslastprognosen ebenfalls in einem ¼-h-Zeitraster. Vereinzelt werden jedoch auch Prognosen im stündlichen Zeitintervall durchgeführt.

Prognoseverfahren lassen sich grundsätzlich in vier Verfahrensgruppen einteilen. Bei manuell empirischen Verfahren erfolgt eine Fortschreibung gegebener Lastgangdaten; Korrekturen erfolgen dabei soweit erforderlich manuell durch einen Anwender.

Bei univariaten Verfahren beschränkt sich die Analyse auf die historischen Lastgangdaten.

Werden komplexe Beziehungen mit mehreren Prognosevariablen, d. h. neben den Lastgangdaten weitere exogene Einflussfaktoren berücksichtigt, so werden diese Verfahren als multivariate Verfahren bezeichnet.

Verfahren, die sich aus der Kombination der genannten Anwendungen ergeben, werden als Kombinationsmodelle bezeichnet.

Bei den befragten Unternehmen verwendeten ca. 15 % univariate Verfahren. Die verbleibenden Verfahren werden zu je einem Drittel in den befragten Unternehmen verwendet.

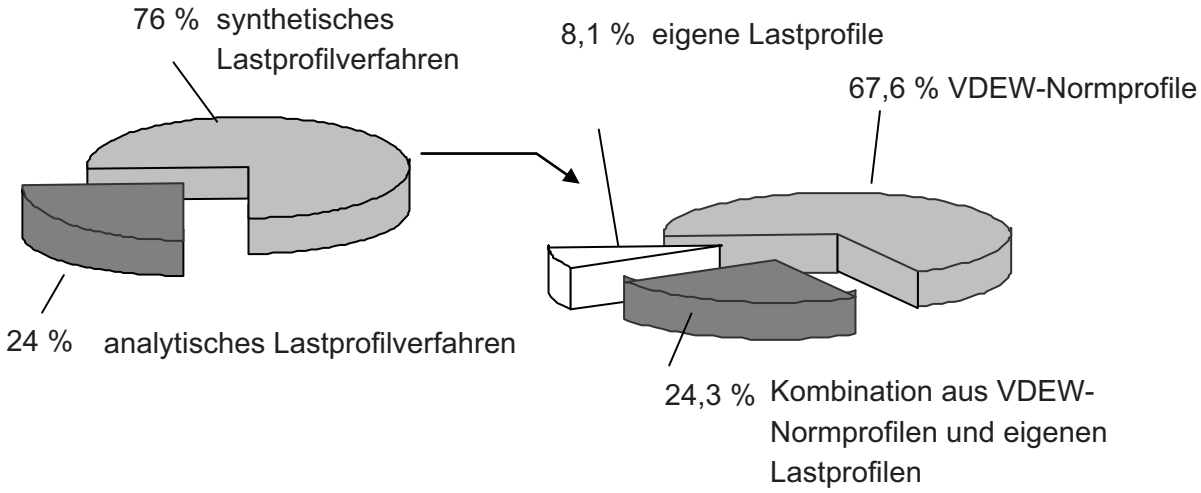


Abbildung 2-8: Verteilung der eingesetzten Lastprofilverfahren

Ziel der Modellbildung im Rahmen von Lastprognosen ist, eine vereinfachte Abbildung der wesentlichen Zusammenhänge zwischen den Einflussgrößen und den gesuchten Ausgangswerten herzustellen. In der heutigen Praxis findet eine Vielzahl von Modellen Anwendung. Im klassischen Regressionsverfahren wird auf die Abhängigkeiten von Einflussvariablen zurückgegriffen und mit Hilfe einer Analyse werden Koeffizienten entsprechend ihrer Gewichtung in einem Gleichungssystem bestimmt. Dieses Verfahren findet für lineare und nichtlineare Systeme Anwendung. Stochastische Zeitreihen umfassen Ansätze zur Beschreibung linearer Abhängigkeiten zwischen vergangenen und aktuellen Werten einer zu prognostizierenden Größe. Diese Verfahren stellen im Regelfall eine Kombination von Autoregressiven (AR) und Moving-Average-(MA) Modellen dar. Die Berücksichtigung extremer exogener Parameter in einer Verfahrenserweiterung erfolgt über Transferfunktionsmodellierung mit Hilfe des Box-Jenkins-Verfahrens.

Gerade in den vergangen Jahren haben neuartige Ansätze wie Fuzzy-Modelle und neuronale Netze zur Prognoseerstellung in der Energieversorgung Anwendung gefunden. Voraussetzung bei der Anwendung der Fuzzy-Modellierung ist ein qualitatives Verständnis zwischen den zu identifizierenden Einfluss- und den zu prognostizierenden Ausgangsgrößen und die sich hieraus ergebende Bildung eines regelbasierten Systems in Form von linguistischen Aussagen wenn Prämisse dann Konklusion. Auf Basis dieser Aussagen werden die Abhängigkeiten der Systemgrößen in qualitativer Form und daraus die Zusammenhänge zwischen Ein- und Ausgangsgrößen beschrieben. Künstliche neuronale Netze sind universelle Approximatoren, bestehend aus nichtlinearen Berechungseinheiten, die wie Neuronen in biologischen neuronalen Netzen miteinander verknüpft sind. In der Lastprognose hat sich der Multilayer-Ansatz der neuronalen Netze durchgesetzt, der häufig analog zum weitverbreiteten Regressionsverfahren zur Modellierung der Zusammenhänge zwischen Ein- und Ausgangsgrößen verwendet wird. Im Rahmen der Befragung wurden in Abhängigkeit der Marktrolle die in Abbildung 2-9 dargestellten Verfahren in den befragten Unternehmen eingesetzt.

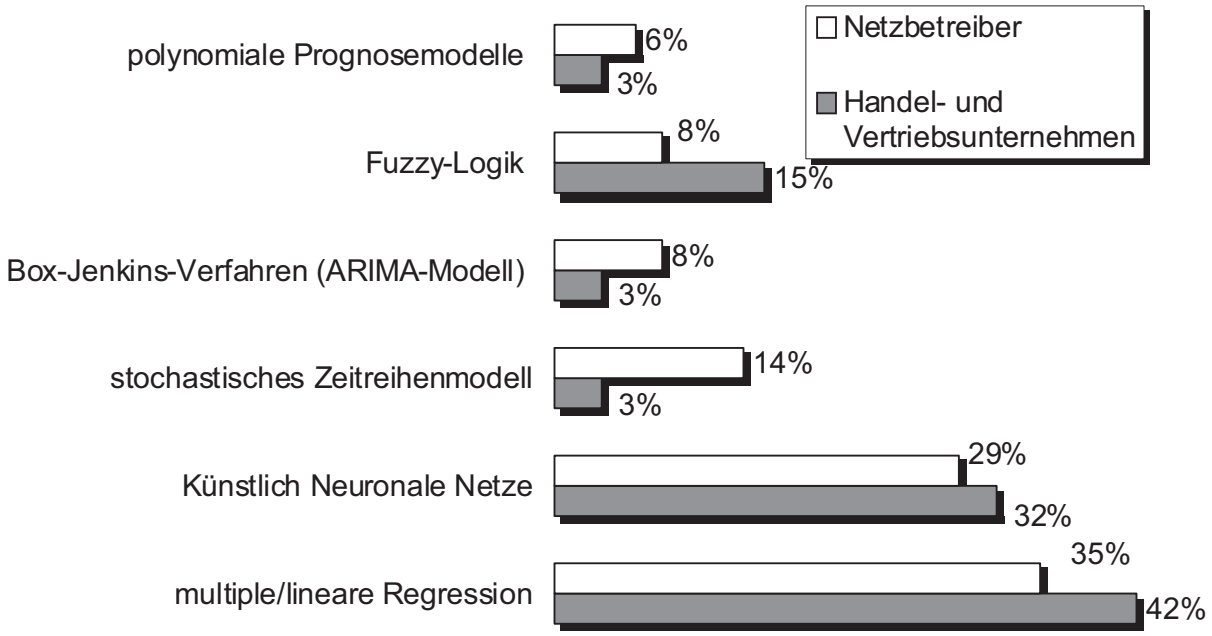


Abbildung 2-9: Einsatz von Prognosemodellen

Bei der Verwendung von multivariaten Verfahren und Kombinationsverfahren ist es erforderlich, exogene Einflussgrößen im Prognoseverfahren zu berücksichtigen. Wesentlicher Grund hierfür ist, dass die Gesamtlast eines Versorgungsnetzes durch exogene Einflussgrößen bestimmt wird und diese Größen vom jeweiligen EVU und stark von den strukturellen Gegebenheiten innerhalb ihres Versorgungsgebietes abhängig sind. Von den teilnehmenden Unternehmen wurden die in Abbildung 2-10 dargestellten Einflussgrößen benannt, die bei ihrer Prognoseerstellung berücksichtigt werden.

Einzelne der teilnehmenden Unternehmen haben weitere Parameter benannt, die aus ihrer Sicht Einfluss auf die Prognoseerstellung nehmen. Dabei wurden EEX-Preise, Luftfeuchte, Schaltprogramme, Niederschlagsmengen aufgrund der Eigenerzeugung aus Wasserkraft, Öffnungszeiten, Schichtzeiten der Großindustrie und Tageslichtdauer sowie Helligkeit benannt. Da eine große Anzahl wetterempfindlicher Verbraucher wie Heizungs-, Klima- und Beleuchtungsanlagen usw. Einfluss auf die elektrische Last nimmt, hat die Qualität der Vorhersage meteorologischer Größen unmittelbaren Einfluss auf die Lastprognosequalität. Von den beteiligten Unternehmen nutzen rund 44 % Wettervorhersagen des Deutschen Wetterdienstes, 47 % nutzen Angebote privater Wetterdienste und rund 11 % die Dienstleistung beider.

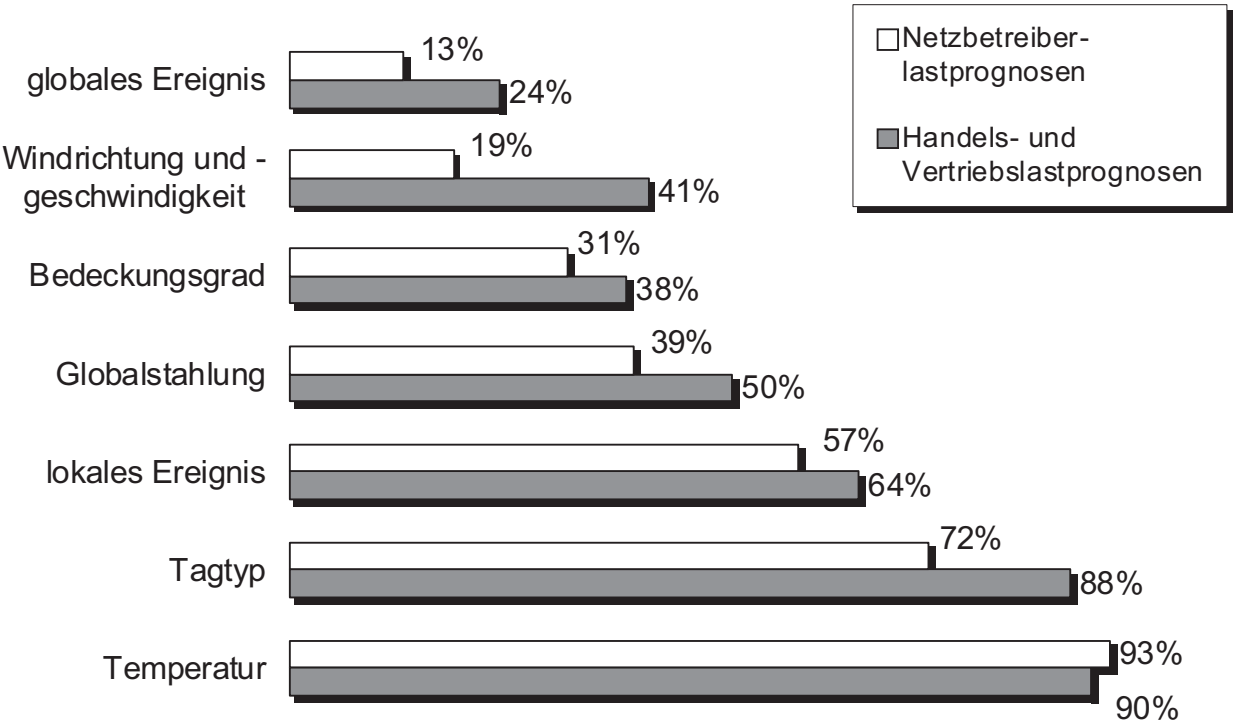


Abbildung 2-10: Berücksichtigung exogener Größen zur Prognoseerstellung

Eine ex-post-Beurteilung von Prognosen kann mit Hilfe unterschiedlichster statistischer Verfahren erfolgen. Aufgrund dieser Tatsache wurde im Rahmen der Befragung auf die Nennung gemessener statistischer Größen verzichtet und der Schwerpunkt auf eine qualitative Beurteilung gelegt. Dabei gaben 15 % der befragten Unternehmen an, sehr zufrieden mit Ihren Prognoseergebnissen zu sein, 70 % zeigen sich als zufrieden, die verbleibenden 15 % sind mit den Prognoseergebnissen wenig oder gar nicht zufrieden.

2.6 Werkzeuge und IT- Systeme

Seit Inkrafttreten der Deregulierung des deutschen Strommarktes wurden die Energieversorger vor neue Aufgaben gestellt, den möglichen Kundenwechsel nicht nur organisatorisch sondern auch energiewirtschaftlich zu lösen. Um die Zeitgleichheit von Entnahme und Produktion im Netz energiewirtschaftlich zu verrechnen, konnte vor der Liberalisierung auf die jeweiligen Versorgungsgebiete zurückgegriffen werden, deren stabiler Kundenbestand mit seinen Verbrauchern aggregiert über die Netzlast erfasst wurde. Die Zurechenbarkeit des Stromverbrauches und die Lieferverpflichtung waren eindeutig geregelt. Mit der Liberalisierung löste sich diese eindeutige Beziehung auf. Es muss nun gewährleistet werden, dass für jeden Kunden der zugehörige Lieferant identifiziert wird und jedem Lieferanten ein Erzeuger zugeordnet und mit diesem abgerechnet wird. In diesem Markt sind daher Zählwerte und Zeitreihen von herausragender Bedeutung. Da der zeitliche Verlauf des Planens von Transaktionen aller Marktteilnehmern und das Produzieren, Liefern und Abrechnen von Energie die nunmehr zentralen Aufgaben zwischen den Marktteilnehmen sind, ist es erforderlich, große Datenmengen untereinander auszutauschen. Eine automatische Abwicklung der vollständigen Prozesskette von der Messung über den Datenaustausch bis hin zur Abrechnung ist daher aus wirtschaftlichen Überlegungen unumgänglich.

Die Bundesnetzagentur hat in ihrem Festlegungsverfahren BK6-06-009 einheitliche Geschäftsprozesse und Datenformate für die Elektrizitätswirtschaft vorgegeben. Als Basis für den elektronischen Datenaustausch wurde das Datenformat UN/EDIFACT gewählt [StromNZV05] [GPKE06].

In Zusammenhang mit der Entwicklung von IT-Lösungen, die diese Anforderungen erfüllen, hat sich der Begriff EDM - Energiedatenmanagementsystem herausgebildet. Der Funktionsumfang dieser Systeme ist jedoch unterschiedlich stark in Abhängigkeit der Marktrolle ausgeprägt.

Die hierfür zurzeit eingesetzten Systeme sind im Regelfall modular aufgebaut, um einzeln oder kombiniert den Interessen der unterschiedlichen Marktteilnehmer gerecht zu werden. Die wesentlichen Bestandteile eines Systems bestehen aus den Modulen Netznutzung- und Bilanzierungssystem sowie dem Stromhandels- und Lastprognosesystem. Die Funktion des Netznutzungs- und Bilanzierungssystems besteht in der Übernahme der Messwerte aus der Zählerdatenerfassung und der Aggregation der Lastgänge nach Kundengruppen und Händlern. Im Stromhandelssystem erfolgt die Kalkulation und Durchführung der Energielieferung sowie das Portfolio- und Handelsfahrplanmanagement. Notwendiger Bestandteil beider Module sind Funktionen der Kommunikation mit Marktteilnehmern und die Verwaltung von Markt- und Vertragsbeziehung. Im Lastprognosesystem werden auf Basis der Zeitreihen Prognosen für den Handel und Vertrieb und zur Regelenergiebeschaffung des Netzbetreibers erstellt.

Die EDM-Systeme und die zählpunktbezogene Kommunikation nehmen bereits heute und werden künftig eine zentrale und wichtige Komponente in der IT-Landschaft aller am Energiemarkt beteiligten Teilnehmer darstellen. Dabei ist es für das Unternehmen nicht relevant, ob es als reines Vertriebsunternehmen oder Netzbetreiber arbeitet oder ob es beide Marktrollen einnimmt. Unabdingbar für die Aufgabenerfüllung dieser Rolle wird die Planung der Energiebereitstellung durch qualitativ hochwertige Prognosen sein, die eine maßgebliche Rolle in einem integrierten System einnehmen werden. Die Funktion dieser Systeme und die auf Mandantenebene getrennte integrierte Nutzung durch die geforderte Unternehmensstruktur sind in Abbildung 2-11 dargestellt.

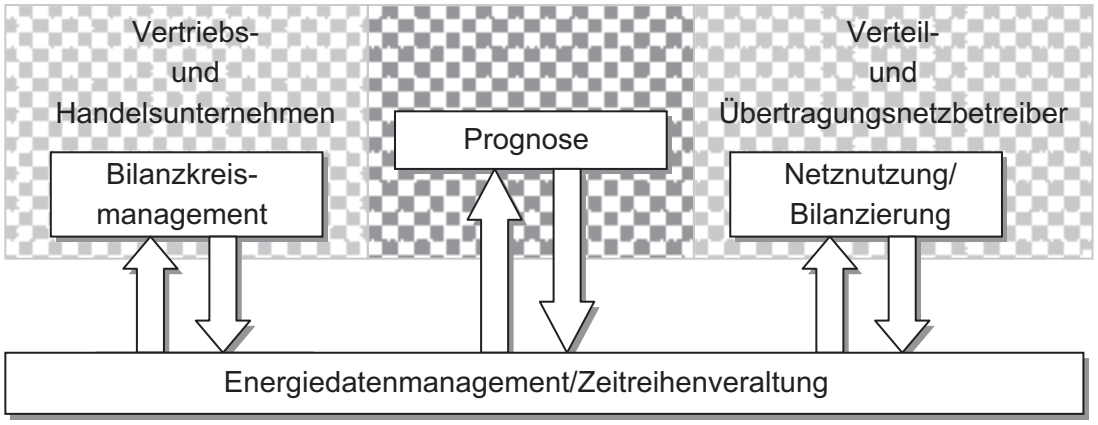


Abbildung 2-11: Aufbau eines Energiedatenmanagementsystems

Aufgrund der andauernden Weiterentwicklung des Energiewirtschaftsgesetzes und dessen Durchführungsverordnungen erfolgen Detaillierung und Standardisierung der Prozessabläufe, die in naher Zukunft eine kontinuierliche Systemweiterentwicklung erfordern. Ausschlaggebend für eine erfolgreiche Fortentwicklung wird sein, dass durch EDM-Systeme die Prozesse für die Kommunikation mit allen Marktpartnern vollständig automatisiert ablaufen und somit entstehende Prozesskosten reduziert werden.

3 Methodische Grundlagen der Zeitreihenanalyse

In diesem Kapitel werden die in der Literatur bekannten methodischen Grundlagen für die dann folgenden Untersuchungen dargestellt. Das Kapitel gliedert sich dabei in sieben Abschnitte. Einleitend werden die allgemeinen Grundlagen, Einflüsse und Einflussgrößen auf das elektrische Lastverhalten dargestellt. Anschließend werden statistische Verfahren, die Zeitreihenanalyse und stochastische Prozesse - die bereits heute im Einsatz befindlichen Modelle und die als Referenzmodelle fungieren - vorgestellt. Abschließend werden die theoretischen Grundlagen neuronaler Netze ausführlicher erläutert, da diese Verfahren primär im Fokus der Untersuchungen stehen.

3.1 Einflüsse und Einflussgrößen auf die elektrische Last

Die Gesamtlast des Energiebedarfs eines Kollektivs wird durch eine Vielzahl unterschiedlicher Faktoren beeinflusst. Der Beeinflussungsgrad hängt hierbei stark von der Zusammensetzung des Kollektivs als solches und des jeweilig eingesetzten Energieträgers ab. Die im Rahmen dieser Arbeit durchgeführten Untersuchungen beschränken sich primär auf den Anwendungsfall der elektrischen Energie. Da jedoch diese einen umfassenden Anwendungsbereich abdecken, können die Ergebnisse der Analyse der Einflussgrößen analog auf jeden anderen Energieträger Anwendung finden.

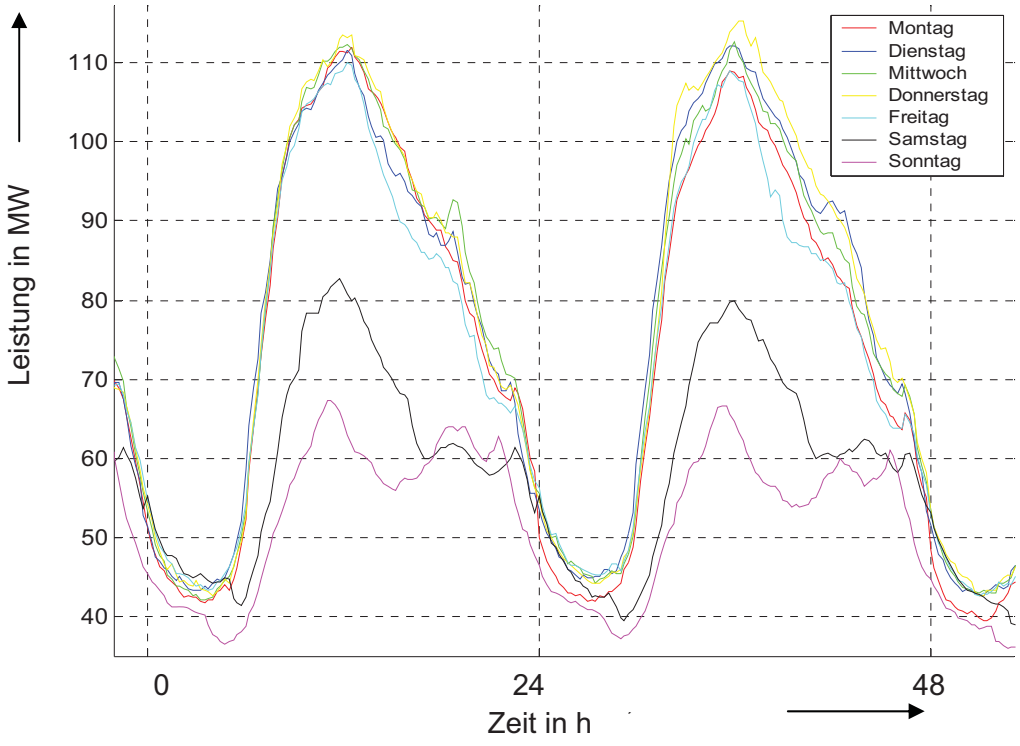


Abbildung 3-1: Elektrischer Lastgang einer Woche

Da die elektrische Last durch die Nachfrage bestimmt wird, besteht grundsätzlich eine kausale Beziehung zwischen dem Verbrauchsverhalten - welches auch durch exogene Einflussgrößen bestimmt werden kann und der Last, die für eine vollständige Beschreibung des Lastgangs zu ermitteln ist. Eine Einteilung einer exogenen Größe auf dieses Verhalten erscheint daher sinnvoll.

Langfristige Trends wie z. B. die Änderung der Demographie, des Versorgungsgebietes oder Veränderungen in der Verbraucherstruktur, z. B. durch Zu- und Abgang leistungsstarker Industriebetriebe, haben langdauernde Veränderungen der Gesamtlast zur Folge [Hau93]. Bei der Erstellung von Tages- und Wochenlastprognosen haben diese Einflüsse für das Ergebnis keine Bedeutung. Bei einem Vergleich der vorhergehenden Jahre haben diese Lastzuwächse Einfluss und sind bei der Ermittlung von langfristigen Trends zur regelmäßigen Aktualisierung des Prognosemodells unverzichtbar.

Die Last eines Kollektivs stellt das Verhalten einer Vielzahl verschiedener Verbraucher dar. Das stochastische Verbrauchsverhalten sowie die Wirkung unvorhergesehener Ereignisse wie Regenschauer, Gewitter, Bewölkung mit Dunkelheit etc. bewirken stets Zufallsschwankungen im Belastungsverlauf. Neben kleineren Netzstörungen bewirken Störungen großer Erzeuger oder Industriebetriebe Abweichungen von einer zuvor erstellten Prognose. Da nur wenige industrielle Verbraucher dem Lastverteiler ihre Fahrpläne im Voraus zur Verfügung stellen, können solche Verbraucher durch große, unprognostizierbare Schwankungen zur Prognoseunsicherheit beitragen. Populäre Ereignisse wie WM-Fußballspiele [Ber07] oder Großveranstaltungen können einen signifikanten Einfluss auf den Lastverlauf nehmen. Obwohl der Zeitpunkt solcher Ereignisse bekannt ist, ist deren Ausmaß nicht eindeutig prognostizierbar.

Die aktive Einflussnahme eines Energieversorgers oder Kunden auf den Verbrauch und somit auf die Lastganglinie wird als Laststeuerung bezeichnet. Diese Laststeuerung kann durch den Einsatz unabhängiger Erzeugungseinrichtungen sowie Lastreduzierung und Lastverteilung erfolgen. Die Zielsetzung der Maßnahmen erfolgt unter der Prämisse einer möglichst gleichmäßigen und dadurch wirtschaftlichen Betriebsweise des Netzes.

Der elektrische Lastgang hat eine Periodizität im Jahres-, Wochen-, und Tagesrhythmus. Feiertage und deren benachbarte Tage stören den zyklischen Wochengang. Der jahreszeitliche Gang der elektrischen Last ist im Wesentlichen auf folgende Faktoren zurückzuführen:

- Änderung der Temperatur im Laufe des Jahres, die eine starke Auswirkung auf den Heizbedarf hat
- Änderung der Tageslänge im Laufe des Jahres, die eine starke Auswirkung auf den Beleuchtungsbedarf hat

- soziale Faktoren, wie z. B. andere Verbrauchergewohnheiten in der Ferienzeit oder Saisonarbeit.

Saisonale Einflüsse bewirken daher nicht nur eine Änderung der Lasthöhe, sondern auch der Lastlinienform [Als95]. Abbildung 3-2 zeigt die Autokorrelationsfunktion der elektrischen Last einer Woche. Ein Einfluss der täglichen Zyklen resultiert aus dem Arbeits- und Freizeitverhalten des Versorgungskollektivs. Die Autokorrelationsfunktion zeigt hierbei, dass die Wochenperiode von 168 Stunden wie auch eine Grundperiode von 24 Stunden (Tagesperiode) vorgegeben ist. Die Periodizität von Tag zu Tag ist weniger stark ausgeprägt, da das Lastverhalten zwischen Wochenenden und Wochentag unterschiedlich ist. Unregelmäßigkeiten treten in diesen Zyklen durch Feier- und Brückentage auf. Bedingt durch das veränderte Verbraucherverhalten werden Höhe und Form der Lastganglinie beeinflusst. Zusätzlich wird Einfluss auf die umliegenden Tage genommen, so dass auch hier Unregelmäßigkeiten im Lastverlauf zu Vergleichstagen zu beobachten sind.

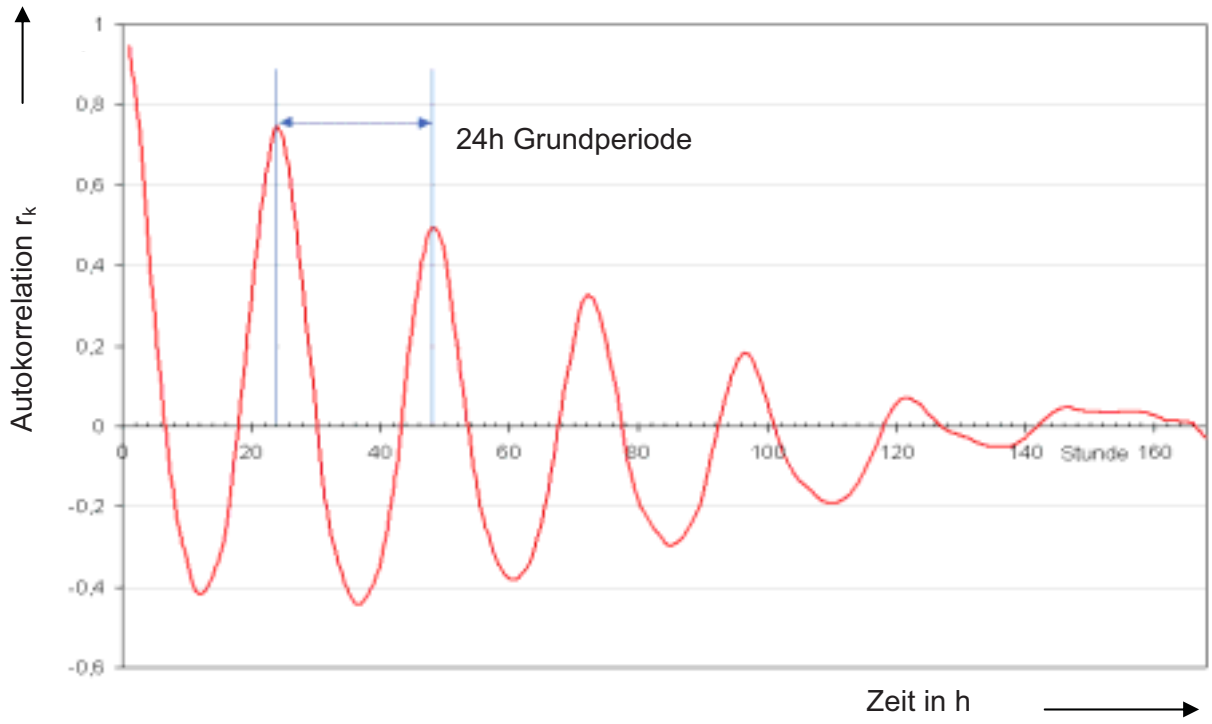


Abbildung 3-2: Autokorrelation des Wochenlastganges

Es gibt eine große Anzahl wetterempfindlicher Verbraucher wie Heizungs- und Beleuchtungsanlagen, Haushaltsgeräte, Klimaanlagen usw. Daher haben meteorologische Größen Einfluss auf den Belastungsverlauf [Hau93]. Eine Analyse des Wettereinflusses auf den elektrischen Energie- und Gasverbrauch wurde in verschiedenen Untersuchungen bereits mehrfach vorgenommen [Dav59], [Asb75], [Gei01]. Der Einfluss der Temperatur ist hierbei die wichtigste Wettergröße, die die Last beeinflusst. Des Weiteren ist die Belastung nicht nur von der momentanen Temperatur abhängig, sondern wird auch stark vom Temperaturverlauf der vorangegangenen Tage

geprägt [Gei01]. Durch Wärmespeicherung der Gebäude, die je nach Art und Bauweise zwischen acht und 50 Stunden andauern kann, wird die Belastung verzögert.

Der Einfluss der Temperatur auf den Lastverlauf ist im Allgemeinen nicht linear. Steigende Temperaturen bewirken im Winter einen Rückgang des Verbrauchs aufgrund sinkender Heizlast und bei Netzbetreibern mit hohem Anteil an Klimageräten einen erhöhten Verbrauch an elektrischer Energie im Sommer.

Die Abhängigkeit der verschiedenen Wettergrößen auf die Belastung lässt sich nur schwer analysieren, da stets eine Abhängigkeit zwischen den einzelnen Wettergrößen besteht. Untersuchungen haben gezeigt, dass ein kräftiger Wind in Verbindung mit niedrigen Temperaturen den gleichen abkühlenden Effekt hat, wie eine um viele Grad niedrigere Temperatur bei Windstille [Huf97]. Beispielsweise nimmt die Windgeschwindigkeit bei +20 ^{o}C fast keinen Einfluss auf die Belastung, während bei -20 ^{o}C der Einfluss erheblich ist. Zusätzlich ist das Empfinden des Menschen ein wichtiger Einflussfaktor auf die Belastung. Die Angabe der reinen Lufttemperatur stimmt häufig nicht mit der gefühlten Temperatur überein. Schon bei einer geringen Luftbewegung kann das Gefühl entstehen, es sei deutlich kälter als es der tatsächlichen Lufttemperatur entspricht.

Als weitere Einflussgröße auf den Energieverbrauch ist das Zusammenwirken von Luftfeuchtigkeit und Temperatur zu nennen. Eine hohe Lufttemperatur in Verbindung mit hoher Luftfeuchtigkeit führt zum verstärkten Einsatz von Klima- und Ventilationsgeräten. Weiterhin hat die Helligkeit einen großen Einfluss auf die elektrische Last. Gerade an Frühlingstagen führt die direkte Sonneneinstrahlung zur Aufwärmung der Gebäude und verringert somit ihren Energiebedarf. Im Artikel von [Mar85] werden zwei Gruppen von Einflussgrößen identifiziert, die bei der Erstellung von Lastprognosen von Bedeutung sind: Einflüsse, die das Verbrauchsverhalten beschreiben, und Witterungseinflüsse. In der Veröffentlichung von [PIR96] unterteilt dieser die Einflussgrößen in vier Gruppen: Allgemeine Einflüsse (Jahreszeiten und Wochentag), Wettereinflüsse, funktionalbedingte Einflüsse und autoregressive Einflüsse. Auch [Als95] identifiziert diese Einflüsse, ergänzt jedoch die bereits benannten noch um die Effekte der Laststeuerung. Daher erscheint zusammenfassend eine Einteilung der Einflüsse in drei Gruppen als sinnvoll:

1. Verbrauchsverhalten: Das Verbrauchsverhalten lässt sich vor allem aus den kalendarischen Daten (Wochen-, Feier-, Brückentage, Ferien, Jahreszeiten etc.) und dem entsprechenden Verhalten der jeweiligen Verbrauchergruppe ableiten. Längerfristige Einflüsse wie beispielsweise veränderte industrielle Produktionszeiten oder verlängerte Ladenöffnungszeiten können bei der kurzfristigen Lastprognose nicht berücksichtigt werden. Diese Einflüsse können lediglich zur Korrektur von Vergangenheitswerten bei der Prognoseerstellung herangezogen werden [Mar85].

Verbraucherverhalten	**Einflussgrößenbeispiel**
- Jahreszeit	
Saison	Winter, Sommer und Übergangszeiten
Monat	Januar, Februar, März.....
Tag/Nacht	Sonnenauf- und Untergang/Tageslichtdauer
Zeit	Sommer- und Winterzeit
- charakteristische Tage	
Wochentagtyp	Montag, Dienstag....
Tagestyp	Wochentag, Wochenenden
Sondertage	Brückentage, Feiertage, Ferientage
- Kundenverhalten	
Privatkunden	Verhalten, Einschaltquote
Gewerbekunden/Handel	Ladenöffnungszeiten, verkaufsoffene Tage
Großkunden/Industrie	Schichtzeiten, Auslastung der Produktion, Betriebsversammlungen, schaltbare Großlasten
Sonderabnehmer	Jahrmärkte, Stadien, Öffnungszeiten, Zeiträume der Veranstaltungen
- Trends	
Konjunktur	Bruttoinlandsprodukt oder andere Indikatoren
Struktur des Kollektivs	Zusammensetzung des Kollektivs (ländlich/ städtisch oder gemischte Abnahmestruktur)
Innovation	Veränderung aufgrund technischer Neuerungen
Witterungseinflüsse	**Einflussgrößenbeispiel**
- Temperatur	Boden und Lufttemperatur, Stunden- , Tagesdurchschnitts-, Dreitagesmittelwerte, Maximum, Minimum, Gradient
- Globalstrahlung	Tagesmittelwert
- Bedeckungsgrad	Differenz zwischen max. und min. Tagestemperatur sowie tatsächlicher und möglicher Globalstrahlung
- Niederschläge	Niederschlagsmengen, Dauer der Niederschläge
- Luftdruck und Luftfeuchtigkeit	Stundenwerte, Tagesmittel, Maximum und Minimum, Gradient
- Wind	Windgeschwindigkeit und –richtung, Tagesmittel- und Stundenwert, Maximum, Minimum
Autoregressive Zusammenhänge	**Einflussgrößenbeispiel**
- Vergangenheitswerte der Leistung	Leistung vor einem Tag, einer Woche, einem Jahr sowie Maximal- und Minimallast des Vortages, der Vorwoche, des Vorjahres
- Vergangenheitswerte der Arbeit	Arbeitsmengen des Vortages, der Vorwoche oder des Vorjahres

Tabelle 3-1: Klassifizierung von Einflussgrößen auf die elektrische Last

2. Autoregressive Zusammenhänge: Diese Zusammenhänge bestehen zwischen Lasten zu verschiedenen Zeitpunkten. Diese Zeitpunkte sind bei der elektrischen Last die Tages-, Wochen- und Jahresperiode.
3. Witterungseinflüsse: Die Witterungseinflüsse auf das Verbrauchsverhalten sind in Abhängigkeit des jeweiligen Parameters unterschiedlich stark ausgeprägt.

Die innerhalb der aufgeführten drei Hauptgruppen identifizierten Einflussgrößen auf die elektrische Last werden in Tabelle 3-1 dargestellt, erläutert und in exemplarischen Beispielen veranschaulicht [Huf00]. Bei der Berücksichtigung dieser exogenen Einflussgrößen für die Erstellung von Lastprognosen muss ein quantitativer Zusammenhang zwischen Einfluss- und Prognosegröße bestehen. Hierzu ist es erforderlich, dass einzubeziehende Einflussfaktoren folgende Bedingungen erfüllen [Edw81]:

- sie müssen qualitativ beschreibbar sein,
- zwischen Last- und Einflussgröße muss eine starke Korrelation bestehen,
- sie müssen als absolute und relative Größe messbar sein,
- die Einflussgrößen müssen quantitativ gut prognostizierbar sein.

Im Rahmen der folgenden Untersuchungen werden Methoden entwickelt und angewendet, um die Bedeutung der Einflussgrößen qualitativ und quantitativ zu beschreiben. Gemäß der Darstellung der einzelnen Einflussgrößen auf die elektrische Last kann die zu prognostizierende Last dabei in einen stochastischen und deterministischen Prozessanteil zerlegt werden. Damit ist die Höhe des stochastischen Anteils maßgeblich für die Größe des Gesamtfehlers und setzt dadurch die Grenze für die Prognosegenauigkeit.

3.2 Zeitreihenprognosen

Prognosen finden in allen Bereichen des täglichen Lebens Anwendung. Zur Lösung von Planungsproblemen werden oft Kenntnisse von künftigen Größen benötigt. Liegen Daten aus der Vergangenheit in Form von Zeitreihen vor, so können diese auf Gesetzmäßigkeiten hin untersucht werden und unter der Annahme, dass diese auch weiterhin gelten, können Vorhersagen für die Zukunft gemacht werden. Die Statistik stellt hierfür ein umfangreiches Instrumentarium bereit, das je nach dem zugrunde liegenden Zeitreihenmodell angewendet werden kann.

Die Prognose trifft somit Aussagen über ein oder mehrere künftige Ereignisse, die entweder auf Beobachtung oder auf einer Theorie beruhen. Die Theorie für ein künftiges Ereignis erfordert eine sachlogische Begründung der Prognose und Planungsprämissen, unter denen sie abgegeben wird. Diese Theorie muss nicht unbedingt umfangreich oder detailliert ausgearbeitet sein [Han83]. Ein Prognosemodell ist dabei ein System, welches beobachtete Werte der zu prognostizierenden Variablen und möglicher anderer Variablen als Elemente enthält und nach bestimmten Regeln

miteinander verknüpft, um als Ergebnis der Verknüpfung Prognosewerte zu erhalten [Thi98]. Die Vielfalt verschiedener Prognosemodelle lässt sich durch Klassifizierung auf eine überschaubare Anzahl von Modelltypen reduzieren, deren wichtigste im Folgenden gegenübergestellt werden.

Grundsätzlich wird zwischen qualitativen und quantitativen Prognosetechniken differenziert. Qualitative Prognosetechniken sind subjektive Einschätzungen, die von Experten mit einem gereiften Fachwissen intuitiv erstellt werden. Dabei ist eine mögliche Variante die lineare Extrapolation: Vergangenheitswerte werden grob in die Zukunft projiziert. Weitere Varianten hierzu sind Meinungsbefragungen oder Lebenszyklusanalysen mit dem Ergebnis, Trends vorherzusagen. Quantitative Prognosetechniken hingegen bestehen hauptsächlich aus der mathematischen Aufarbeitung von Datenmaterial der Vergangenheit und geben konkrete, zahlenmäßige Resultate für die Zukunft.

Darüber hinaus unterscheidet man univariate und multivariate Prognosemodelle. Bei einer univariaten Zeitreihenprognose liegt für jeden Zeitpunkt eine einzige Beobachtung vor, es wird also nur ein Merkmal betrachtet. Die Prognose des künftigen Systemverhaltens wird anhand der vorliegenden (vergangenen und aktuellen) Beobachtungen ermittelt. Dem liegt die Annahme zugrunde, dass die vorliegenden Beobachtungen über die Zeit hinweg eine interne Struktur oder Regelmäßigkeit besitzen, welche durch die Analyse herausgearbeitet und durch die Prognose genutzt werden soll [Pri05]. Üblicherweise werden die letzten n-Beobachtungen als Stützbereich benutzt, um die nächste, nicht mehr vorliegende Beobachtung zu prognostizieren. Bei multivariaten Zeitreihenprognosen betrachtet man nicht eine Variable isoliert sondern das Zusammenwirken mehrerer Variablen zugleich und damit ihre Abhängigkeitsstruktur. Damit liegen für jeden Zeitpunkt mehrere Beobachtungen oder verschiedene Merkmale vor, die zu prognostizierende endogene Zeitreihe und die erklärenden exogenen Zeitreihen. Der Ansatz der multivariaten Zeitreihe versucht auch der Tatsache gerecht zu werden, dass die meisten Prozesse nicht isoliert auftreten und auch nicht isoliert prognostizierbar sind, sondern nur als Menge mehrerer miteinander in Abhängigkeit stehender Prozesse.

Ein weiteres wichtiges Kriterium zur Klassifikation von Zeitreihenprognosen ist der Zeithorizont, über den diese Modelle künftige Aussagen treffen. Üblicherweise wird hierbei zwischen kurz-, mittel- und langfristigen Modellen differenziert [Mak98]. Jedoch ist die Klassifizierung einerseits vom jeweiligen Anwendungsgebiet, andererseits auch vom zeitlichen Abstand der vorliegenden Beobachtungs- und Vorhersagewerte (Stunde, Tag, Woche, Monat, etc.) abhängig. Eine Klassifizierung von Prognosehorizonten der Energiewirtschaft erfolgte bereits in Kapitel 2.2.

3.3 Prozess der Prognoseerstellung

Eine systematische Entwicklung von Prognosen benötigt ein Modell, welches weitgehend unabhängig von der Lösung und den Kenntnissen des Bearbeiters ist. Die Prognoseerstellung ist ein festgelegter Prozess, der in verschiedenen Schritte durchgeführt wird. Unter dem Begriff sicentific forecasting wird bei [Gil77] in einer kompakten und übersichtlichen Abfolge dieser Prozess in vier Teilschritte gegliedert, die in Abbildung 3-3 dargestellt sind.

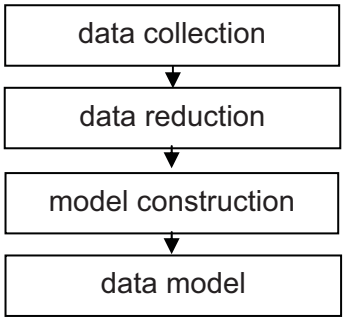


Abbildung 3-3: Scientific forecasting

Im Rahmen des ersten Hauptschrittes der „data collection“ werden alle für die Prognoseerstellung relevanten Daten gesammelt und für die weitere Analyse aufbereitet. Das heißt, es erfolgt eine Plausibilisierung und eine Ersatzwertbildung der vorliegenden Daten. Der dann folgende entscheidende Schritt der „data reduction“ identifiziert die Datenbestände, die höchste Relevanz für die Prognoseerstellung aufweisen. Dabei erfolgt eine Zerlegung jedes einzelnen Datensatzes in das Niveau, den Trend, die Zyklizität, die Saisonalität, die Extrasaisonalität, die Sondereinflüsse und die irregulären Komponenten der Datenbestände.

Nach Analyse aller für die Prognose infrage kommenden Datenbestände besteht in diesem Schritt die Möglichkeit, die zunächst als wenig relevant eingestuften Datenbestände zu verwerfen und diese bei einem weiteren Prozessdurchlauf zu berücksichtigen.

Im Schritt der ‚model construction’ werden die zuvor als relevant klassifizierten Daten in einer Struktur mit der zu prognostizierenden Größe verbunden. Innerhalb dieses Schrittes erfolgt eine Vorverarbeitung der Daten. Abschließend werden die Zusammenhänge zwischen diesen in einem mathematischen Modell spezifiziert. Dieser Schritt lässt sich in drei Teilschritte verfeinern, die in Abbildung 3-4 dargestellt sind [Joe96].

Der Teilprozessschritt „Indikatorbildung“ beinhaltet dabei die Transformation von Zeitreihen, die im Hauptschritt „data reduction“ als Größen mit hohem Einfluss für das Prognoseergebnis eingestuft wurden. Dabei werden die für die Prognose relevanten

Datenreihen als Indikatoren bezeichnet, die als Input für das Prognosemodell Anwendung finden und dieses Modell in geeigneter Form beschreiben.

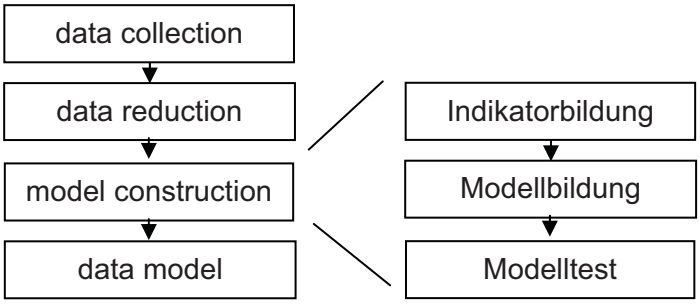


Abbildung 3-4: Model construction

Im Schritt Modellbildung wird das Prognosemodell dann entwickelt, welches die Indikatoren als Inputgrößen verwendet und zu einer Prognose verarbeitet. Beim folgenden Modelltest werden auf Basis von Vergangenheitswerten, die dem Modell als Inputwerte zur Verfügung gestellt werden, erwartete Outputwerte der Vergangenheit gegenübergestellt. Dabei werden Abschnitte von Datenreihen angewendet, die zur Modellbildung bisher nicht herangezogen wurden.

Für den letzten Hauptschritt der „model extrapolation“ wird das Modell aus dem vorherigen Schritt angewendet. Durch Extrapolation in die Vergangenheit festgestellte Gesetzmäßigkeiten - die durch relevante Daten repräsentiert werden - können dann Prognosen für die Zukunft erstellt werden.

3.4 Beurteilungsmaß und Qualität von Prognosen

Das Ziel einer statistischen Modellierung natürlicher Systeme besteht darin, Prognosen über künftige Entwicklungen abzugeben. Dabei besteht zum einen die Möglichkeit, einen Vorhersagewert unter ceteris-paribus-Annahme zu prognostizieren, der am wahrscheinlichsten erscheint. Es ist auch möglich, unter gleichen Annahmen Vorhersagen innerhalb eines Prognoseintervalls zu treffen, in dem der Wert mit einer gegebenen Wahrscheinlichkeit liegen kann. Vorhersagen enthalten in spezifizierten oder geschätzten Modellen erhebliche Ungewissheiten aufgrund der Störterme des verwendeten Modells oder Varianzen der geschätzten Eingangsparameter. Die Prognoseunsicherheit, die zwangsläufig in jedem Modell implementiert ist, nimmt somit erheblichen Einfluss auf das Prognoseintervall [And97].

Aussagen zur Prognosequalität lassen sich in ex-ante- und ex-post-Beurteilungen klassifizieren [Thi98]. Bei den ex-ante-Beurteilungen erfolgt eine Überprüfung der Prognose, ohne dass konkrete Vergangenheitswerte vorliegen. Die Beurteilung erfolgt aufgrund der Analyse der kausalen Variablen, dem Prognosezeitraum sowie der Funktionsform. Die abschließende Beurteilung der Qualität erfolgt durch Vergleich der

prognostizierten mit den tatsächlichen Werten und kann als ex-post-Beurteilung bezeichnet werden. Der Schwerpunkt einer ex-post-Beurteilung von Prognosen liegt in der Quantifizierung des Fehlers eines Erwartungswertes, wobei unterschiedliche statistische Verfahren verwendet werden können.

3.4.1 Ursachen von Lastprognosefehlern

Beim prognostizierten Lastgang sind natürliche Fehler zu erwarten, die sich in deterministische und stochastische Fehleranteile zerlegen lassen [Als95]:

$$E = E_{det} + E_{stoch} \tag{3-1}$$

Stochastische Fehleranteile E_{stoch} lassen sich nur zu einem geringen Anteil eliminieren, insbesondere bei der Lastprognose, da das Verhalten der Verbraucher und damit die Schwankung der Lastgänge zu einem erheblichen Teil dem Zufall unterliegen. Diese Zufallsschwankungen setzen somit eine Grenze für die Prognosegenauigkeit, die auch unter Berücksichtigung aller Einflussfaktoren nicht unterschritten werden kann. Die Größe des deterministischen Fehleranteils E_{det} wird ausschließlich durch die auf das System einwirkenden tatsächlichen und relevanten Einflussgrößen sowie Modellierungsfehler bestimmt. Die deterministischen Fehlereinflussfaktoren lassen sich in folgende Arten klassifizieren [Bra02]:

- Mangelhafte Qualität der Vorhersage meteorologischer Einflussgrößen auf den Lastgang.
- Fehlerbehaftete Datenbasis der gemessenen Wettergrößen sowie der gemessenen Lastverläufe.
- Modellierungsfehler.
- Inkonsistente Modellierung zwischen Einfluss- und Prognosegrößen.
- Mangelnde Datenclusterung.

3.4.2 Fehlerkenngrößen

Die Güte eines Prognosemodells wird gemessen, indem für einen Datensatz mit bekannter zu prognostizierender Größe die Vorhersage des Modells mit der tatsächlichen Größe in Beziehung gesetzt wird. Die ex-post ermittelte Prognosequalität wird überwiegend statistisch gemessen, wobei der Prognosefehler im Mittelpunkt der Betrachtung steht. Der Prognosefehler (wahrer Fehler WE) ist die Differenz zwischen der prognostizierten Last (P_{prog}) und der tatsächlichen Last (P_{tats}) an einem prognostizierten Zeitpunkt. Er kann bei einer zu hoch prognostizierten Last ein negatives, bei Lastunterschätzung ein positives Vorzeichen annehmen.

$$WE_{(t)} = P_{tats(t)} - P_{prog(t)} \text{ mit } t = T+1,...,T+n \tag{3-2}$$

Da in den meisten Fällen eine Betrachtungsschranke zur Bewertung der Prognosegüte erforderlich ist, wird der Fehlerbetrag zu Hilfe genommen. Dieser ist als absoluter Fehler (AE) bezeichnet und ist wie folgt definiert:

$$AE_{(t)} = \left|WE_{(t)}\right| = \left|P_{tats(t)} - P_{prog(t)}\right| \tag{3-3}$$

Da die Betrachtungsschranke zur Beurteilung der Güte maßgeblich ist, wird ebenfalls der Fehlerbetrag zu Hilfe genommen. Dieser wird als absoluter prozentualer Prognosefehler (APE) bezeichnet und ist wie folgt definiert:

$$APE_{(t)} = \frac{100}{P_{tats(t)}} \cdot \left|P_{tats(t)} - P_{prog(t)}\right| = \frac{100}{P_{tats(t)}} \cdot \left|WE_{(t)}\right| \tag{3-4}$$

Der Bezug des Prognosefehlers auf die tatsächlich aufgetretene Last ist der relative Prognosefehler (RE):

$$RE_{(t)} = \frac{P_{tats(t)} - P_{prog(t)}}{P_{tats(t)}} = \frac{WE_{(t)}}{P_{tats(t)}} \tag{3-5}$$

Um auch diesen mit einer Bewertungsschranke zu beurteilen, kann der relative absolute Prognosefehler (RAE) wie folgt angegeben werden:

$$RAE_{(t)} = \frac{\left|P_{tats(t)} - P_{prog(t)}\right|}{P_{tats(t)}} = \frac{\left|WE_{(t)}\right|}{P_{tats(t)}} = \frac{AE_{(t)}}{P_{tats(t)}} \tag{3-6}$$

Die Konstruktion der statistischen Fehlerkennmaße erfolgt nun im Folgenden auf Basis der definierten Prognosefehler. In der Literatur ist hierzu eine Vielzahl von Kenngrößen definiert [Han83] [Hüt86] [Sch89] [Sch89] [Web90] [Zöf01]. Im Rahmen der Arbeit werden nur die relevanten Kenngrößen vorgestellt und nachstehend näher erläutert. Der Fehlerbetrag der in einem Prognoseintervall i=1...n ermittelten Werte wird als mittlerer Prognosefehler (ME) bezeichnet und ist wie folgt definiert:

$$ME = \frac{1}{n} \cdot \sum_{t=T+1}^{T+n} WE_{(t)} = \frac{1}{n} \cdot \sum_{i=T+1}^{T+n} \left(P_{tats(t)} - P_{prog(t)}\right) \tag{3-7}$$

Der hieraus abgeleitete mittlere absolute Prognosefehler (MAE) lässt sich wie folgt beschreiben:

$$MAE = \frac{1}{n} \cdot \sum_{t=T+1}^{T+n} AE_{(t)} = \frac{1}{n} \cdot \sum_{t=T+1}^{T+n} \left|P_{tats(t)} - P_{prog(t)}\right| \tag{3-8}$$

Der mittlere relative Prognosefehler (MRE) kann wie folgt beschrieben werden:

$$MRE = \frac{1}{n} \cdot \sum_{t=T+1}^{T+n} \frac{100 \cdot AE_{(t)}}{P_{tats(t)}} = \frac{1}{n} \cdot \sum_{t=T+1}^{T+n} \frac{\left|P_{tats(t)} - P_{prog(t)}\right|}{P_{tats(t)}} \cdot 100 \tag{3-9}$$

Um im Rahmen von Untersuchungen mit Hilfe einer Bewertungsschranke Gütevergleiche zu ermöglichen, ist der mittlere absolute prozentuale Prognosefehler (MAPE) wie folgt definiert:

$$\mathrm{MAPE} = \frac{1}{n} \cdot \sum_{t=T+1}^{T+n} \frac{AE_{(t)}}{\left|P_{tats(t)}\right|} = \frac{1}{n} \cdot \sum_{t=T+1}^{T+n} \frac{\left|P_{tats(t)} - P_{prog(t)}\right|}{\left|P_{tats(t)}\right|} \tag{3-10}$$

Summe der quadratischen Prognosefehler (SSE):

$$\mathrm{SSE} = \sum_{t=T+1}^{T+n} WE_{(t)}^2 = \sum_{t=T+1}^{T+n} \left(P_{tats(t)} - P_{prog(t)}\right)^2 \tag{3-11}$$

Mittlerer quadratischer Prognosefehler (MSE):

$$\mathrm{MSE} = \frac{1}{n} \cdot \sum_{t=T+1}^{T+n} WE_{(t)}^2 = \frac{1}{n} \cdot \sum_{t=T+1}^{T+n} \left(P_{tats(t)} - P_{prog(t)}\right)^2 \tag{3-12}$$

Wurzel des mittleren quadratischen Prognosefehlers (RMSE):

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}} = \sqrt{\frac{1}{n} \cdot \sum_{t=T+1}^{T+n} WE_{w(t)}^2} = \sqrt{\frac{1}{n} \cdot \sum_{t=T+1}^{T+n} \left(P_{tats(t)} - P_{prog(t)}\right)^2} \tag{3-13}$$

Bei der Ermittlung der quadratischen Fehler nehmen Abweichungen des prognostizierten Wertes zum Istwert aufgrund ihrer Gewichtung größeren Einfluss auf den Fehlerbetrag. Diese Bewertung ist bei einer zeitkontinuierlichen Abweichung gerechtfertigt und erwünscht. Beruht die Differenzen zwischen Prognose- und Istwert jedoch auf singulären Störungen, ist die Bewertung ungerechtfertigt.

3.5 Zeitreihenanalyse

3.5.1 Allgemein

Zeitreihen begegnen uns in den verschiedensten Bereichen des täglichen Lebens, wo Grafiken mit zeitlichen Verläufen von Anzahlen präsentiert werden. Daher ist es kaum verwunderlich, dass gerade in den letzten Jahrzehnten die Anzahl der hierfür entwickelten Methoden stark gewachsen ist [Sti01]. Die Zeitreihenanalyse ist eine der wichtigsten Anwendungsgebiete im Bereich der statistischen Methoden. Einfache deskriptive (heuristische) Methoden wie Graphiken, Vorjahresvergleiche, Berechnung von absoluten Zuwächsen oder Differenzen oder relativen Zuwächsen (Wachstumsraten) sowie Autokorrelationen usw. können hilfreich sein zur Beschreibung der Daten und für die Suche nach einem adäquaten Modell für das Zustandekommen dieser Daten. Weitergehende Erklärungen und auch Prognosen setzen im Regelfall eine Modellbildung voraus.

Ganz allgemein kann dabei eine Zeitreihe als eine zeitlich geordnete Abfolge von Beobachtungswerten eines einheitlichen Merkmals von statistischen Massen definiert werden. Die statistische Zeitreihenanalyse befasst sich mit diesen Beobachtungswerten, die nach bestimmten Merkmalen erhoben wurden, wobei eines dieser Merkmale die Zeit ist. Sie ist damit eine besondere Form der Regressionsanalyse, bei der die Analyse mit folgender Zeitsetzung erfolgt [Kar05]:

- Deskription: Die Beschreibung von Zeitreihen erstreckt sich über die Berechnung der empirischen Momente über die Darstellung der Zeitreihe als Komponentenmodell bis zu der Bestimmung der einzelnen Komponenten. Ziel der Deskription ist die Entdeckung elementarer Charakteristika der untersuchten Zeitreihe.
- Modellierung: Bei der Modellierung betrachtet man die Zeitreihe als Realisation eines stochastischen Prozesses. Aufgabe der Modellierung ist die Bestimmung der Parameter der zugrunde gelegten stochastischen Modelle sowie deren Wirkung.
- Prognose: Für die Prognose von Werten ist die Gültigkeit des angepassten Modells eine zentrale Annahme, wobei die zu prognostizierenden Werte durch die Extrapolation der angepassten Modelle in die Zukunft erzeugt werden.
- Kontrolle: Die Analyse der Zeitreihe identifiziert deren charakteristische Parameter. Durch die Steuerung dieser Parameter kann das Verfahren des untersuchten Systems kontrolliert werden. Oft ist dies jedoch nicht oder nur schwer möglich bei der Analyse von Phänomenen wie beispielsweise Sonnenflecken oder Vulkanausbrüchen. In diesem Fall liefern die gewonnen Informationen eine Vorbereitung auf bevorstehende Ereignisse.

Bei der Erstellung der Prognose wird jedoch vorausgesetzt, dass die Bestimmung der Struktur der Zeitreihe erfüllt werden kann. Wichtig für die Bestimmung der Struktur einer Zeitreihe ist, dass sich innerhalb dieser die Beobachtungen auf statistische Massen beziehen müssen, die sich lediglich im zeitlichen Merkmal unterscheiden, da ansonsten die Vergleichbarkeit der Zeitreihenwerte untereinander erschwert oder - im Extremfall - gar nicht möglich ist. Der Fokus dieser Arbeit beschränkt sich auf die Analyse von energiewirtschaftlichen Zeitreihen und damit auf äquidistante Größen, also Zeitreihen mit gleicher zeitlicher Entfernung zwischen den Zeitindizes. Jeweils nach Art des Modells lassen sich verschiedene Zeitreihenanalyseverfahren unterscheiden, deren Klassifikation in Abbildung 3-5 dargestellt ist.

Es wird hierbei zwischen univariaten und multivariaten Verfahren unterschieden, je nachdem ob die zu beschreibende Zeitreihe y_t ein- oder mehrdimensional oder ein Vektor mit den Daten y_{t1}, y_{t2} ist. Dabei werden innere Verfahren immer nur univariaten Analyseverfahren zugeordnet. Innere Verfahren erklären eine Zeitreihe y_t allein als Funktion der Zeit oder früherer Werte y_{t-d} einer gleichen Zeitreihe. Äußere Methoden ziehen auch andere Variablen - x_t, x_{t-d}, z_t usw. - zur Erklärung der Ausgangsfunktion heran. Zu den inneren Verfahren gehören auch die weit verbreiteten "Box-Jenkins-

Methoden" der Zeitreihenanalyse, auf die im Rahmen der vorliegenden Arbeit nicht weiter eingegangen werden soll. Zu den äußeren Verfahren zählen statistische Testverfahren sowie Regression und Korrelationsanalyse. Die inneren Verfahren werden wiederum unterteilt in die Gruppe der Komponentenmodelle sowie der Fourieranalyse-Methoden. Dabei interpretieren die Komponentenmodelle eine Zeitreihe y_t als Überlagerung einfacher Funktionen der Zeit, die formal aufgrund ihrer Periodizität definiert sind und "Komponenten" genannt werden. Hierauf beruhende Verfahren der Zeitreihenanalyse eliminieren sukzessive die einzelnen Komponenten wie Trend-, Zyklus- und weitere Komponenten, wobei die dann verbleibende "Restkomponente" r_t meist als verfahrensbedingter Rest übrig bleibt.

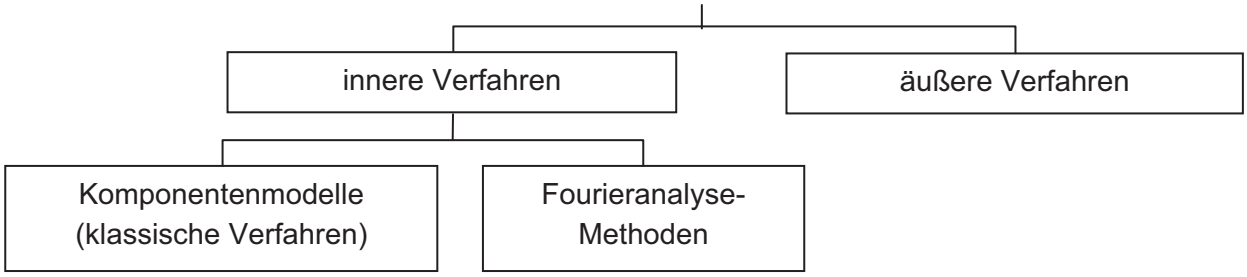


Abbildung 3-5 : Übersicht von Zeitreihenanalyseverfahren

Die Gruppe der Fourieranalyse-Methoden, zu denen unter anderem auch die harmonische Analyse, Spektralanalyse gezählt werden können, ist eine Alternative zum "klassischen Verfahren“. Die Fourieranalyse betrachtet eine Zeitreihe, mit Ausnahme der nichtzyklischen (monotonen) Trendkomponente, als Summe vieler Schwingungen unterschiedlicher Frequenz und Amplitude, die "direkt" (simultan) geschätzt werden. Somit kann ein beliebiges periodisches Signal durch eine Summe von Sinus- und Kosinusfunktionen (eine so genannte Fourier-Reihe) erzeugt werden und eine Zerlegung des Signals in Frequenzanteile ist möglich.

3.5.2 Darstellung von Zeitreihen

Bei der Untersuchung von Zeitreihen empfiehlt es sich, grundsätzlich mit einer grafischen Darstellung zu beginnen, um schon vor Beginn der Analyse rein deskriptiv erkennbare Eigenschaften der Zeitreihe zu erkennen. Um hierbei interessante Merkmale zu detektieren, ist es sinnvoll, Darstellungsform wie auch Skalierung zu variieren.

Eine Zeitreihe ist die zeitliche genormte Folge von beobachteten Werten. Diese kann in der Form $(y_t : t = 1,....,n)$ oder $(y_{(t_i)} : i = 1,....,n)$ beschrieben werden. Diese Beschreibung setzt jedoch voraus, dass der Abstand zwischen den Messgrößen äquidistant ist. Im Rahmen von Untersuchungen elektrischer Lastprofile ist der Zeitraum zwischen den

Messpunkten aufgrund des ¼-h-Abrechnungsrasters des Lastgangs äquidistant. Die Visualisierung der Zeitreihe kann dabei in Abhängigkeit des Analyseziels in unterschiedlichen Darstellungsformen erfolgen. Der Zeitpfad der Reihe kann mit linearer (Abbildung 3-1) oder logarithmischer Skalierung der Ordinate erfolgen. Bei einem frequenzbezogenen Analyseverfahren hingegen werden Merkmalswerte der Zeitreihe in Abhängigkeit von Frequenzen beobachtet und nicht in deren zeitlicher Abfolge. Der Einsatz dieser Verfahren ist erst dann nützlich, wenn eine Zeitreihe Periodizitäten aufweist. Ziel dieser Darstellungsform ist, die Intensität von Schwingungen in Abhängigkeit ihrer zugehörigen Frequenzen darzustellen, exemplarisch ist in Abbildung 3-6 das Periodogramm eines elektrischen Lastgangs dargestellt.

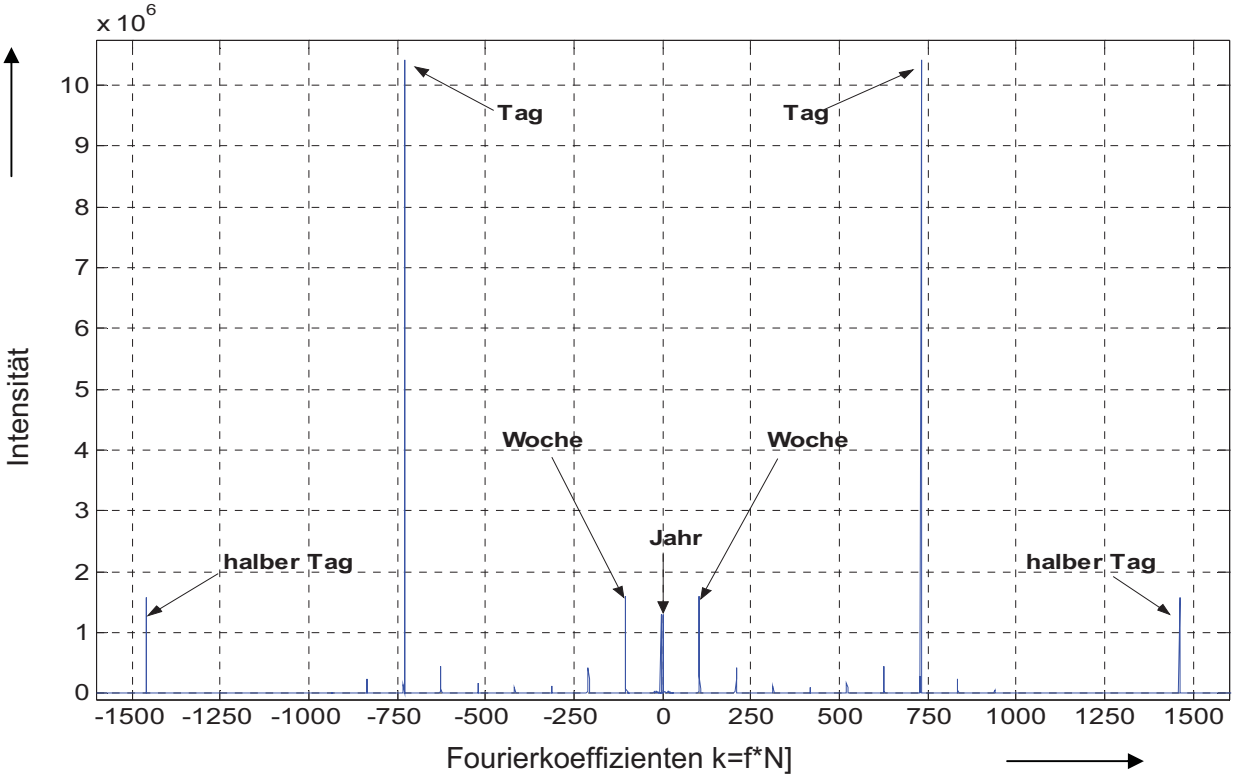


Abbildung 3-6: Periodogramm von Wochentagen der elektrischen Last

Die Abbildung 3-6 zeigt die hauptsächlich vorkommenden periodischen Anteile der Lastgangzeitreihen zweier Jahre. Es handelt sich dabei um die zu einem Tag, einem halben Tag und einer Woche zugehörigen Frequenzen. Beim Korrelogramm hingegen erfolgt die graphische Darstellung der Autokorrelation einer Zeitreihe. Exemplarisch ist das Korrelogramm eines elektrischen Lastganges in Abbildung 3-2 dargestellt. Hierbei werden die Korrelationskoeffizienten gegen die Dauer der Zeitverschiebung aufgetragen. In einem statistischen Modell wird grundsätzlich die geordneten Folge von Zufallsvariablen vorausgesetzt. Vergleicht man deren Abfolge mit sich selbst, so spricht man von Autokorrelation. Da jede unverschobene Folge mit sich selbst am ähnlichsten ist, hat die Autokorrelation für die unverschobenen Folgen den höchsten Wert. Wenn zwischen den einzelnen Gliedern der Folge eine Beziehung existiert, die mehr als zufällig ist, hat auch die Korrelation der ursprünglichen Folge mit der verschobenen

Folge in der Regel einen Wert, der signifikant von Null abweicht. In einem solchen Fall sind die Glieder der Folge autokorreliert.

Genutzt wird die Autokorrelation unter anderem in der Regressionsanalyse und der Zeitreihenanalyse [Kre06]. Die Zeitreihenanalyse zweier Zeitreihen zu unterschiedlichen Zeitverschiebungen τ zwischen den Reihen wird hierbei als Kreuzkorrelationsfunktion bezeichnet. Die Kreuzkorrelationsfunktion ermittelt damit den Zusammenhang zwischen den unterschiedlichen Merkmalen zweier Zeitreihen x_t und y_t.

3.5.3 Empirische Momente und Stationarität

Bei der Zeitreihenanalyse ist es für weitere Untersuchungen von großem Interesse festzustellen, ob die Zeitreihe als stationär aufgefasst werden kann. Die empirischen Momente bilden die Grundlage einer statistischen Analyse. Für eine Zeitreihe $(x_t)_{t\in T}$ der Länge N gelten folgende Definitionen:

Arithmetischer Mittelwert:

$$\bar{x} = \frac{1}{N}\sum_{t=1}^{N} x_t \tag{3-14}$$

Empirische Varianz:

$$s^2 = \frac{1}{N-1}\sum_{t=1}^{N}\left(x_t - \bar{x}\right) \tag{3-15}$$

Standardabweichung:

$$s = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}\left(x_t - \bar{x}\right)} \tag{3-16}$$

Die Bestimmung der Stärke des linearen Zusammenhangs der Zeitpunkte untereinander erfolgt für N Beobachtungspaare (x_i, y_i):

Empirische Kovarianz:

$$c = \frac{1}{N-1}\sum_{t=1}^{N}\left(x_t - \bar{x}\right)\left(y_t - \bar{y}\right) \tag{3-17}$$

Die Korrelation r ermittelt ein Maß für den Grad der Stärke der linearen Abhängigkeit zwischen zwei Merkmalen einer Zeitreihe. Durch die Normierung der empirischen Kovarianz, mit dem Produkt der Standardabweichung, erhält man den Korrelationskoeffizienten nach Bravis-Pearson:

Korrelation:

$$r = \frac{\frac{1}{N}\sum_{t=1}^{N}(x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(x_t - \bar{x})^2} \cdot \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(y_t - \bar{y})^2}} \qquad (3\text{-}18)$$

hierbei folgt $-1 \le r \le 1$ aus der Cauchy-Schwarz-Ungleichung $\left(\sum a_i b_i\right)^2 \le \left(\sum a_i^2\right)\left(\sum b_i^2\right)$.
Die Messung des linearen Zusammenhangs einander folgener Beobachtungen einer Zeitreihe kann mit Hilfe einer Modifizierung der Kovarianz und der Korrelation erfolgen:

Autokovarianz:

$$c_\tau = \frac{1}{N}\sum_{t=1}^{N-\tau}(x_t - \bar{x})(x_{t+\tau} - \bar{x}) \text{ mit } \tau = 0,1,2,.....,N-1 \qquad (3\text{-}19)$$

mit $c_0 = s^2$ und $c_{-\tau} = c_\tau$

Die Autokorrelationsfunktion ist eine Normierung der Autokovarianzfunktion. Mit der Autokovarianzfunktion lassen sich um τ-Zeiteinheiten gegeneinander versetzte Beobachtungen innerhalb derselben Zeitreihe auf lineare Abhängigkeiten hin untersuchen.

Autokorrelation:

$$r_\tau = \frac{c_\tau}{c_0} = \frac{\sum_{t=1}^{N-|\tau|}(x_t - \bar{x})(x_{t+\tau} - \bar{x})}{\sum_{t=1}^{N}(x_t - \bar{x})^2} \qquad -1 \le r_\tau \le 1 \qquad (3\text{-}20)$$

Die Zeitreihe ist dann schwach stationär, wenn alle Messwerte - unabhängig von der Zeit - immer um einen konstanten Mittelwert pendeln, mit nahezu konstanter Varianz. Die Variablen $\bar{x}_t$ und $\bar{x}_{t-l}$ sind die arithmetischen Mittelwerte der jeweiligen Komponente. Ist die Zeitreihe stationär, so kann $\bar{x}_t = \bar{x}_{t-l} = \bar{x}$ gesetzt werden [Ger04].
In der Theorie der Zeitreihenanalyse werden Zeitreihen auch als stochastische Prozesse untersucht. Eine der ersten Forderungen an ein lineares Zeitreihenmodell ist die Stationarität, auf die im Rahmen der stochastischen Prozesse im Kapitel 3.7 weiter eingegangen wird.

3.6 Klassische Zeitreihenanalyse

Zeitreihen aus den unterschiedlichsten Bereichen zeigen deutlich, dass diese eine Vielfalt an möglichen Variationsmustern besitzen können. Die Temperatur-Zeitreihe beispielsweise ist vor allem durch die periodische Komponente des Jahresganges gekennzeichnet. Ähnliche komplexere, jedoch auch periodische Anteile zeigen Zeit-

reihen wie Verkehrsdurchsätze - Tagesrhythmus mit Spitzen im morgendlichen und abendlichen Berufsverkehr - sowie Arbeitsmarktzahlen, die Sommer- und Wintervariabilität mit einem starken und einem schwachen Maximum darstellen. Im Gegensatz dazu sind solche Komponenten in Zeitreihen von Börsendaten überhaupt nicht zu finden, da diese eine Variation auf sehr großen Zeitskalen zeigen, die einer stochastischen Dynamik auf kleineren Zeitskalen überlagert sind. Die Beschreibung dieser Verläufe hat in den vergangen Jahren eine Vielzahl von einschlägigen Methoden hervorgebracht, von denen die klassische Zeitreihenanalyse zu den elementaren Methoden zählt. Diese beschreibt die Zerlegung der Zeitreihe durch miteinander verknüpfte Komponenten. Diese Komponenten werden mit deterministischen Verfahren wie beispielsweise der „Methode der kleinsten Quadrate“ oder der Fouriertransformation bestimmt oder parametrisiert.

3.6.1 Komponentenmodell

Die zuvor beschriebenen Zeitreihen zeigen zwei Arten von deterministischer oder quasi-deterministischer Langzeit-Variabilität:

- Trends, die eine langfristige Grundrichtung und damit die systematische Veränderung der zeitlichen Entwicklung einer hinreichend langen Zeitreihe erfassen.
- Zyklen, die einen mittelfristig wiederholenden Einfluss auf die Dynamik der Zeitreihe erfassen.

Unter einem Zyklus im engeren Sinne kann ein Variationsmuster verstanden werden, welches einer harmonischen Funktion, wie beispielsweise Sinus oder Cosinus, folgt. Im erweiterten Sinne kann ein Zyklus aber auch ein beliebiges, periodisches Muster bedeuten, was dadurch gekennzeichnet ist, dass die entsprechende Komponente sich nicht mehr durch die Variation mit einer einzigen Frequenz beschreiben lässt. Solche periodischen Zyklen können aufgrund ihrer festen Periodenlänge als saisonale Komponente bezeichnet werden. Die Saisonalität charakterisiert damit regelmäßige Schwankungen, die innerhalb eines Jahres, eines Quartals oder eines Monats oder einer festgelegter Dauer liegen können [Hüt86]. Neben zyklischen Komponenten, die regelmäßige, beispielsweise jahreszeitliche Einflüsse erfassen, existieren auch solche, die immer wieder auftreten, jedoch keine fest definierte Periode besitzen. Diese unterliegen beispielsweise konjunkturellen Einflüssen.

Während man Komponenten mit fester Periode mit geeigneten statistischen Verfahren explizit detektieren kann, lassen sich vor dem Hintergrund beschränkter Beobachtungszeit Trends und aperiodische oder konjunkturelle Zyklen häufig nur schwer nachweisen. Aus diesem Grund fasst man beide in der Regel zu einer einzigen ”glatten” Trendkomponente zusammen. Liegen Trendkomponenten wie auch saisonale oder konjunkturelle Zyklen vor, so ist die Zeitreihe nicht mehr stationär, d. h. ihre statistischen Eigenschaften hängen vom beobachteten Zeitintervall ab. Insbesondere

sind die Mittelwerte der Beobachtungsgrößen zeitabhängig, was etwa am Beispiel der Tagestemperaturen in den Sommer- und Wintermonaten einfach zu verdeutlichen ist.

Trends und Zyklen beeinflussen also immer die statistischen Eigenschaften einer Zeitreihe und sollten daher vor einer weitergehenden Analyse in geeigneter Form separiert bzw. eliminiert werden. Insbesondere gilt dies für Trends. Der verbleibende Anteil der Zeitreihe unterliegt Extrasaisonalitäten oder Kalendersaisonalitäten, die durch die Kalenderstruktur bedingt sind [Lew74]. Hierzu zählt die unterschiedliche Anzahl von Arbeitstagen in einem Monat, einem Jahr oder gleiche Monate verschiedener Jahre. Die dann verbleibende Restkomponente enthält alle Schwankungen, die nicht bereits durch die obigen Komponenten erfasst sind. Das sogenannte „Klassische Komponentenmodell“ einer Zeitreihe kann somit im Allgemeinen die folgenden Bestandteile enthalten:

1. Eine Trendkomponente $\mathrm{m(t)}$, die systematische Veränderungen des mittleren Niveaus der Zeitreihe repräsentiert.
2. Aperiodische Konjunkturzyklen $\mathrm{k(t)}$, die mehrjährige nicht regelmäßige Schwankungen berücksichtigen.
3. Eine saisonale Komponente $\mathrm{s(t)}$, mit einer klar definierten Periodenlänge.
4. Eine Restkomponente $\mathrm{r(t)}$, deren Mittelwert per Definition Null sein soll.

Trend- und Konjunkturkomponente werden als glatte Komponente ($\mathrm{g(t) = m(t) + k(t)}$) und saisonale und konjunkturelle Komponente werden zu einer zykischen Komponente $\mathrm{z(t)}$ zusammengefasst ($\mathrm{z(t) = k(t) + s(t)}$). Die Varianz der Restkomponente $\mathrm{r(t)}$, die häufig Eigenschaften eines stochastischen Prozesses zeigt und daher als ”Rauschen” aufgefasst werden kann, hängt stark vom statistischen Verfahren ab, das zur Separation der deterministischen Komponenten benutzt wird [See08]. Die einzelnen Komponenten treten im Regelfall in tatsächlich beobachteten Zeitreihen vermischt auf. Die Verknüpfung der Komponenten kann dabei in Abhängigkeit der Zeitreihe auf unterschiedlicher Weise erfolgen. Im Normalfall werden die Komponenten mit Hilfe des additiven Modells, in dem sich Einflüsse überlagern, oder des multiplikativen Modells miteinander verknüpft.

$$\mathrm{x(t) = m(t) + k(t) + s(t) + r(t) = g(t) + z(t) + r(t)} \tag{3-21}$$

$$\mathrm{x(t) = m(t) \cdot k(t) \cdot s(t) \cdot r(t) = g(t) \cdot z(t) \cdot r(t)} \tag{3-22}$$

mit
$\mathrm{x(t)} : /N_0 \rightarrow /R \mathrel{\hat{=}}$ Zeitreihe
$\mathrm{m(t)} : /N_0 \rightarrow /R \mathrel{\hat{=}}$ Trendkomponente
$\mathrm{g(t)} : /N_0 \rightarrow /R \mathrel{\hat{=}}$ glatte Komponente
$\mathrm{k(t)} : /N_0 \rightarrow /R \mathrel{\hat{=}}$ Konjunkturkomponente
$\mathrm{s(t)} : /N_0 \rightarrow /R \mathrel{\hat{=}}$ Saisonalkomponente
$\mathrm{r(t)} : /N_0 \rightarrow /R \mathrel{\hat{=}}$ Restkomponente

Durch Logarithmieren der Gleichung (3-21) lässt sich das multiplikative Modell in ein additives transformieren.

$$x(t) = g(t) \cdot z(t) \cdot r(t) \Rightarrow \log(x(t)) = \log g(t) + \log z(t) + \log r(t) \qquad (3\text{-}23)$$

Die einzelnen Komponenten treten im Regelfall in tatsächlich beobachteten Zeitreihen vermischt auf. Die Verknüpfung der Komponenten kann dabei in Abhängigkeit der Zeitreihe auf völlig unterschiedliche Weise erfolgen. Die Zeitreihenanalyse ermittelt primär das Verhalten einzelner oder mehrerer Komponenten. Da diese unterschiedliche statistische Eigenschaften besitzen, ist es erforderlich, diese zur weiteren Analyse in eine geeignete Form zu überführen.

3.6.2 Trendbestimmung

In der Statistik existieren verschiedenste Verfahren, um die Trendkomponente einer Zeitreihe zu bestimmen. Bei diesen Verfahren wird die zeitliche Entwicklung der Zeitreihe festgestellt, das heißt, steigende, fallende oder horizontale Entwicklungen werden erkannt und diese mathematisch beschrieben. Ziel dieser Verfahren ist die Isolierung von Trends oder die Bestimmung einer trendbereinigten Zeitreihe, die Basis zur Durchführung weiterführender Analysen ist [See08].

Zur Bestimmung des Trends geht man zunächst davon aus, dass sich die betrachtete Zeitreihe in Form $x(t) = m(t) + r(t))$ mit $t = 1...N$ darstellen lässt, $m(t)$ repräsentiert dabei eine Geradenfunktion oder kann durch Polynome beschrieben werden. Werden Polynome höheren Grades für die Trendfunktion angewendet, so lassen sich unterschiedliche Trendverläufe darstellen. In diesen Darstellungen können Wendepunkte sowie lokale Minima und Maxima beschrieben werden. Die Beschreibung solcher polynomialer Trends erfolgt häufig durch Anwendung von Regressionstechniken, wobei sich die Trendbestimmungsmodelle in lineare und nichtlineare Modelle klassifizieren lassen.

Zur Trendbestimmung eignen sich im Allgemeinen glatte mathematische Funktionen, die nicht periodisch sind oder eine große Periodenlänge besitzen. Die Ermittlung der Trendlinie kann durch die Freihandmethode, die Methode der halben Durchschnitte, die Methode der gleitenden Durchschnitte oder die Methode der kleinsten Quadrate beschrieben werden. Allgemein können Trendfunktionen durch Polynome n-ten Grades abgeschätzt werden. Diese können ebenfalls mit der Methode der kleinsten Quadrate geschätzt werden. Generell sind die Schätzungen für $m(t)$ besser, je höher der Polynomgrad p gewählt wird. In der Praxis werden jedoch meist nur Polynome bis zum zweiten Grad verwendet. Die Verwendung von Polynomen zur Abbildung der Trendfunktion birgt auch Nachteile, denn eine globale Gültigkeit des verwendeten Modells wird vorausgesetzt, lokale Eigenheiten werden nur schwach gewichtet, alle Beobachtungen gehen gleich stark ein und polynomiale Trends gehen schnell gegen

$\pm\infty$, was zur Folge hat, dass der prognostizierbare Zeitraum begrenzt ist. Der Vollständigkeit halber sei erwähnt, dass neben den linearen und polynomialen auch logarithmische und exponentielle Funktionen zum Einsatz kommen. Sie besitzen aber einen geringeren Verbreitungsgrad [Sch97] und finden bei der Analyse elektrischer Lastgangzeitreihen keine Anwendung.

3.6.3 Methode der kleinsten Quadrate

Die Trendfunktion repräsentiert die langfristige Entwicklung einer Zeitreihe. Oft ist die Elimination des Trends notwendig, um lineare Abhängigkeiten zwischen den Zeitpunkten einer Zeitreihe zu untersuchen. Zur Bestimmung des Trends geht man zunächst davon aus, dass sich die betrachtete Zeitreihe in der Form $x_t = m_t + e_t \quad t = 1...N$ darstellen lässt, m_t repräsentiert eine Funktion, wie z. B. eine Gerade oder Parabel. Eine geeignete Methode zur Parametrisierung der Funktion ist die „Methode der kleinsten Quadrate“ nach Gauss unter anderem zur Ermittlung einfacher linearer oder polynominaler Funktionen zur Beschreibung des Trendanteils.

Die Methode der kleinsten Quadrate ist das mathematische Verfahren zur Ausgleichungsrechnung und wird in der Statistik als Schätzmethode in der Regressionsanalyse eingesetzt. Die Anwendung der Methode zur Ermittlung der Trendfunktion ist in der Praxis weit verbreitet. Sie bietet den Vorteil, dass eine Schätzfunktion explizit erzeugt werden kann. Oft besteht ein Problem, dessen mathematische Beschreibung es zu ermitteln gilt aus abhängigen Variablen x, deren Zustandekommen von einer vorgegeben Variablen t oder von mehreren Messpunkten t_1 bis t_n abhängig sein kann. Dabei sollten die $x(n)$-Werte der zu ermittelnden Modellfunktion

$$x(n) = f(t_1,...,t_a;a_1,...,a_b), \tag{3-24}$$

die von den Variablen a_i und b_i abhängig sind, die Messwerte approximieren. Die Wahl des Typs der Modellfunktion erfolgt häufig durch eine Parabel oder Exponentialfunktion, dabei dient der Parameter a_i zur Anpassung der Funktion des gewählten Typs an die beobachteten Werte $x(n)$.

Voraussetzung zur Lösbarkeit des Gleichungssystems ist, dass es mindestens so viele Messwerte wie Parameter der Funktion gibt. Es werden die Parameter $a_1,....,a_b$ ausgewählt werden, bei denen die Summe der Quadrate der Abweichung zwischen entsprechender Modellkurve und Daten minimal wird. Der Sachverhalt kann wie folgt beschrieben werden:

$$\min_{\vec{a}} \sum_{i=1}^{n} \left(x_{m,i} - a_i\right)^2 \tag{3-25}$$

Dieses Minimierungsproblem hängt von der Modellfunktion ab. Ist die Modellfunktion linear in den unabhängigen Modellvariablen $t_1,.....,t_n$, erhält man ein Gleichungssystem der Form: $x(n) = f(t_1,...,t_b;a_0,...,a_b) = a_0 + a_1t_1 + \cdots + a_bt_b$, das auf das überbestimmte lineare Gleichungssystem

$$\begin{aligned} a_0 + a_1t_{11} + \cdots + a_jt_{1j} \cdots + a_bt_{1b} &= x_1 \\ a_0 + a_1t_{21} + \cdots + a_jt_{2j} \cdots + a_bt_{2b} &= x_2 \\ &\vdots \\ a_0 + a_1t_{i1} + \cdots + a_jt_{ij} \cdots + a_bt_{ib} &= x_i \\ &\vdots \\ a_0 + a_1t_{n1} + \cdots + a_jt_{nj} \cdots + a_bt_{nb} &= x_n \end{aligned} \tag{3-26}$$

führt. Indem man die Koeffizienten t_{ii} zur Datenmatrix A, die Parameter a_i zum Parametervektor $\vec{a}$ und die Beobachtungen x_i zum Vektor $\vec{b}$ zusammenfasst, kann man das lineare Gleichungssystem in Matrixform darstellen.

$$\begin{pmatrix} 1 & t_{11} & \cdots & t_{1j} & \cdots & t_{1b} \\ 1 & t_{21} & \cdots & t_{2j} & \cdots & t_{2b} \\ \vdots & & & & & \\ 1 & t_{i1} & \cdots & t_{ij} & \cdots & t_{ib} \\ \vdots & & & & & \\ 1 & t_{n1} & \cdots & t_{nj} & \cdots & t_{nb} \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_2 \\ \vdots \\ a_b \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} \tag{3-27}$$

Dieses Gleichungssystem lässt sich somit auch in der Form $\underline{A} \cdot \vec{a} = \vec{b}$ darstellen. Dabei führt dann wieder der Ansatz der kleinsten Quadrate zu einem linearen Ausgleichsproblem der Form $\min_x \|Ax - b\|_2$, welches zu lösen ist. Sind die Spaltenvektoren der Matrix A linear unabhängig, so ist die Matrix $\underline{A}^T\underline{A}$ positiv definit. Es gibt genau eine Lösung durch das Lösen der Normalengleichung: $\underline{A}^T \cdot \underline{A}\vec{x} = \underline{A}^T \cdot \vec{b}$ bzw. $\vec{x} = \left(\underline{A}^T \cdot \underline{A}\right)^{-1} \underline{A}^T \cdot \vec{b}$ [Eld90].

3.6.4 Filter

Die Aufgabe eines Filters im Sinne der Zeitreihenanalyse ist, neben der Bestimmung globaler Trends, die Glättung der Zeitreihe und die Elimination von irregulären Schwankungen innerhalb der Zeitreihe. Filter werden dann auch zur Bestimmung des Grades von Trendpolynomen eingesetzt. Ein Filter ist eine Funktion, die die eingehende

Zeitreihe x_t in eine ausgehende Zeitreihe y_t transformiert. $y_t = F(x_t)$ für $t \in T$. Ein Filter wird als linear bezeichnet, wenn gilt:

$$F(a \cdot x_t) = a \cdot y_t \text{ und}$$
$$F(x_{1t} + x_{2t}) = F(x_{1t}) + F(x_{2t}) = y_{1t} + y_{2t} \quad (3\text{-}28)$$

Ein weiterer Filtertyp ist der Differenzenfilter der dazu dient, polynomiale Trends durch Differenzenbildung benachbarter Werte einer Zeitreihe zu beseitigen. Für den Differenzenfilter 1. Ordnung einer Zeitreihe gilt:

$$\Delta x_t = x_t - x_{t-1} \text{ für } t = 2,3,...,N \quad (3\text{-}29)$$

Für den Differenzenfilter $p-ter$ Ordnung gilt:

$$\Delta^p x_t = \Delta^{p-1} x_t - \Delta^{p-1} x_{t-1} \text{ für } t = p+1,...,N \quad (3\text{-}30)$$

Zur Elimination von Trends einer Zeitreihe bildet man so lange Differenzen, bis eine Zeitreihe keine Trendkomponente mehr enthält. Ein Differenzenfilter reduziert den Grad eines Polynoms jeweils um Eins. Wendet man das Verfahren entsprechend oft an, erhält man für ein beliebiges Polynom einen konstanten Wert [Sch97]. Die Anzahl der Differenzenbildung entspricht dann dem Grad des Ausgangspolynoms.

3.6.5 Zyklische Schwankungen

Zeitreihen elektrischer Lastgangdaten besitzen oft ein ausgeprägtes periodisches Verhalten, dessen Eigenschaften durch zyklische Anteile des Komponentenmodells repräsentiert werden. Eine Funktion ist periodisch mit Periode $T \neq 0$ wenn gilt: $f(t \pm T) = f(t)$ $t \in /R$, weiter gilt $f(t \pm k \cdot T) = f(t)$ mit $t \in /R$, $k \in Z$. Unter recht allgemeinen Voraussetzungen lassen sich gerade periodische Funktionen als Überlagerung von harmonischen Schwingungen verschiedenster Amplitude und Frequenz abbilden:

$$x(t) = A \cdot \cos(\omega_0 \cdot t - \varphi)$$
$$x(t) = A_1 \cdot \cos(2\pi \cdot f \cdot t) + A_2 \cdot \sin(2\pi \cdot f \cdot t) \quad (3\text{-}31)$$
$$A = \sqrt{A_1^{\,2} + A_2^{\,2}}; \varphi = \arctan\left(\frac{A_2}{A_1}\right)$$

mit A: Amplitude; ω_0: Kreisfrequenz; φ: Phasenverschiebung. Das Periodogramm dient dem Auffinden von manifesten oder verborgenen Periodizitäten in einer Zeitreihe und ist eine Funktion $I(f)$ der Frequenz, die angibt, mit welcher Intensität die jeweilige Frequenz innerhalb der Zeitreihe vorhanden ist. Es wird angenommen, dass die Zeitreihe $(x_t)_{t \in T}$ mit $T = (1,...,N)$ sich aus der Überlagerung von harmonischen Schwingungen beliebiger Frequenzen zusammensetzt. Zur Bestimmung der Intensität

einer Frequenz bietet sich die Regression mit Hilfe der „Methode der kleinsten Quadrate“ an. Für das zwei- bzw. einseitige Spektrum gilt dann:

$$I(f) = \frac{1}{N}\left[\left(\sum_{t=0}^{N} x_t \cos(2\pi f t)\right)^2 + \left(\sum_{t=0}^{N} x_t \sin(2\pi f t)\right)^2\right] \qquad (3\text{-}32)$$

Das Periodogramm hat folgende Eigenschaften [Sch87]:

- $I(f)$ verschwindet für $f = 0 : I(0) = 0$
- $I(f)$ ist eine gerade Funktion: $I(-f) = I(f)$
- $I(f)$ ist eine nicht negative Funktion: $I(f) \geq 0$
- $I(f)$ ist periodisch mit Periode 1: $I(f+1) = I(f)$

Bei der weiteren Analyse eines Periodogramms müssen zusätzlich weitere Eigenschaften beachtet werden. Mit dem Begriff Leakage wird das Phänomen bezeichnet, dass in der Umgebung einer realen Schwingungskomponente eine Erhöhung der Ordinatenwerte beobachtet werden kann. Dieser Effekt ist wesentlich auf die Endlichkeit der Zeitreihe zurückzuführen. Weiter kommen Diskontinuitäten hinzu, wie sie z. B. bei der Modifikation der Zeitreihe für die Fast-Fourier-Transformation entstehen können. Unter Aliasing wird das Phänomen verstanden, dass bei der diskreten Abtastung eines Signals Schwingungen höherer Frequenz nicht von Schwingungen niedrigerer Frequenz unterschieden werden können. Die größte noch eindeutig bestimmbare Frequenz wird als Nyquist–Frequenz bezeichnet. Laut dem Shannon–Theorem muss die erforderliche Abtastrate eines Signals mindestens das Doppelte der höchsten Signalfrequenz sein. Oberschwingungen einer Fundamentalfrequenz f sind Frequenzen mit $k \cdot f$ mit $k = 1,2,3...$. Sie resultieren aus dem nicht sinusförmigen Charakter zyklischer Schwankungen innerhalb der Zeitreihe. Sie sind keinesfalls als eigenständige Periodizitäten zu betrachten, sondern repräsentieren vielmehr die Aspekte der Saisonalkomponenten.

3.7 Stochastische Prozesse

Die Stochastik ist die mathematische Beschreibung von zufälligen Vorgängen, die zwar einer Gesetzmäßigkeit unterliegen können, deren Ausgänge aber trotzdem als zufällig betrachtet und beobachtet werden. Eine Zeitreihe wird dabei als ein dynamischer Prozess verstanden, der neben den deterministischen Komponenten auch einen zufälligen Charakter besitzt. Der stochastische Prozess beschreibt dabei zeitlich geordnete, zufällige Vorgänge der Zeitreihe. Während sich stochastische Prozesse als Teilgebiet der Statistik verstehen, versuchen die Methoden der klassischen Zeitreihenanalyse eine Zeitreihe mit einer speziellen Struktur der Zufallsfunktionen zu beschreiben, wobei die Untersuchung von Abhängigkeiten der Variablen hierbei von besonderem Interesse ist.

Ein stochastischer Prozess ist eine Folge von Zufallsvariablen X_t mit $t \in T$ und $T = /N$ oder $T = /R$, die Zufallsvorgänge repräsentieren. Bei jeder Realisierung eines Ereignisses ω des Zufallsvorgangs nimmt jede Zufallsvariable X_t des Prozesses einen reellen Wert $X_t(\omega)$ an [Sch87]. Für jede Realisation entsteht eine Folge $(X_t(\omega))_{t \in T}$ von Werten. Der stochastische Prozess ordnet also jedem Ereignis ω eine Zeitreihe $(X_t(\omega))$ zu und kann auf zwei Arten interpretiert werden: als „Ensemble“ von Zeitreihen $(X_t(\omega))_{t \in T}$ oder als eine Folge von Zufallsvariablen, die jedem Zeitpunkt t eine Zufallsvariable zuordnet.

Für physikalische Vorgänge, wie z. B. das Lastverhalten, lassen sich in der Regel die Rahmenbedingungen, die zum Zeitpunkt der Realisierung vorherrschten, nicht reproduzieren. In der Regel wird daher von einer Folge von Zufallsvariablen $X_{t_1}, X_{t_2,}, ..., X_{t_N}$ ausgegangen. Ein stochastischer Prozess wird definiert, indem für jede Auswahl beliebiger Zeitpunkte $t_1, t_2, ..., t_n \in T$ der Folge das System der Verteilungsfunktionen $F_{t_{1,...,}, t_n}(x_{1_1}, x_2, ..., x_n)$ vorgegeben wird. Um dies zu ermöglichen, müssen die Verteilungsfunktionen eine gewisse „Konsistenz“ (Stationarität) besitzen. Eine elementare Bedingung bei der Anwendung stochastischer Prozesse ist, dass die Zeitreihe die Stationaritätsbedingungen erfüllt. Bei einem Prozess im Rahmen von Lastprognosen der Länge N wäre dies ein enormer Aufwand, da jedes statistische Moment einer Zufallsvariablen $(X_t)_{t \in T}; T = \{1,2,3,...N\}$ bestimmt werden müsste. Im Rahmen der Untersuchungen wurde die Stationaritätsbedingungen der Lastgangzeitreihen mit Hilfe des Mittelwertstationaritäts- sowie des Varianzstationaritätstests (Anlage 9.8) nachgewiesen.

Lineare stochastische Prozesse liefern eine Näherung der untersuchten Zeitreihe durch lineare Verknüpfung der Vergangenheitswerte, wobei die Gewichte den jeweiligen linearen Zusammenhang zwischen zwei Zeitpunkten x_t und x_{t-i}, mit repräsentieren. Diesen Zeitreihen liegen lineare Prozesse endlicher Ordnung zugrunde. Drei Typen, die hierzu Anwendung finden, werden nachfolgend kurz beschrieben.

Der Moving-Average-Prozess ist der gleitende Durchschnitt und ergibt sich aus einem Rauschprozess $e_t \approx N(0, \sigma^2)$, wobei die Zeitreihe als gewichtetes Mittel dieser Rauschkomponenten (Zufallschocks) und deren Vorgängern gebildet wird. Ein Moving-Average-Prozess q der Ordnung $MA(q)$ ist definiert durch:

$$x_t = e_t - \sum_{i=1}^{q} \beta_i \cdot e_{t-i} \text{ mit } \beta_i \in IR \qquad (3\text{-}33)$$

wobei der Parameter β_i den linearen Zusammenhang zwischen einem zurückliegenden Zufallsschock und dem aktuellen Wert x_t repräsentiert.

Ein autoregressiver Prozess beschreibt den linearen Zusammenhang zwischen der aktuellen Beobachtung x_t und den vorausgegangenen Beobachtungen unter Einbeziehung eines Fehlerterms (Rauschkomponente) $e_t \approx N(0,\sigma^2)$. Ein autoregressiver Prozess $AR(p)$ ist definiert durch:

$$x_t = \sum_{i=1}^{p} \alpha_i \cdot x_{t-i} + e_t \text{, mit } \alpha_i \in IR \tag{3-34}$$

wobei der Parameter α_i den linearen Zusammenhang zwischen zwei Zeitpunkten repräsentiert. Der Autoregressive-Moving-Average-Prozess ($ARMA$-Prozess mit den Ordnungen p und q) ist die Kombination von $AR(p)$- und $MA(q)$-Prozessen und wird definiert durch:

$$x_t = \sum_{i=1}^{p} \alpha_i \cdot x_{t-i} + e_t - \sum_{j=1}^{q} \beta_i \cdot e_{t-j} \,. \tag{3-35}$$

Das $ARIMA(p,d,q)$-Modell (Autoregressive-Integrierte-Moving-Average-Prozesse) stellt eine Erweiterung des $ARMA$-Modelles auch für instationäre Prozesse dar. Durch Umformung der AR-Gleichung lässt sich zeigen, dass durch Differenzenbildung White Noise e_t erzeugt werden kann, was einen statistischen Prozess darstellt. Damit kann sichergestellt werden, dass aus einem instationären Prozess durch geeignete Bildung von Differenzen ein stationärer $ARIMA(p,d,q)$-Prozess erzeugt wird [Hus05].

4 Neuronale Netze

Ein neuronales Netz besteht aus einer Vielzahl einzelner Zellen, die den Nervenzellen eines biologischen neuronalen Netzes entsprechen. Jede einzelne Zelle ist bereits eine einfache Verarbeitungseinheit. Durch das Zusammenwirken dieser Einheiten entsteht ein leistungsfähiges Modul, welches in der Lage ist, anspruchsvolle Aufgaben wie Mustererkennung, Prognosefindung usw. zu erfüllen. Ebenso wie bei dem biologischen Vorbild sind zwischen diesen Zellen Verbindungen vorhanden, über die Reizimpulse geleitet werden. Die Art und Weise, wie diese Reize weitergeleitet werden und wie sie in der Empfängerzelle berücksichtigt werden, bestimmt die logische Funktionalität des Netzes in seiner Gesamtheit. Neuronale Netze oder „künstlich“ neuronale Netze sind damit informationsverarbeitende Systeme. Die Forschung im Themengebiet der künstlichen neuronalen Netzen verfolgt hauptsächlich zwei Ziele: die Erforschung und das Verständnis der Funktionsweise natürlich neuronaler Netze und deren Eignung für industrielle Anwendungen.

Der wesentliche Unterschied zwischen den herkömmlichen algorithmischen Ansätzen und neuronalen Netzen besteht darin, dass beim neuronalen Netz zunächst nur die Struktur und das Lernverfahren vorgegeben und nicht für die zu lösende Aufgabe ein Algorithmus angepasst wird, sondern das Netz sich selbst im laufenden Lernprozess konfiguriert. Neuronale Netze entwickeln somit anhand von Beispielen und Lernverfahren Lösungsstrategien. Die Methoden, neuronale Netze in Prognosesysteme zu integrieren, sind bekannt [Sto00] [Hof04] [Thi98] [Dau05]. Es existiert für die klassischen Anwendungen eine Vielzahl effizienter Lernalgorithmen zur Bestimmung der inneren Parameter (Gewichte). Theoretische Untersuchungen belegen, dass schon ein einfach strukturiertes künstliches neuronales Netz (Multilayer Perzeptron) verschiedene nichtlineare, stetige Zusammenhänge beliebig genau approximieren kann, wenn es nur hinreichend groß gewählt wird [Dau05]. Darüber hinaus wurde der Nachweis erbracht, dass insbesondere bei Verwendung sigmoidialer Funktionen als neuronale Transferfunktionen die künstlichen neuronalen Netze der klassischen multiplen Regression überlegen sind. All dies rechtfertigt den Einsatz der Technik neuronaler Netze in der Zeitreihenanalyse und besonders in der Lastprognoseproblematik.

Im folgenden Abschnitt wird zunächst ein Überblick über das komplexe Themengebiet der künstlichen neuronalen Netze gegeben. Es werden deren Aufbau und Funktion, verschiedene Netzstrukturen und diverse Lernalgorithmen beschrieben und deren Funktionsweise erläutert. Die Ausführungen beschränken sich jedoch dabei auf Themengebiete, die für die weiteren Untersuchungen von Interesse sind.

4.1 Grundlagen und historische Entwicklung

Eine allgemein gültige Charakterisierung und Systematisierung neuronaler Netze existiert nicht, da sie aus sehr unterschiedlichen Betrachtungsweisen interpretiert werden können. Die klassische Sichtweise versteht neuronale Netze als einen Hauptzweig der künstlichen Intelligenz [Pol05]. Eine Systematisierung von Netzwerktypen ist nach dem Funktionsumfang des Netzes (Funktionsapproximation, Klassifikation, Assoziativspeicher) möglich, wobei alternativ die Möglichkeit besteht, eine Systematisierung anhand der Netztopologien vorzunehmen [Kuh04] [Reh95].

Erfolgt eine Unterteilung nach Funktionsumfang, so können unter der Kategorie Funktionsapproximation solche Netze eingruppiert werden, die besonders gut geeignet sind, ein Modell der Funktion $y(t) = f(x_t) + \varepsilon(t)$ zu schätzen. Mit $y(t)$ wird die zu erklärende abhängige Variable des t-ten Beobachtungsobjektes zum t-ten Zeitpunkt bezeichnet, x_t bezeichnet den Vektor unabhängiger Variablen und $\varepsilon(t)$ den zufälligen Störterm. Voraussetzung für die Funktionsapproximation ist, dass Daten in zeitlicher Abfolge vorliegen und diese einem funktionalen Zusammenhang unterliegen. In diesem Fall können neuronale Netze dazu eingesetzt werden, den impliziten Zusammenhang und künftige Zeitreihendaten vorherzusagen. Hierzu ist es erforderlich, dass das neuronale Netz einen Abschnitt aus der Zeitreihe mit vorgegebenem Output lernt zu assoziieren. In einer dann folgenden Verarbeitungsphase ergänzt das Netz die Zeitreihe um Vorhersagewerte.

Bei der Aufgabenstellung der Klassifikation geht es im Allgemeinen darum, einem Eingangsmuster ein davon abweisendes Ausgangsmuster zuzuordnen. Anhand der Ausprägungen von m Merkmalen des Eingangsvektors werden diese k Klassen zugeordnet. Dabei sind die verschiedenen Eingabemuster statisch und unabhängig voneinander und unterliegen keiner zeitlichen Reihenfolge. Die Merkmale können dabei einen n-dimensionalen Merkmalsraum aufspannen. Ein Objekt wird dann durch die Ausprägung eines Merkmals $1,\ldots n$, das heißt einem Merkmalsvektor im Merkmalsraum, repräsentiert. Die Objektklassen ergeben sich durch eine Zerlegung des Merkmalraumes in zusammengehörende Teilmengen sowie einer Restklasse, der alle diejenigen Objekte zugeordnet werden, die keiner der anderen Klassen zuzuordnen sind. Bei der Klassifikation wird nun der Merkmalsvektor eines Objekts genau einer der vorher definierten Klassen oder aber der Restklasse zugeordnet. Hierbei kann eine Klasse durchaus durch mehrere Punktwolken im Merkmalsraum repräsentiert werden. Die zu klassifizierenden Objekte werden durch ihre Eigenschaften bzw. Merkmale beschrieben. Es kann durchaus vorkommen, dass die Klassen so angeordnet sind, dass eine Trennung durch einfache Abstandsmaße nicht ohne weiteres möglich ist. Die Definition der Klassen muss natürlich sinnvoll erfolgen. Ein Klassifikator im allgemeinen Sinn ist also eine Verfahrensweise, die eine Zuordnung eines Merkmalsvektors, dessen Klassenzugehörigkeit unbekannt ist, in eine der vorher definierten Klassen bewerk-

stelligt. Die Dimensionierung des Klassifikators bezeichnet man bei neuronalen Klassifikatoren, auf die wir uns im Weiteren beschränken werden, auch häufig als Lernen. Traditionelle Klassifikatoren lassen sich in die Kategorien geometrische und statistische Klassifikatoren unterteilen [Mat08].

Als ein weiterer Funktionsumfang neuronaler Netze kann die Eigenschaft des assoziativen Speichers aufgeführt werden. Ein Assoziativspeicher oder auch inhaltsadressierbarer Speicher (engl. Content Addressable Memory, CAM) ist eine Speicherform, bei der mit der Assoziation von Inhalten gearbeitet wird, um auf einzelne Speicherinhalte zuzugreifen. Dieses ist kein besonderes Merkmal neuronaler Netze, denn auch andere Speichermodelle können dieses leisten. Die Besonderheit von Speichermodellen auf Basis neuronaler Netz besteht jedoch in der Fehlertoleranz. Innerhalb der Lernphase werden dem Speicher verschiedene Muster vorgelegt. Hierbei werden dem assoziativen Speicher die Frage-Antwort-Tupel in der Lernphase zur Verfügung gestellt, und durch Wahl der Lernregel die Korrelation dieser Muster zueinander im Speicher codiert. Ist dann bei einer späteren Abfrage des Speichers die Anfrage verfälscht oder unvollständig, kann dieser zumindest die hierzu richtige Antwort reproduzieren. Vornehmlich wird dieser Funktionsumfang im Rahmen von Bilderkennungssystemen angewendet [Ros96].

Eine alternative Systematisierung neuronaler Netze kann nach der Art ihrer Netzarchitektur erfolgen. Grundsätzlich können neuronale Netze in zwei große Bereiche, die Netzwerke mit Rückkopplung (Feedback-Netze) und solche ohne Rückkopplung unterteilt werden. Bei rückgekoppelten Netzen wird noch zwischen Netzen mit direkter und indirekter Rückkopplung unterschieden.

Bei direkter Rückkopplung wird der Ausgang eines Neurons direkt auf seine Eingangsneuronen zurückgeführt. Dieses beeinflusst beim nächsten Bearbeitungsschritt direkt die Berechnung des neuen Aktivierungszustandes.

Bei indirekter Rückkopplung existieren Verbindungen von einer Schicht zu einer Vorgängerschicht, so dass bestimmte Eingangsbereiche besonders hervorgehoben werden, d. h. diese Netztypen können trainiert oder konstruiert werden [Law92]. Ein Nachteil dieser Netztypen ist, dass durch die Existenz eines Zufallselements nicht immer dieselbe exakte Lösung für identische Eingänge gefunden wird.

Die zweite große Kategorie sind die vorwärtsgekoppelten Netzwerke. Netze aus dieser Gruppe werden in heutigen Anwendungen als häufigste Netztopologien eingesetzt, nicht zuletzt auch aufgrund der sehr hohen Speicherbelegung der Gruppe rückgekoppelter Netze.

Einer der Schwerpunkte dieser Arbeit liegt in der Entwicklung und Anwendung der Kohonenkarte. Dabei handelt es sich um ein zweischichtiges neuronales Netz mit einer

Eingabe- und einer Kartenschicht, die meist als zweidimensionale Matrix dargestellt wird. Der Signalfluss ist vorwärts gerichtet, weil keine Verbindungen zur Eingabeschicht zurückführen. Die Neuronen der Kartenschicht sind dagegen sehr wohl miteinander verbunden, und sie beeinflussen sich gegenseitig. Das Lernen wird nicht überwacht, es wird demnach nicht auf ein bestimmtes Ergebnis hingelernt. Im Rahmen der Untersuchung werden vorwärtsgerichtete Netztopologien, wie beispielsweise Feedforward-, Elman- und Jordan-Netze eingesetzt, die mit Hilfe unterschiedlicher Lernalgorithmen (Feedforward- und Backpropagation) zur Erstellung von Lastprognosen analysiert werden.

Die historische Entwicklung neuronaler Netze lässt sich in vier Perioden gliedern [Zel94]. Als erste leisteten McCulloch und Pitts in den 40er Jahren Pionierarbeit auf diesem Gebiet. In ihrem Aufsatz „A Logical Calculation of the ideas imminent in nervons activity" [McP43] zeigten sie, dass bereits eine kleine Klasse von Neuronen jede algorithmische Funktion berechnen kann. 1949 wurde von Donald O. Hebb in der Publikation „The Organisation of Behaviour" [Heb49] die nach ihm benannte Hebb´sche Lernregel veröffentlicht, die bis heute in ihrer allgemeinen Form als einfaches universelles Lernkonzept für alle neuronalen Netze gilt. Auch das Konzept der Zellgruppen geht auf Hebb zurück.

In der zweiten Periode wurde von Frank Rosenblatt und Charles Wightman [Ros58] in den Jahren 1957 - 1959 am MIT der erste Neurocomputer entwickelt. Neben dieser technischen Leistung wurde Frank Rosenbatt durch sein Buch „Principles of Neurodynamics" bekannt, in dem er das Perzeptron-Konvergenz-Theorem beschreibt. Die dritte große Periode fand in den 60er Jahren mit der Theorie von Minsky sowie den Arbeiten von Papert statt, welche die Grenzen einfacher Perceptrons aufzeigten. Ihre Schlussfolgerung, dass auch mächtigere Modelle als das Perzeptron die gleichen Probleme aufweisen, stellte die Forschung auf dem ganzen Gebiet der neuronalen Netze in Frage, was aus heutiger Sicht falsch war. Sie bewirkten jedoch, dass die Forschungen auf dem Gebiet der neuronalen Netze und insbesondere der Computerwissenschaften für die nächsten Jahre fast zum Erliegen kamen.

1972 stellte Teuvo Kohonen in seinen Arbeiten „correlation matrix memories" [Koh72] ein Modell eines Assoziativspeichers vor. Christof von Malsberg veröffentlichte mit "Self-organization of orientation sensitve cells in the striata cortex" [Mal73] ein nichtlineares Neuronenmodell, bei dem sich die Zellen mit ihrem Lernverfahren ähnlich wie das neurobiologische Vorbild ausbilden. Speziell die Gruppe der Self-Organizing-Feature-Maps (SOM) wurde in den 80er Jahren von Teuvo Kohnen auf Basis der Überlegungen von Rosenblatt und Malsberg weiterentwickelt. Primäres Ziel der Anwendung von SOM besteht in der topologieerhaltenden Abbildung hochdimensionaler Merkmalsräume in einem Ausgaberaum niedriger Dimension.

Die vierte Phase wurde maßgeblich durch den Einfluss von John Hopfield geprägt. In seinem Artikel „Neural Computation of Decisions in Optimization Problems“ [Hob85] zeigte er, dass mit dem Hopfieldsnetz Optimierungsaufgaben gelöst werden können. Des Weiteren waren entscheidende Publikationen der Backpropagation Lernalgorithmus für Multilayer Perceptrons (multilayer feedforward networks), vorgeschlagen von Werbos, und letztendlich von Rumelhart „Learning representations by back-propagation errors“ [Rum86] 1986 publiziert.

4.1.1 Das biologische Vorbild

Die Entwicklung künstlich neuronaler Netze wurde ursprünglich durch den Versuch geprägt, Funktionsweisen des menschlichen Gehirns nachzubilden. Die elementare Verarbeitungseinheit in der Hirnrinde sind Nervenzellen, die als Neuronen bezeichnet werden. In der Großhirnrinde des Menschen ist eine flache Schicht von Nervenzellen mit ca. 2 bis 3 Millimeter Stärke und einer Fläche von etwa 2.200 cm^2. Die Anzahl der Neuronen in der Großhirnrinde wird auf rund 10^{11} geschätzt [Kil92] [Hof92]. Alle Neuronen mit Ihren Verbindungen zu den Nachbarneuronen bilden ein neuronales Netz. Jedes Neuron ist mit durchschnittlich 10^3 bis 10^4 Nachbarneuronen verbunden, so dass ein menschliches Gehirn im Neokortex bis zu 10^{14} Verbindungen besitzt, wodurch diese eine hochgradig parallele Struktur ausbilden. Die Verbindungsvarianten zwischen den Neuronen können mit $10^{1.000.000}$ Erregungszuständen beziffert werden [Ede98].

Ein Neuron (Nervenzelle) ist eine spezielle biologische Zelle, die in der Lage ist, Informationen zu verarbeiten. Das Neuron besteht aus einem Zellkörper (Soma), der eigentlichen Einheit zur Informationsverarbeitung, und zwei verschiedenen Typen von nach außen verzweigten Ästen: den Axonen und den Dendriten. Der schematische Aufbau ist in Abbildung 4-1 dargestellt.

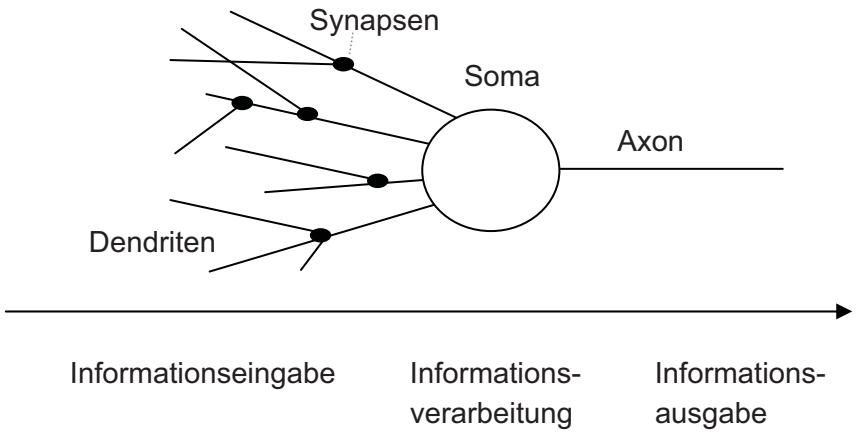


Abbildung 4-1: Schematische Darstellung eines biologischen Neurons

Die Dendriten leiten die Information an das Soma. An diese sind Synapsen gekoppelt, die als Verbindungsknoten fungieren. Diese können entweder die Enden eines Axons, eines Neurons oder die Dendriten eines anderen Neurons miteinander koppeln. In seltenen Fällen erfolgt hierüber die Verbindung zwischen Dendriten unterschiedlicher Neuronen (dendrodritische Synapsen) oder gar zwischen Axonen (axoaxonische Synapsen). Synapsen lassen sich in zwei Arten gruppieren: die erregenden (exzitatorischen) Synapsen und die hemmenden (inhibitorischen) Synapsen. Die exzitatorischen Synapsen wirken mit ihrem empfangenen Signal besonders stark auf die Soma. Inhibitorische Synapsen hingegen wirken dämpfend in ihrem Übertragungsverhalten zwischen Eingangsignal und Übertragungsimpuls auf die Soma [Zel94]. Wenn ein Impuls eine Synapse erreicht, werden Chemikalien (Neurotransmitter) freigesetzt. Die Neurotransmitter verteilen sich über den Synapsenspalt und erhöhen oder vermindern die Tendenz der eigenen elektrischen Impulse. So wird eine Anpassung an die Umgebung realisiert. Die Effektivität kann durch die durchgehenden Signale angepasst werden, so dass die Synapse von der Umgebung „lernt“.

Die von den Dendriten an die Soma übermittelten Signale werden dort verarbeitet. Der Zellkörper besitzt hierfür ein Nucleid, welches Informationen über vererbliche Züge beinhaltet, und Plasma, welches das molekulare Werkzeug zur Erzeugung des für das Neuron notwendigen Materials enthält. Ein Neuron erhält über seine Dendriten Signale (Impulse) von anderen Neuronen und sendet Signale, welche vom Zellkörper erzeugt werden, durch das Axon an andere Neuronen. Es fungiert somit als Aggregator unterschiedlicher Eingangssignale. Wird dabei ein kritischer Schwellwert der Eingangssignale erreicht, so sendet das Soma seinen Ausgangsimpuls über das Axon zu einer anderen Nervenzelle. Die Art des Ausgangssignals wie beispielweise Intensität wie auch deren zeitliche Fortpflanzung hängt dabei von dem jeweiligen Soma ab. Über das Axon werden die Informationen an weitere Neuronen weitergeleitet.

Inspiriert von biologischen neuronalen Netzen wurden bereits unzählige Fortschritte bei der Entwicklung intelligenter Systeme erzielt. Die Wissenschaft beschäftigt sich mit neuronalen Netzen in den unterschiedlichsten Bereichen wie Klassifikation, Approximation von Funktionen, Optimierung usw. Im Laufe der Evolution wurden dem menschlichen Gehirn viele erwünschte Eigenschaften zuteil, welche in von-Neumann-Rechnern nicht anzutreffen sind.

Gegenüber dem biologischen Vorbild ist die Simulation von Neuronen in künstlichen neuronalen Netzen sehr stark idealisiert. Implementationen künstlicher neuronaler Netze ähneln ihren biologischen Vorbildern in einer ganzen Reihe von Eigenschaften [Zel04] [Sto00]. Sie verfügen über eine massive Parallelität, das heißt, über eine große Anzahl von Verarbeitungseinheiten, die gleichzeitig und nicht wie in den klassischen von-Neumann-Rechnerstrukturen sequenziell arbeiten. Die Grundelemente der Netze bestehen aus recht einfachen Elementen, die erst durch die Verknüpfung und Anordnung mit umliegenden Elementen sehr komplexe Aufgaben übernehmen können.

Gerichtete Verbindungen zwischen Neuronen über Synapsen werden in der Simulation durch Verbindungen mit Gewichten dargestellt. Die Plastizität der Synapsen des biologischen Vorbildes wird während des Lernvorganges durch Anpassung der Gewichte erzielt. Das neuronale Netz verfügt über eine recht hohe Konnektivität, das heißt zwischen den einzelnen Neuronen bestehen eine Vielzahl von Verbindungen.

Die wesentlichen Unterschiede von künstlichen neuronalen Netzen zu biologischen Netzen sind die geringere Anzahl von Neuronen und Verbindungen zwischen den Elementen. Des Weiteren werden keine zeitlichen Phänomene der synaptischen Kopplung oder Einflüsse auf Neurotransmitter und auch zeitlichen Signalabläufe der Nervenleitungen berücksichtigt. Die chemischen Einwirkungen räumlich benachbarter Neuronen werden ebenfalls nicht berücksichtigt. Die in biologischen Netzen auftretenden Unplausibilitäten im Laufe des Lernvorganges werden bei der Simulation ebenfalls nicht berücksichtigt.

Es gibt sicherlich noch eine Vielzahl weiterer Unterschiede zwischen künstlichen neuronalen Netzen und deren biologischen Vorbildern. Festzuhalten ist dabei jedoch, dass derzeit viele technische Fragestellungen mit neuronalen Netzen zu lösen sind. Derzeit ist nicht zu erkennen, dass biologisch adäquate Simulationen für diese Gruppe von Problemen bessere oder gar schnellere Lösungen bieten [Zel04]. Für die weitere Vorgehensweise wird das Neuron als eine Prozessor-Einheit (Unit) verstanden, die Informationen in einfacher Weise verarbeitet und deren mathematische Beschreibung einfach möglich ist.

4.1.2 Formal mathematische Beschreibung eines Neurons

Das Grundmodell eines Neurons stützt sich im Wesentlichen auf die Vereinfachungen von McCulloch und Pitts aus dem Jahre 1943 [McP43]. In Ihren Publikationen wurden Neuronen als eine Art Addierer mit Schwellwert betrachtet. Trotz der immensen Unterschiede in den Architekturen, Arbeitsweisen, Lernalgorithmen und Anwendungsgebieten der unterschiedlichen Netztypen lassen sich einige ureigene Grundelemente identifizieren, die in Abbildung 4-2 dargestellt sind. Diese Grundelemente der verschiedenen Netztypen bestehen darin, dass jedes Netzwerk letztlich aus einer Vielzahl eigener Verarbeitungseinheiten, sogenannten Units besteht, welche hochgradig miteinander verknüpft sind.

Die Informationsverarbeitung innerhalb eines Neurons erfolgt mit drei Schritten; der Bildung des Nettoeingangssignals, der Ausbildung des Aktivierungsniveaus und letztlich der Bildung des Outputs. Nach Anlegen eines Eingangsvektors $\bar{x}$, der aus den Werten von m unabhängigen Variablen besteht, erfolgt dessen Gewichtung. Die Bedeutung des Gewichtsvektors w_i kann in Abhängigkeit der zugrunde gelegten Aktivierungsfunktion (Sigmoidfunktion, harte Schwellwertfunktion) maßgeblichen

Einfluss auf den Eingangsvektor nehmen. Das zu bildende Nettosignal net_i berechnet sich bei vielen Netzwerken als Linearkombination des Eingangsvektors $\bar{x}$ des Neurons und dessen Gewichtsvektors w_i wie folgt:

$$\mathrm{net}_i = \sum_{j=0}^{m} w_{ij} x_i \quad \text{bzw. in Matrixschreibweise} \quad \mathrm{net}_i = w_i^T x \tag{4-1}$$

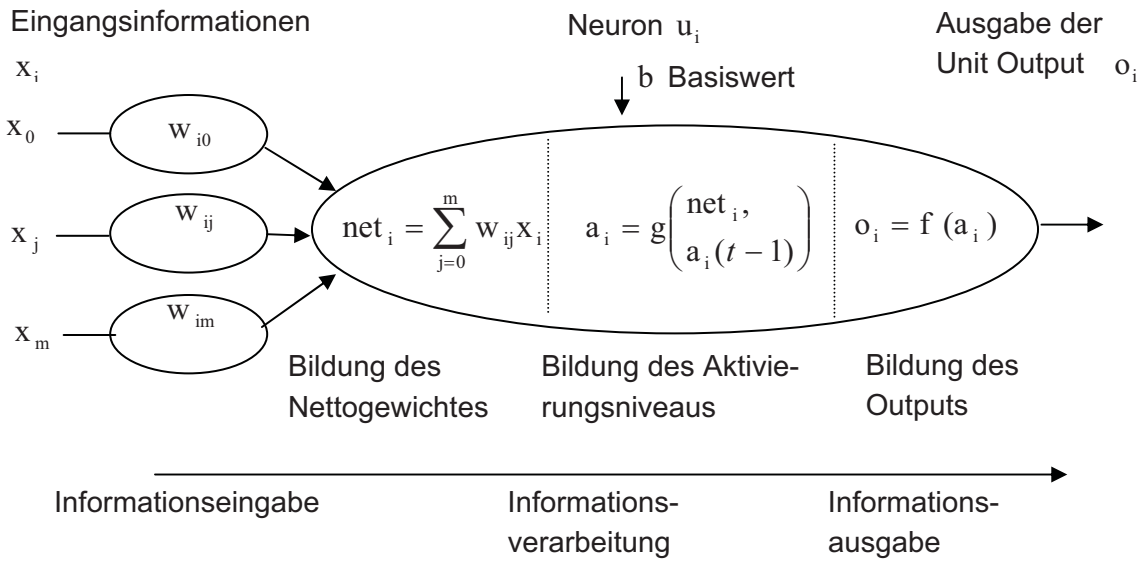


Abbildung 4-2: Schematische Darstellung eines künstlichen Neurons

Dabei kann die Berechnung des Nettogewichtes bei manchen Netzwerktypen (SOM) nach einem anderen Verfahren erfolgen. Das Nettogewichtssignal net_i wird in dem dann folgenden Verarbeitungsschritt in das Aktivierungsniveau überführt, in dem die Aktivierungsfunktion g dabei komplexe Formen annehmen kann, die vom Aktivierungsniveau a_i und den zeitlichen Verarbeitungsschritten $(t-1)$ abhängen. Das Aktivierungsniveau a_i berechnet sich wie folgt:

$$a_i(t) = g(\mathrm{net}_i, a_i(t-1)) \tag{4-2}$$

Man unterscheidet lineare und sigmoide Aktivierungsfunktionen sowie Schrittfunktionen. Die Schrittfunktionen geben ab einem Schwellwert θ_j, bei dem das entsprechende Neuron aktiv werden soll, den Wert 1 zurück. Diese Funktion lässt sich wie folgt beschreiben:

$$a_j = \begin{cases} 1; \text{falls} & \mathrm{net}_j \geq \theta_j \\ 0; & \text{sonst} \end{cases} \tag{4-3}$$

Schrittfunktionen sind allerdings nur für binäre Aktivierungen geeignet, da sie an der Sprungstelle nicht differenzierbar sind. Da diese Eigenschaft jedoch bei vielen Lernalgorithmen benötigt wird, kann durch Anwendung einer sigmoiden Aktivierungsfunktion

ein stetiger Übergang erzielt werden. Damit ist sichergestellt, dass ein neuronales Netz häufig sowohl auf Signale kleiner als auch sehr großer Amplituden reagiert, wobei es bei kleinen Amplituden wesentlich sensibler reagiert. Sigmoide Aktivierungsfunktionen haben daher ihre größte Sensibilität, d. h. ihre größte Steigung, im Bereich um den Arbeitspunkt. Die populärsten Funktionen [Bra95] sind zurzeit die logarithmische Aktivierungsfunktionen und der Tangens Hyperbolicus. Für diese Aktivierungsfunktionen gilt:

$$f_{log}(x) = \frac{1}{1+e^{-x}} \qquad f_{tanh}(x) = \tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{4-4}$$

Die Berücksichtigung einer zeitlichen Abhängigkeit des Aktivierungsniveaus a_i vom vorhergehenden Verarbeitungsschritt erlaubt es, in Verbindung mit einer geeignet gewählten Aktivierungsfunktion g ein komplexes, dynamisches Eigenverhalten in das Netzwerk zu induzieren. Hierdurch entsteht ein ausgesprochen schwierig zu behandelndes Modell, weshalb sich in vielen Netztypen die Vereinfachung nach (4-5) finden lässt [Pol05]:

$$a_i = net_i \tag{4-5}$$

Aus Vereinfachungsgründen wird im Folgenden von $a_i = net_i$ ausgegangen. Der Basiswert b der Unit ist der Schwellwert für die Aktivierungsfunktionen. Dieser Wert gibt an, ab wann ein Neuron (stark) aktiv ist. Es ist auch möglich, den Basiswert durch ein sogenanntes „on“-Neuron zu ersetzen. Um dieses zu realisieren, verwendet man einen zusätzlichen Eingang, der immer den Wert 1 erhält und ebenfalls über Gewichte mit allen inneren Neuronen verbunden ist. Diese Gewichte müssen dann allerdings zusammen mit den anderen Gewichten trainiert werden [Daw05].

Der Ausgabewert des Neurons o_i ist das Ergebnis der Outputfunktion f auf das Aktivierungsniveau a_i und den Basiswert b. Die Outputfunktion ist im Regelfall nicht linear. In Abhängigkeit des gewählten Netzes ist diese für das gesamte Netzwerk einheitlich gewählt oder für jedes Neuron mit unterschiedlichen Outputfunktionen ausgestattet.

$$o_i = f(a_i + b) \tag{4-6}$$

Zusammengefasst können die einzelnen vorgenannten Teilschritte der Verarbeitung miteinander und untereinander zusammengefasst werden. Unter Berücksichtigung der Vereinfachungsregel aus (4-6) ergibt sich folgende mathematische Beziehung, mit der ein Neuron verdichtet beschrieben werden kann:

$$o_i = f\left(\left(\sum_{i=1}^{n} w_{ij} x_i\right) + b\right) \tag{4-7}$$

4.2 Aufbau neuronaler Netze und Netztopologien

Die Leistungsfähigkeit von neuronalen Netzen entsteht erst durch die Verknüpfung vieler Neuronen, welche auf verschiedenste Weise miteinander verbunden sein können. Zur besseren Strukturierung sowohl der Netzwerkerläuterungen wie auch der graphischen Darstellung neuronaler Netze werden diese in der Regel als gerichtete Graphen dargestellt. Hierbei werden die einzelnen Neuronen (ohne Gewichte der Eingänge) als Knoten und die Gewichte der Eingänge, Neuronenverbindungen und der Informationsfluss als gewichtete Kanten in einem Graphen betrachtet.

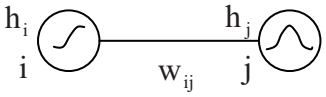


Abbildung 4-3: Darstellung eines neuronalen Netzes als gewichteter Graph.

In Abbildung 4-3 wird ein neuronales Netz als gewichteter Graph dargestellt. Die Gewichte beziehungsweise die Beziehungen der Gewichte werden direkt an die zugehörige Kante des Modells geschrieben. Für gewöhnlich ist es ausreichend, die Netztopologie, die Art der Neuronen sowie deren exakte Verbindungsstruktur zu visualisieren. Abbildung 4-3 zeigt die gewichtete Verbindung w_{ij} zwischen dem Neuron i mit sigmoider Aktivierungsfunktion h_i und dem Neuron j mit symmetrischer Aktivierungsfunktion h_j. Abbildung 4-4 verdeutlicht, dass sich Neuronen innerhalb des Netzes durch Ihre Position eindeutig beschreiben lassen. Beim Aufbau der neuronalen Netze gibt es im Wesentlichen drei Arten von Schichten, die Eingabeschicht mit den Eingangsneuronen, die inneren bzw. verdeckten Schichten, die auch als „hidden untis“ bezeichnet werden, und die Ausgabeschicht.

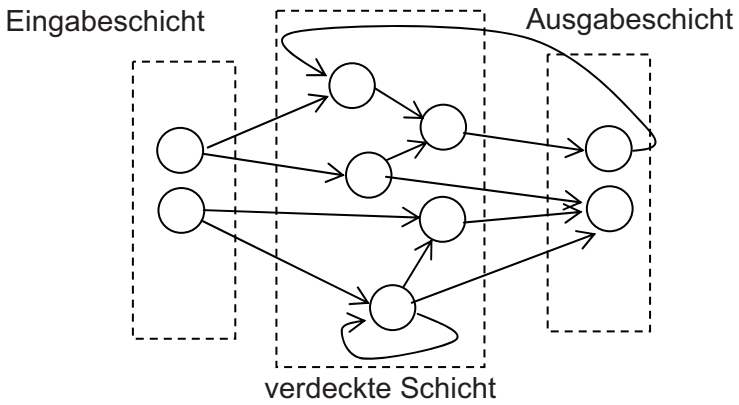


Abbildung 4-4: Schichtenmodell

Da die Eingabeschicht in der Regel nur als Informations- und Signalquelle dient, sind die Neuronen dieser Schicht mit keinerlei Funktionalität ausgestattet [Tag01]. Neuronen der verdeckten oder inneren Schicht erhalten Ihre Eingaben ausschließlich von anderen Neuronen des Netzwerkes und geben ihre Outputs ausschließlich an andere Units

desselben Netzwerkes weiter. In dieser Schicht des Netzes erfolgt hauptsächlich die Verarbeitung des Eingangsmusters. Ihre Ausprägung, d. h. Anzahl und Art der Verbindungen, ist für die Leistungsfähigkeit des gesamten Netzes maßgeblich. Die Ausgabeschicht enthält ihre Eingaben von der verdeckten Schicht und repräsentiert letztlich das Ergebnis des gesamten Netzwerkes.

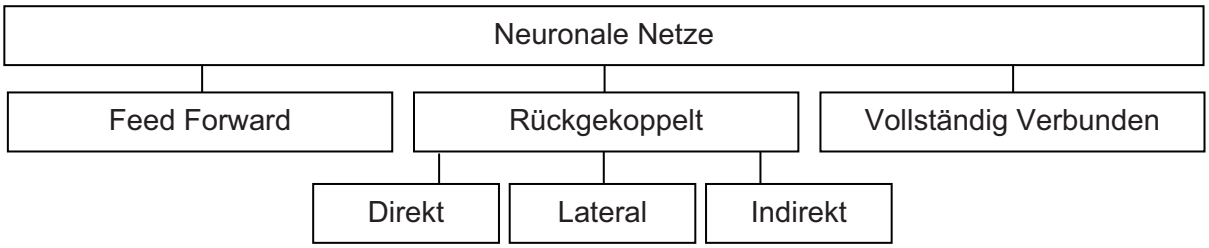


Abbildung 4-5: Topologische Klassifikation

Sämtliche Netzparameter, die Aufbau und interne Struktur beschreiben, werden unter dem Begriff Netztopologie zusammengefasst. Die Klassifikation von Netztopologien ist in Abbildung 4-5 vorgenommen. Die Beschreibung der Parameter kann wie folgt vorgenommen werden:

- Hierarchische Struktur und Vernetzung
- Definition, welche der Verbindungen fest bzw. adaptierbar sind
- Typ der Neuronen
- Anzahl der Neuronen
- Netztypspezifische Parameter

Das wesentliche Merkmal neuronaler Netze ist die hierarchische Netzstruktur, die festlegt, welche Neuronen miteinander verbunden sind und den Informationsfluss definiert. In der Praxis sind Netze mit unstrukturiertem Aufbau, wie in Abbildung 4-4 dargestellt, unüblich [Tag01]. Häufig werden Netzstrukturen mit einer oder mehreren Neuronenschichten eingesetzt, wie in Abbildung 4-6 ersichtlich, was als n-schichtige Multi-Layer-Perceptron (MLP) bezeichnet werden.

Bei der Definition der Anzahl der Neuronenschichten herrscht in der Literatur kein einheitliches Bild, ob z. B. die Eingabeschicht, die häufig keine weitere Funktionalität besitzt, bei der Zählung mit berücksichtigt wird. Im Rahmen der vorliegenden Arbeit wurde die Konvention getroffen, die Eingabeschicht mitzuzählen. Neben der Klassifikation der Schichten ist eine weitere Einteilung der Netztopologien möglich. Hierbei werden der Informationsfluss innerhalb des Netzes und die Verknüpfung der Neuronen betrachtet. Sind Neuronen eines Netzes vollständig miteinander verbunden, so wird dieses Netz als vollständig verknüpft bezeichnet. Neuronenverbindungen über eine oder gar mehrere Ebenen hinweg werden als „cross connections“ bezeichnet.

Ein weiteres Merkmal ist die Richtung des Informationsflusses. Die Definition der Flussrichtung von Information bezieht sich dabei auf die Anwendungsphase, denn die Lernphase, bei der die Gewichte eines Netzes neu gesetzt werden, dient zur Definition der Netztopologien. Netze werden als vorwärtsgerichtete oder Feedforward-Netze bezeichnet, wenn der Informationsfluss vom Eingang zum Ausgang läuft. Besteht ein umgekehrter Informationsfluss in der Anwendungsphase, vom Ausgang zum Eingang, ergibt sich ein rückgekoppeltes oder rekurrentes Netz. Zusätzlich können zur Richtung des Informationsflusses in bestimmten Netztopologien feste Verbindungen existieren, die während der Lernphase nicht adaptiert werden.

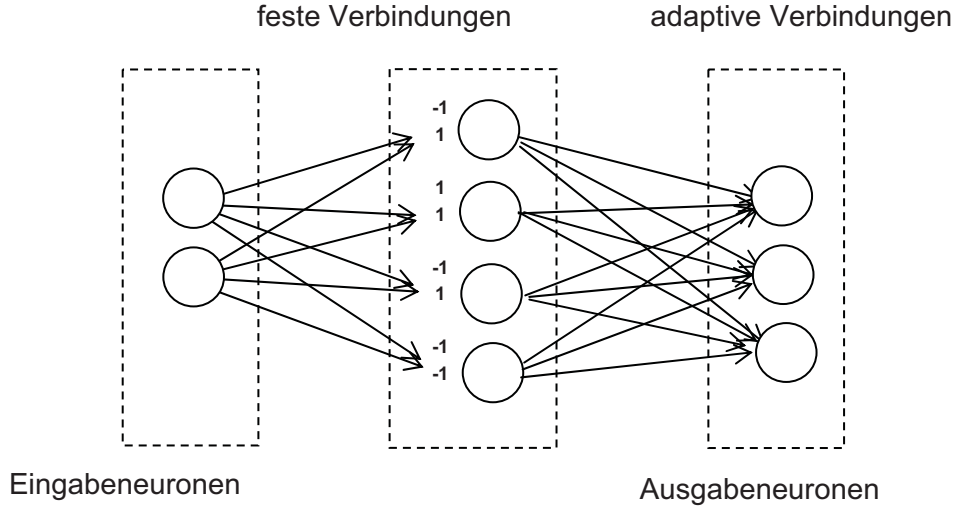


Abbildung 4-6: n-schichtige Multi-Layer-Perceptron (MLP)

Der Zweck, unterschiedliche Netztopologien anzuwenden, liegt darin begründet, dass bestimmte Netztypen für spezielle Problemstellungen besser geeignet sind. In den folgenden Abschnitten soll nun ein kurzer Überblick über die im Rahmen der Untersuchungen eingesetzten Netztopologien (Feedforward-Netze, Elman- und Jordan-Netze) sowie unterschiedlichen Netzstrukturen gegeben werden.

4.2.1 Feedforward-Netze

Das Feedforward-Netz oder auch Multilayer Perzeptron (MLP) besteht aus einer Eingabeschicht und einer Ausgabeschicht. Dazwischen befinden sich beliebig viele innere Schichten (Verarbeitungsschichten) von Neuronen (Abbildung 4-7). Die Verbindungen von jedem Neuron dürfen bei einem Feedforward-Netz ausschließlich ein Neuron der nächsten Schicht treffen. Die Neuronen einer Schicht sind jeweils mit jedem Neuron der folgenden Schicht verbunden. Diese Verbindung ist jedoch nur nach vorne gerichtet und es gibt keinerlei Rückkopplungen. Der Signalfluss läuft in der Anwendungsphase im gesamten Netz nur in eine Richtung, wobei die Neuronen nur Signale aus den vorgelagerten Schichten empfangen und nur Signale an die nach-

gelagerte Schichten weitergeben. Es hat sich gezeigt, dass für die meisten Problemstellungen eine Schicht innerer Neuronen völlig ausreichend ist, wenn für diese Schicht genügend Neuronen gewählt werden.

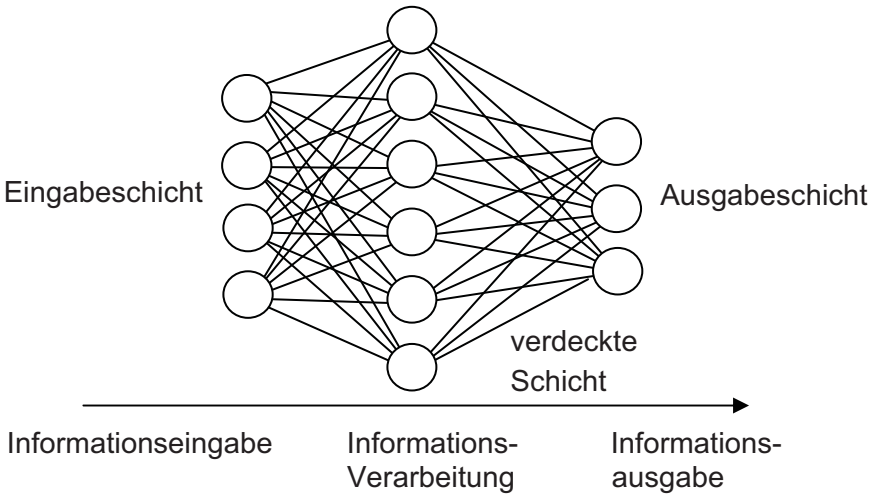


Abbildung 4-7: Feedforward-Netz mit drei Neuronenschichten

Für die Erkennung und Klassifikation zeitveränderlicher Muster und für Prognosen von Zeitreihen sind partiell rekurrente Netztopologien besonders geeignet [Moz93]. Diese Netztopologien zeichnen sich durch ihre Kontext-Einheiten aus. Kontext-Einheiten sind Neuronen, die sich quasi auf gleicher Ebene wie die Neuronen der Eingabeschicht befinden und von Neuronen der Zwischen- oder Ausgabeschicht Informationen erhalten und an diese Schichten weiterleiten. Diese Informationen werden dann dort verarbeitet und um einen Schritt verzögert an die Zwischenschicht zurückgegeben. Dabei entspricht die Anzahl der Kontext-Einheiten genau der Anzahl der Einheiten der Ausgabeschicht oder der Zwischenschichten, wobei zwischen den Einheiten eine 1:1-Verbindung besteht.

Die Gewichte der Kontext-Zellen werden kontinuierlich auf Eins gesetzt, um sicherzustellen, dass eine exakte Kopie des Ausgangs- oder der Zwischenschicht dem Eingang zugeführt wird. Kontext-Einheiten mit den dazugehörigen modifizierbaren Gewichten werden auch in der Literatur als dynamisches Gedächtnis des neuronalen Netzes bezeichnet [Cic93] [Rey08], denn ein identischer Input im Netz wird durch die Existenz der Kontextabhängigkeit modifiziert. Zur Gruppe der rekurrenten Netze können Jordan- und Elman-Netze gezählt werden.

4.2.2 Jordan- und Elman-Netze

Bei den Jordan-Netzen handelt es sich um Netze einer einfachen Feedforward-Netzarchitektur, die mit Kontext-Zellen erweitert wurden. Eine solche Netzarchitektur ist in Abbildung 4-8 dargestellt.

Die Kontextzellen speichern den Ausgabezustand und geben diesen zusammen mit den Eingabezellen auf die inneren Neuronen. Dabei erfolgt die Speicherung der Ausgabe über eine 1:1-Rückkopplung der Ausgangsneuronen. Zusätzlich besitzen die Kontextzellen eine direkte Rückkopplung auf sich selbst mit einer festen (nicht trainierbaren) Stärke von λ. Die Verbindungen zu den inneren Neuronen jedoch werden zusammen mit den Gewichten der Eingangsneuronen trainiert. Alle Verbindungen sind unidirektional. Zu Beginn der Verarbeitung von Eingabemustern werden die Aktivierungsfunktionen der Kontextzellen auf einen festen Wert gesetzt.

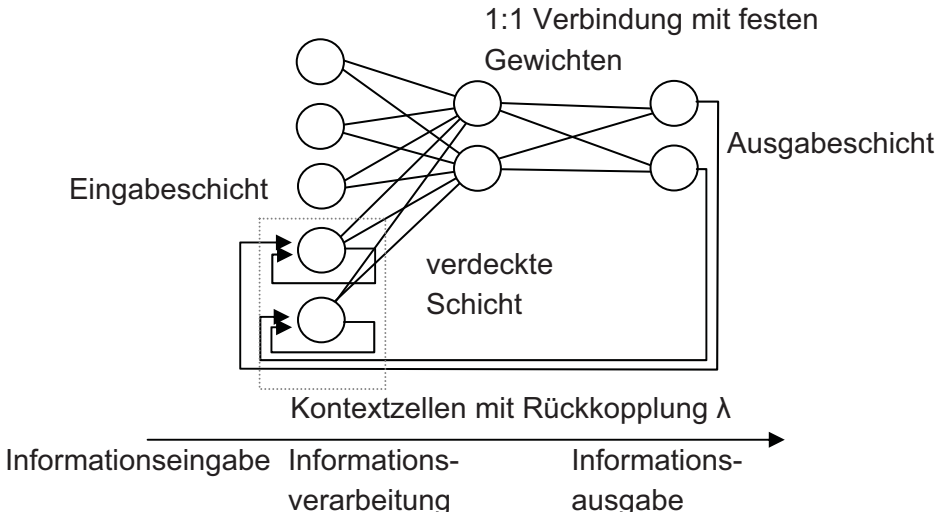


Abbildung 4-8: Jordan-Netz

Durch diese Anordnung kann eine zeitliche Abfolge der Eingabesignale berücksichtigt werden. Werden die zeitabhängigen inneren Zustände des Netzes mit dem Vektor $S(t)$ bezeichnet, so kann man das Verhalten des Netzes durch die folgenden beiden Gleichungen beschreiben:

$$O(t) = F(S(t), I(t)) \tag{4-8}$$

$$S(t+1) = G(S(t), I(t)) \tag{4-9}$$

Dabei ist die Funktion F die Aktivierungsfunktion, O die Ausgabefunktion und G die Übergangsfunktion des Netzes. Beide Funktionen hängen vom aktuellen Zustand $S(t)$ und von der Eingabe $I(t)$ ab und damit auch der Folgezustand $S(t+1)$ vom aktuellen Zustand $S(t)$ und der Eingabe $I(t)$, so dass sich der Folgezustand aus dem momentanen Zustand und der Ausgabe ergibt.

Die Struktur des Jordan-Netzes weist gegenüber dem ähnlichen Elman-Netz den Nachteil auf, dass die Zahl der Rückkopplungen durch die Anzahl der Ausgabeneuronen und somit durch die Problemstellung festgelegt wird. Beim Elman-Netz werden hingegen anstatt der Ausgabeneuronen die Ausgaben der verdeckten Schicht

rückgekoppelt, womit die Einschränkung des Jordan-Netzes aufgehoben wird und im Allgemeinen mit Elman-Netzen bessere Resultate erzielt werden [Zel94].

Die nach deren Entwickler Jeffrey J. Elman benannten Elman-Netze sind somit eigentlich modifizierte Jordan-Netze. Elman-Netze stellen auch eine Erweiterung normaler Multilayer-Netzwerke dar, indem die Aktivität der versteckten Schicht nicht nur an eine Outputschicht, sondern auch an die Kontextschicht (siehe Abbildung 4-9) weitergeleitet wird.

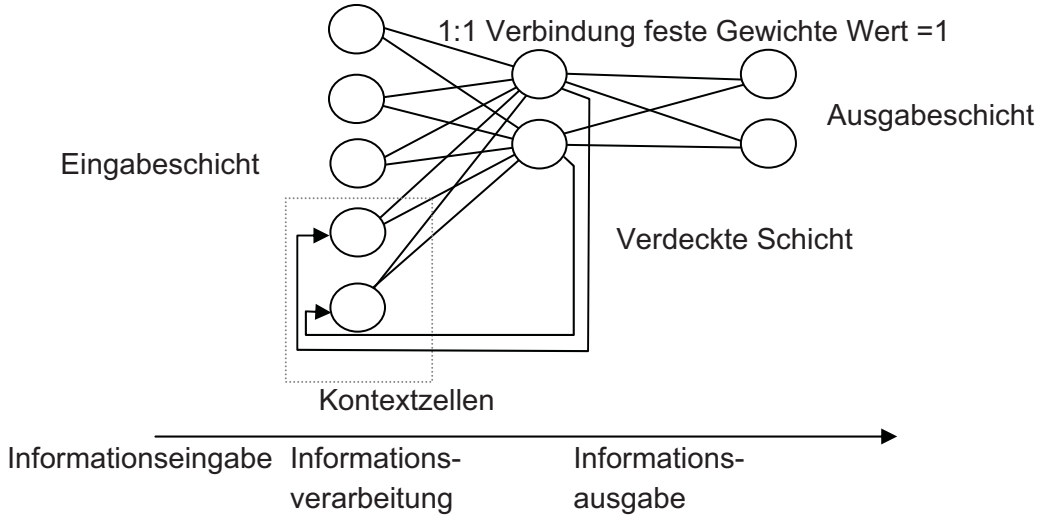


Abbildung 4-9: Elman-Netz

Dazu ist jede Zelle der Kontextschicht mit einer Zelle der versteckten Schicht verknüpft. Im Gegensatz zum Elman-Netz entfallen hierbei auch die Rückkopplungen der Kontextschicht auf sich selbst. Während die Verbindung von der versteckten Schicht zur Kontextschicht fest ist, d. h. deren Aktivität im Verhältnis 1:1 weitergeleitet wird, sind auch hier die Gewichte der Verbindungen von der Kontextschicht zur versteckten Schicht variabel und können während des Lernvorganges adaptiert werden. Elman-Netze speichern in den Kontextzellen die Aktivierung der Zellen der versteckten Schicht des vorangegangenen Musters als internen Zustand, der im folgenden Verarbeitungsschritt das eingehende Muster beeinflusst. Somit hängt wie beim Jordan Netz das Ausgangssignal nicht ausschließlich vom Eingangssignal ab, sondern berücksichtigt bereits präsentierte Muster.

Im Vergleich zu den Jordan-Netzen bieten die Elman-Netze den Vorteil, dass sie nicht an das Ausgabemuster gebunden sind, d. h. die Anzahl der Kontextzellen kann von den Ausgabeeinheiten abweichen, was eine wesentlich flexiblere Verarbeitung unterschiedlicher Ein- und Ausgabemuster ermöglicht. Einfache Elman-Netze bestehen aus nur einer verdeckten Schicht. Es können aber auch Netze aus mehreren verdeckten Schichten erstellt werden. Diese werden dann als hierarchische Elman-Netze bezeichnet, bei denen zu jeder verdeckten Schicht und der Ausgabeschicht Kontextzellen

angelegt werden. Diese Netztopologien werden für komplizierte Problemstellungen eingesetzt [Daw01].

4.3 Lernalgorithmen

Die Arbeitsweise eines neuronalen Netzes kann in die Ausführungsphase (Recall-Modus) und in den Lernmodus klassifiziert werden. Im Folgenden werden die für die Untersuchungen relevanten Lernverfahren kurz dargestellt.

4.3.1 Das Lernproblem

Innerhalb des Lernmodus modifiziert sich das neuronale Netz gemäß einer fest vorgeschrieben Lernregel selbst. Die Lernregel ist dabei ein Algorithmus, der das neuronale Netz verändert. Im Lernmodus werden Gewichte, Schwellwerte, Anzahl der Neuronen etc. auf Basis vorhandener Daten dem Netz zur Verfügung gestellt und somit vom Netz erlernt. Der Lernvorgang kann prinzipiell folgende Schritte beinhalten:

- Entwicklung neuer Verbindungen
- Löschen existierender Verbindungen
- Modifikation der Verbindungsstärke (Veränderung der Gewichte)
- Modifikation des Schwellenwertes
- Modifikation der Aktivierungs- bzw. Ausgabefunktion
- Entwicklung neuer Zellen
- Löschen bestehender Zellen

Die Lernverfahren lassen sich in drei Gruppen einteilen. Beim sogenannten überwachten Lernen (supervised learning) wird das Netz zunächst mit einem Datensatz konfrontiert, bei dem Eingangs- und auch Ausgangsdatensatz eines Paares bekannt sind. Dem Netz wird der Eingangsdatensatz präsentiert und das vom Netz erzeugte Ausgabemuster mit dem Ausgabedatensatz verglichen. Tatsächliche und korrekte Ausgaben werden verglichen. Nach jedem Lernschritt wird der Netzausgang mit dem ursprünglichen Ausgangswert der Originaldaten verglichen. Anhand der Differenz dieser beiden Werte werden die Gewichte einer Lernregel entsprechend angepasst. Innerhalb der Gruppe des bestärkenden Lernens (reinforcement learning) erfolgt im Gegensatz zum überwachten Lernen kein Vergleich von Ausgangs- und Eingangsdatenmusterpaaren. Im Rahmen dieser Verfahren wird dem Netz hier lediglich mitgeteilt, ob seine Ausgabe korrekt oder inkorrekt war, auf die Rückführung einer Differenz wird verzichtet. Bei der letzten Gruppe, dem unüberwachten Lernen (unsupervised learning), versucht das Netz selbstständig Zusammenhänge in den Lerndaten zu detektieren (self-organized learning). Das Netz versucht, ohne Beeinflussung von außen die präsentierten Daten in Ähnlichkeitsklassen aufzuteilen. Die Korrektur der Netzgewichte wird nicht von außen veranlasst, sondern das Netz be-

stimmt durch eine ‚Bewertungsfunktion', ob eine Änderung notwendig ist. Die Gewichte werden hier durch Zufallszahlen leicht verändert, und die Bewertungsfunktion stellt fest, ob der neue Ausgabewert besser ist als der vorherige.

In der Praxis erfolgt das Lernen überwiegend auf Basis der Lernregel von Hebb. Sie besagt, dass wenn eine Zelle j eine Eingabe von Zelle i erhält und beide gleichzeitig stark aktiviert werden, das Gewicht beider w_{ij} verstärkt wird. Dabei ist Δw_{ij} die Änderung des Gewichtes w_{ij} zwischen dem Eingang e_i und Ausgang a_j.

$$\Delta w_{ij} = \sigma \cdot e_i \cdot a_j \qquad (\sigma : \text{Lernrate}) \tag{4-10}$$

In der Praxis wird eine Modifikation der Hebb-Regel verwendet, die so genannte Delta-Regel. Anstatt des Ausganges a_j wird $\Delta a_j = (a_j^{\text{aktuell}} - a_j^{\text{erwartet}})$ verwendet:

$$\Delta w_{ij} = w_{ij}^{\text{neu}} - w_{ij}^{\text{alt}} = \sigma \cdot e_i \cdot (a_j^{\text{aktuell}} - a_j^{\text{erwartet}}) \tag{4-11}$$

Bei der Delta-Regel erfolgt die Gewichtsänderung proportional zum aktuellen und dem erwarteten Ausgangswert von Neuron j. Stimmen diese beiden Werte überein, so wird das Gewicht nicht weiter verändert.

4.3.2 Gradientenverfahren

Bei diesem Verfahren wird der Gradient einer Zielfunktion, in diesem Fall die Fehlerfunktion, berechnet. Der Gradient ist definiert als ein Vektor der partiellen Ableitungen der Funktion $f(x_1, x_2, ..., x_n)$. Der positive Gradient zeigt in Richtung des stärksten Anstieges dieser Funktion und der negative Gradient in Richtung des steilsten Abstieges. Die Fehlerfunktion $E(W) = E(w_1, w_2, ..., w_n)$ beschreibt den auftretenden Fehler der Prognosewerte eines neuronalen Netzes in Abhängigkeit der Gewichte w_{ij}. Durch Änderung aller Gewichte um einen Bruchteil des negativen Gradienten der Fehlerfunktion soll möglichst schnell ein Minimum dieser Funktion gefunden werden. Es gilt:

$$\Delta W = -\eta \nabla E(W) \tag{4-12}$$

η ist der Lernfaktor bzw. die Schrittweite mit der der Lernprozess durchgeführt wird, $-\nabla E(W)$ ist der negative Gradient der Fehlerfunktion. Für ein einzelnes Gewicht des Netzes gilt somit:

$$\Delta w_{ij} = -\eta \frac{\partial}{\partial w_{ij}} E(W) = \sum_p -\eta \frac{\partial}{\partial w_{ij}} E_p \tag{4-13}$$

Als Fehlerfunktion E wird häufig der quadratische Abstand zwischen Erwartungswert $a^{erwartet}$ und Prognosewert $a^{trainiert}$ verwendet. Der Gesamtfehler ergibt sich als Summe der Fehler über alle Datensätze p:

$$E = \sum_p E_p \text{ mit } E_p = \frac{1}{2}\sum_j (a_{jp}^{trainiert} - a_{jp}^{erwartet})^2 \tag{4-14}$$

und über alle Ausgangsneuronen j.

4.3.3 Backpropagation

Das Backpropagation-Verfahren ist ein Lernalgorithmus, der auch für mehrstufige nichtlineare Netze angewendet werden kann. Das Verfahren beruht grundsätzlich auf einer Gradienten-Abstiegsmethode, die 1986 erstmalig von Hilton, Rumelhart, und Williams vollständig beschrieben wurde [Hil86]. Bei diesem Lernverfahren sind Ein- und Ausgang des assoziierenden Musters bekannt und die Verbindungsgewichte eines Netzes werden im Laufe des Verfahrens geschätzt. Das Ziel des Algorithmus ist, eine möglichst gute Anpassung aller Verbindungsgewichte zwischen Eingabe und Ausgabe zu erreichen und den Ausgabefehler dabei zu minimieren. Der Lernvorgang lässt sich dabei in zwei Teilschritte gliedern. Im ersten Teilschritt werden die Eingabewerte durch das Netz zur Ausgabeschicht geleitet, um dort die Netzausgabewerte zu ermitteln. Im zweiten Schritt wird der Fehler zwischen Netzausgabewert und der vorgegebenen Ausgabe ermittelt. Auf Basis des Fehlers des jeweiligen Lernmusters erfolgt dann eine Anpassung der Verbindungsgewichte mit dem Ziel, den Fehler zu minimieren. Die Trainingsphase wird im Folgenden in Anlehnung an [Hil86] [Zel94] beschrieben. Sie gilt für Netze mit mehreren Ebenen und semilinearen (d. h. monotonen, differenzierbaren, jedoch im Allgemeinen nichtlinearen) Aktivierungsfunktionen. Der Eingangswert einer Zelle j bei Muster p ist gegeben durch:

$$net_{pj} = \sum_i a_{pi} w_{ij} \tag{4-15}$$

Für die Zelle j mit semilinearer Aktivierungsfunktion f gilt:

$$a_j = f(net_{pj}) \tag{4-16}$$

Analog der Vorgehensweise beim Gradientenverfahren, wie bereits in Abschnitt 4.3.2 (Gleichung 4-13) erwähnt, ergibt sich die gleiche Definition des Fehlers:

$$\Delta w_{ij} = \sum_p -\eta \frac{\partial E_p}{\partial w_{ij}} \tag{4-17}$$

Betrachtet man nun $\frac{\partial E_p}{\partial w_{ij}}$ und wendet die Kettenregel an, so ergibt sich:

$$\frac{\partial E_p}{\partial w_{ij}} = \frac{\partial E_p}{\partial net_{pj}} \cdot \frac{\partial net_{pj}}{\partial w_{ij}} \tag{4-18}$$

$\frac{\partial net_{pj}}{\partial w_{ij}}$ ist die Ableitung von Gleichung 4-15 nach w_{ij}, somit gilt:

$$\frac{\partial net_{pj}}{\partial w_{ij}} = \frac{\partial}{\partial w_{ij}} \sum_i a_{pi} w_{ij} = a_{pi} \tag{4-19}$$

Der erste Faktor der Gleichung (4-18) wird als Fehlersignal δ definiert:

$$\delta_{pj} = -\frac{\partial E_p}{\partial net_{pj}} \tag{4-20}$$

Gleichungen 4-19 und 4-20 werden in Gleichung 4-18 eingesetzt, daraus ergibt sich für die Gleichung 4-17:

$$\Delta w_{ij} = \eta \sum_p a_{pi} \delta_{pj} \tag{4-21}$$

Das Fehlersignal kann umformatiert werden:

$$\delta_{pj} = -\frac{\partial E_p}{\partial net_{pj}} = -\frac{\partial E_p}{\partial a_{pj}} \cdot \frac{\partial a_{pj}}{\partial net_{pj}} \tag{4-22}$$

Der zweite Faktor der Gleichung 4-16 lässt sich einfach beschreiben als erste Ableitung der Aktivierungsfunktion f von Neuron j:

$$\frac{\partial a_{pj}}{\partial net_{pj}} = \frac{\partial}{\partial net_{pj}} f\left(net_{pj}\right) = f'(net_{pj}) \tag{4-23}$$

Um das Δw_{ij} zur Laufzeit zu bestimmen, muss nur noch der erste Faktor aus Gleichung 4-20 bestimmt werden. Hierzu sind allerdings zwei Fälle zu berücksichtigen:

1. j ist Ausgangsneuron
2. j ist verdecktes Neuron

Im ersten Fall ergibt sich nach Definition von δ_{pj} für die Ableitung E_p

$$-\left(\frac{\partial E_p}{\partial a_{pj}}\right) = \left(t_{pj} - a_{pj}\right) \tag{4-24}$$

Im zweiten Fall kann die partielle Ableitung $-(\partial E_p)/(\partial a_{pj})$ nur indirekt berechnet werden, indem die Kettenregel der Gleichungen 4-15 und 4-20 zu Hilfe genommen werden.

$$-\frac{\partial E_p}{\partial a_{pj}} = -\sum_k \frac{\partial E_p}{\partial net_{pk}} \cdot \frac{\partial net_{pk}}{\partial a_{pj}} = \sum_k \left(\delta_{pk} \cdot \frac{\partial}{\partial a_{pj}} \sum_i a_{pi} w_{ik} \right) = \sum_k \delta_{pk} w_{jk} \qquad (4\text{-}25)$$

Der Gesamtfehler für Neuronen j des Musters p wird somit aus den gewichteten Fehlern δ_{pk} der Nachfolgeneuronen k und der Gewichte w_{ij} der Verbindungen von j zu k gebildet. Mit Hilfe der Gleichung 4-20 erfolgt die Summation aller Nachfolgezellen k und j, weil Neuron j bei allen diesen Nachfolgezellen einen Fehler verursacht. Es ergibt sich dann, mit Gleichung 4-20, im ersten Fall (j ist Ausgabezelle):

$$\delta_{pj} = f'_{act}(net_{pj}) \cdot (t_{pj} - a_{pj}) \qquad (4\text{-}26)$$

und im zweiten Fall (j ist verdeckte Zelle):

$$\delta_{pj} = f'_{act}(net_{pj}) \cdot \sum_k \delta_{pk} w_{jk} \qquad (4\text{-}27)$$

Diese beiden Gleichungen ermöglichen eine rekursive Berechnung der δ_{pj} für alle Zellen des Feedforward-Netzes, ausgehend von den Ausgabe- zu den Eingabezellen. Zusammenfassend gilt die Backpropagation-Regel:

$$\Delta_p w_{ij} = \eta a_{pi} \delta_{pj} \qquad (4\text{-}28)$$

mit

$$\delta_{pj} = \begin{cases} f'(net_{pj})\,(t_{pj} - a_{pj}); & \text{für } j \text{ Ausgangsneuron} \\ f'(net_{pj}) \sum_k \delta_{pk} w_{jk}; & \text{für } j \text{ verdecktes Neuron} \end{cases} \qquad (4\text{-}29)$$

Setzt man eine bekannte logarithmische Aktivierungsfunktion ein, ergibt sich für deren Ableitung der Aktivierungsfunktion folgender Zusammenhang:

$$f'(net_{pj}) = f(net_{pj})(1 - f(net_{pj})) = a_{pj}(1 - a_{pj}) \qquad (4\text{-}30)$$

Somit lautet die vereinfachte Form des Fehlersignals:

$$\delta_{pj} = \begin{cases} a_{pj}(1 - a_{pj})\,(t_{pj} - a_{pj}); & \text{für } j \text{ Ausgangsneuron} \\ a_{pj}(1 - a_{pj}) \sum_k \delta_{pk} w_{jk}; & \text{für } j \text{ verdecktes Neuron} \end{cases} \qquad (4\text{-}31)$$

Bei der Anwendung des Algorithmus ergeben sich oft einige Probleme, die auf das Gradientenverfahren, das ein lokales Verfahren ist und dadurch keine Informationen über die gesamte Fehlerfläche besitzt, zurückzuführen sind. Die Abbildung 4-10 verdeutlicht mögliche Probleme des Gradientenverfahrens.

a) In flachen Plateaus ist der Gradient nahe Null. Da die Gewichtsänderung von der Größe des Gradienten abhängig ist, stagniert der Algorithmus oder es sind sehr viele

Iterationsschritte durchzuführen, um ein Minimum zu erreichen. Des Weiteren lässt sich bei dieser Situation nicht differenzieren, ob ein lokales oder globales Minimum erreicht wurde oder das Verfahren auf dem flachen Plateau stagniert.
b) Das Verfahren kann in engen Schluchten der Fehlerkurve oszillieren. Das geschieht, wenn, bedingt durch einen hohen Gradienten am Anfang einer Schlucht, aufgrund der großen Gewichtsänderung, ein Überspringen der Schlucht erfolgt, und auf der entgegengesetzten Flanke der Schlucht sich dieser Vorgang wiederholt.
c) Das Gradientenverfahren detektiert ein lokales Minimum. Neuronale Netze mit hoher Dimension verfügen im Regelfall über eine sehr zerklüftete Fehlerfläche. Aufgrund dieser Struktur ist die Wahrscheinlichkeit recht hoch, innerhalb eines lokalen statt eines globalen Minimums zu landen [Hec90]. Hierzu gibt es nur wenig allgemeingültige Verfahren, die dieses Problem lösen können, da nicht nur vom Verfahren an sich, sondern auch von der Kodierung der Eingaben abhängt, in welchem Maße sich lokale und globale Minima ausbilden. Die praktische Anwendung hat jedoch gezeigt, dass bei der Wahl einer klein gewählten Schrittweite η immer ein Minimum gefunden wird, das für den Anwender ein durchaus akzeptables Ergebnis liefert.
d) Bei zu großer Schrittweite η kann es ebenfalls in engen Schluchten zum Verlassen eines guten Minimums kommen und somit ein schlechterer Fehler in Kauf genommen werden.

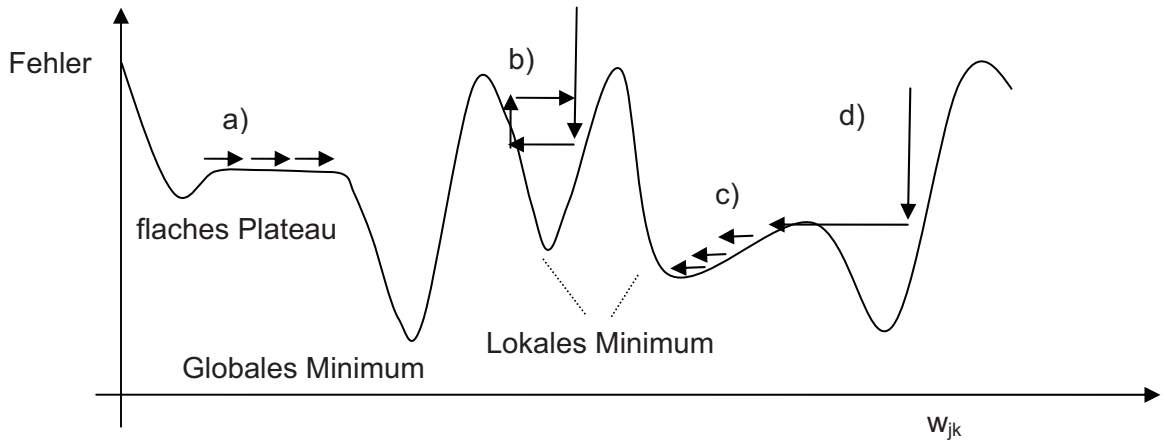


Abbildung 4-10: Probleme des Gradientenverfahrens

Die richtige Wahl des Lernfaktors η (Schrittweite) ist von entscheidender Bedeutung für das Verhalten des Algorithmus. Zu kleine η-Werte bringen oft sehr lange Laufzeiten des Algorithmus mit sich. Hingegen bewirken große Werte des Lernfaktors η große Sprünge auf der Fehlerfläche, mit dem Risiko, enge Täler zu übersehen, aus ihnen herauszuspringen oder Oszillation zu verursachen. In der Literatur gibt es widersprüchliche Aussagen zur Dimensionierung des Lernfaktors wobei unstrittig ist, dass dieser maßgeblich von den Trainingsdaten, der Größe des Netzes und der Topologie abhängt. Es ist also notwendig, bei Anwendung des Backpropagation-Verfahrens das Netz öfter mit diesem Algorithmus zu trainieren und hierbei die Schrittweite

zu variieren. Das Ergebnis kann andernfalls nicht eindeutig beurteilt werden. Um den oben genannten Problemen entgegenzuwirken, gibt es zahlreiche Modifikationen des Backpropagation-Algorithmus. Ein Beispiel ist das auch in den weiteren Untersuchungen eingesetzte Resilient Propagation (Rprop). Dieser Algorithmus verwendet eine eigene adaptive Schrittweite η_{ij} für die Bestimmung der Gewichte. Zudem wird für die Gewichtsbestimmung nur das Vorzeichen des Gradienten verwendet. Der Rprop hat den Vorteil gegenüber der Backpropagation, dass die Trainingsdauer erheblich gesenkt wird.

4.3.4 Feedforward-Algorithmus

Die Feedforward- oder auch Forwardpropagation-Algorithmen sind spezielle Lernverfahren für vorwärtsgerichtete neuronale Netze. Der Algorithmus lässt sich auf beliebige Forward-Netze anwenden und bietet Geschwindigkeitsvorteile gegenüber anderen Algorithmen bei Netzen mit genau einer verdeckten Schicht. Im Gegensatz zum Backpropagation-Algorithmus, der zuerst die Gewichte der Verbindungen zwischen Hidden- und Outputlayer modifiziert, beginnt der Feedforward-Algorithmus damit, die Eingangsgewichte zu bestimmen. Die Berechnung wird nacheinander für jeden Knoten des Hidden-Layers durchgeführt.

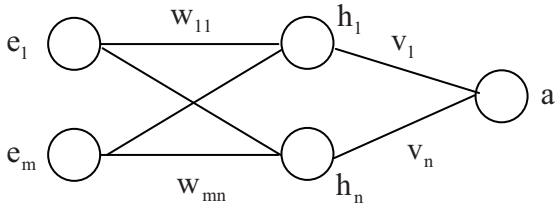


Abbildung 4-11: Struktur des Feedforward-Netzes

Es lässt sich mathematisch zeigen, dass eine Funktion, die in einem bestimmten Intervall stetig ist, durch ein neuronales Netz mit nur einer verdeckten Schicht abgebildet werden kann [Bra02]. Da die Einflussgrößen auf den Energiebedarf im Regelfall stetig verlaufen, ist durch die Verwendung dieses Algorithmus kein Nachteil bei der Erstellung von Lastprognosen zu erwarten. Anhand eines Beispielnetzes mit n Eingängen und einer verdeckten Schicht mit n Ausgangsneuronen wird im Folgenden ein Forwardpropagation-Algorithmus beschrieben, der zunächst iterativ die vorderen und dann die hinteren Gewichte bestimmt [Ger08].
Für das in Abbildung 4-11 dargestellte Netz gelten folgenden Beziehungen es stehen dabei die Lerndatensätze (e_{pk}, a_p) mit $p = 1,...,N$ und $k = 1,...,n$ zur Verfügung:

$$h_{pj} = f\left(\sum_{k=1}^{n} e_{pk} w_{kj}\right) \text{ mit } j = 1,...,m \tag{4-32}$$

Die Ausgangsfunktion eines Neurons lässt sich wie folgt beschreiben:

$$a_p = f\left(\sum_{j=1}^{m} h_{pj} v_j\right) \text{ mit } j = 1,...,m \tag{4-33}$$

Aus dieser Beziehung kann der Hilfsvektor $\vec{S}$ definiert werden:

$$S_p = f^{-1}(a_p) = \sum_{j=1}^{m} h_{pj} v_j \tag{4-34}$$

Die Summenform kann auch in einem linearen Gleichungssystem dargestellt werden:

$$\underbrace{\begin{pmatrix} h_{11} & \cdots & h_{1m} \\ \vdots & \cdots & \vdots \\ h_{N1} & \cdots & h_{Nm} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} = \begin{pmatrix} S_1 \\ \vdots \\ S_N \end{pmatrix}}_{H \circ \vec{v} = \vec{S}} \tag{4-35}$$

Zur Lösung des Gleichungssystems muss folgende Bedingung gelten: $\text{rang}(H) = \text{rang}(H,S)$, sodass die ausreichend linearunabhängigen Gleichungen zur Verfügung stehen. Damit zeigt sich, dass sich die hinteren Gewichte $v_1,....v_m$ durch Lösen des linearen Gleichungssystems bestimmen lassen, falls die Eingangsgewichte ermittelt wurden.

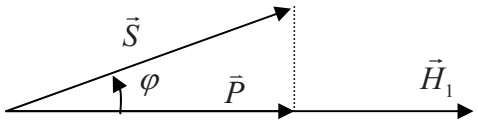
Im Gegensatz zu anderen Lernalgorithmen basiert dieser Algorithmus nicht auf der Minimierung des zu erwartenden Fehlers, sondern auf der Lösbarkeit des Gleichungssystems. Ist $n << N$, wie es in der Praxis der Fall ist, müssen die Gewichte w_{kj} nun so gewählt werden, dass die Lösbarkeitsbedingung möglich ist. Hierzu wird versucht, dass die Spaltenvektoren $\vec{H}_1$ optimal in Richtung des Hilfsvektors $\vec{S}$ auszurichten. Im ersten Schritt wird der erste Spaltenvektor $\vec{H}_1$ so ausgerichtet, dass dieser möglichst parallel zum Vektor $\vec{S}$ ist:

$$\vec{H}_1 = \begin{pmatrix} h_{11} \\ h_{21} \\ \vdots \\ h_{N1} \end{pmatrix} \quad \| \quad \vec{S} \tag{4-36}$$

Das heißt, der Winkel zwischen den Vektoren $\vec{S}$ und $\vec{H}_1$ sollte möglichst klein sein. Die Gewichte $w_{11},....w_{1n}$ sind demzufolge so zu dimensionieren, dass der Winkel zwischen $\vec{H}_1$ und $\vec{S}$ minimal wird, was bedeutet, dass der $\cos(\angle)$ und auch $\cos^2(\angle)$ maximal zu wählen ist, woraus folgt, dass:

$$\frac{\left(\vec{H}_1 \cdot \vec{S}\right)^2}{\left|\vec{H}_1\right|^2 \left|\vec{S}\right|^2} \to \max \text{ und } \frac{\left(\vec{H}_1 \cdot \vec{S}\right)^2}{\left|\vec{H}_1\right|^2} \to \max \tag{4-37}$$

Mit Hilfe eines Gradientenverfahrens werden zunächst die Gewichte w_{1k} so eingestellt, dass die Gleichung (4-35) maximale Werte annimmt. Alle weiteren Gewichte werden auf Null gesetzt. Das Gewicht $\vec{v}_1$ wird so gewählt, dass $\vec{v}_1 \cdot \vec{H}_1$ dem Projektionsvektor $\vec{P}$ von $\vec{S}$ in Richtung $\vec{H}_1$ entspricht.



Für die Länge des Projektionsvektors $\vec{P}$ gilt dann:

$$P = \left|\vec{S}\right| \cdot \cos(\varphi) \cdot \frac{\vec{H}_1}{\left|\vec{H}_1\right|} = \left|\vec{S}\right| \cdot \frac{\left(\vec{S} \cdot \vec{H}_1\right)}{\left|\vec{S}\right| \cdot \left|\vec{H}_1\right|} \cdot \frac{\vec{H}_1}{\left|\vec{H}_1\right|} = \frac{\left(\vec{S} \cdot \vec{H}_1\right)}{\left|\vec{H}_1\right|^2} \cdot \vec{H}_1 = v_1 \cdot \vec{H}_1 . \qquad (4\text{-}38)$$

Also kann für die v_j folgende Formel angegeben werden:

$$\vec{v}_j = \frac{\left(\vec{S} \cdot \vec{H}_j\right)}{\left|\vec{H}_j\right|^2} . \qquad (4\text{-}39)$$

Dieser Vorgang wird für jeden Eingang wiederholt. Nach jedem gesetzten Gewicht passt sich der Gesamtvektor der $\vec{H}_j, j = 1,...,m$ dem Ergebnisvektor $\vec{S}$ besser an. Dieser Vorgang wird so oft wiederholt, bis der Fehler klein genug ist. Das Verhalten lässt sich auf mehrere Ausgangsgrößen verallgemeinern. Die Anzahl der Ausgänge entspricht dann der Anzahl der Spalten von $\vec{w}$ und $\vec{S}$. So gilt in allgemeiner Form:

$$\vec{S}_J = \vec{S} - \sum_{k=1}^{j-1} v_k \vec{H}_k \qquad (4\text{-}40)$$

Das mathematische Verfahren hierzu ist ausführlich in [Ger08] beschrieben. Der Algorithmus bietet den Vorteil gegenüber dem verbreiteten Backpropagation-Algorithmus, dass er stabil und schneller abläuft und keine lokalen Minima angestrebt. Nachteilig bei diesem Algorithmus ist der hohe Speicheraufwand, und es können ausschließlich Feedforward-Netzstrukturen mit ihm trainiert werden.

4.3.5 Backpropagation für rekurrente Netze

Aufgrund der zusätzlich vorhandenen Kontextzellen müssen partiell rekurrente Netze mit einem leicht abgewandelten Backpropagation-Lernalgorithmus oder verwandten Algorithmen (Super SAB, Quickprop, Rprop) trainiert werden und es werden Verfahren zur Bestimmung der Gewichte benötigt. Nach Minsky und Papert [Min88] gibt es zu jedem rekurrenten Netz mit diskreten Zeitschritten für eine feste Zeitdauer ein äquivalentes Feedforward-Netz, welches allerdings extrem viele Knoten aufweist. Bei

rekurrenten Netzen werden üblicherweise die Eingabewerte in Folge angelegt. Die Kontextzellen werden lediglich als zusätzliche Eingabezellen betrachtet. Der Eingabevektor besteht dann aus dem normalen Eingangsvektor und dem Zustandsvektor der Kontextzellen. Dieser Zustandsvektor wird aufgrund der festen Übergangsfunktion, mit der er gekoppelt ist, in jedem Schritt neu definiert. Aufgrund des beschriebenen Verfahrens kann folgender modifizierter Backpropagation-Algorithmus angewendet werden [Zel94]:

1. Initialisierung der Kontextzellen.
2. Anlegen des Eingangsmusters und Vorwärtspropagierung bis zur Ausgabe des Netzes (rekurrente Zellen werden hierbei nicht beachtet).
3. Ermittlung des Fehlersignales durch Vergleich der Eingabe- mit den Ausgabewerten.
4. Rückwärtspropagierung der Fehlersignale (von den Ausgabezellen zu den Eingabezellen).
5. Berechnung der Gewichtsänderung und deren Adaption.
6. Ermittlung des Folgezustandes aller Kontextzellen gemäß Ihrer Eingabeverbindung.

Anhand einer einfachen Netztopologie (Abbildung 4-12) soll verdeutlicht dargestellt werden, wie dieser Algorithmus arbeitet.

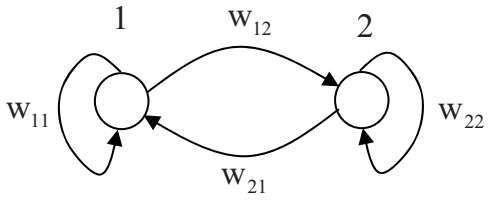


Abbildung 4-12: Rekurrentes Netz mit zwei Neuronen

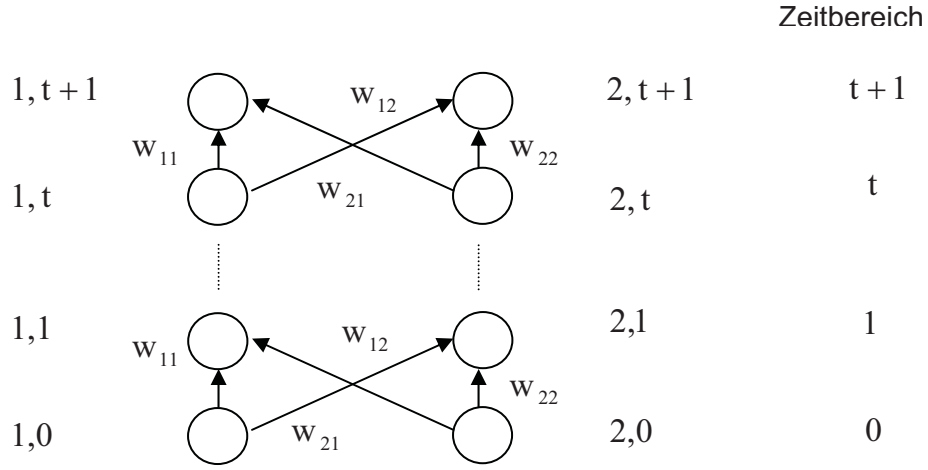


Abbildung 4-13: Äquivalentes Feedforward-Netz

Durch eine zeitliche Entfaltung werden für ein äquivalentes Feedforward-Netzwerk alle Neuronen und Gewichte des rekurrenten Systems für jeden diskreten Zeitschritt durch eigene Neuronen und Gewichte ersetzt. Dazu muss die maximale Zahl der Zeitschritte,

die man untersuchen will, vorher bekannt sein. Die Verbindungsgewichte w_{ij} werden dann immer auf allen Ebenen gleich den Gewichten des Intervalls $t+1$ gesetzt (Abbildung 4-13). Diese Gewichte müssen auf allen Ebenen gleich sein und den Gewichten des rekurrenten Systems entsprechen [Zell94]. Die Propagierungsregel für dieses Netzwerk lautet:

$$a_j(t+1) = f\left(net_j(t)\right) = f\left(\sum_i a_i(t)\, w_{ij} + in_j(t)\right) \tag{4-41}$$

Wobei $a_i(t)$ der Ausgabe des Neurons j zum Zeitpunkt t entspricht und $in_j(t)$ hier die Eingabe des Neuron j zum selben Zeitpunkt darstellt. Existiert noch keine Eingabe $in_j(t)$, so wird diese auf 0 gesetzt. Für die Fehlerfunktion aller Ausgabeneuronen k über die gesamte Zeit $t+1$ bis t_{max} gilt:

$$E = \sum_{t=1}^{t_{max}} E(t) = \sum_{t=1}^{t_{max}} \frac{1}{2} \sum_k (E_k(t))^2 \tag{4-42}$$

mit:

$$E_k(t) = \begin{cases} t_k(t) - a_k(t); & \text{falls für k zum Zeitpunkt t eine Ausgabe spezifiziert wurde} \\ 0; & \text{sonst} \end{cases}$$

Der Übersichtlichkeit halber wurde auf die Summation über die Muster p verzichtet. Sie wird an den notwendigen Stellen implizit angenommen. Es gilt weiterhin:

$$\Delta w_{ij} = -\eta \frac{\partial E_p}{\partial w_{ij}} \tag{4-43}$$

Das Fehlersignal wird definiert als:

$$\delta_j(t) = -\frac{\partial E}{\partial net_j(t)} \tag{4-44}$$

Damit ergibt sich:

$$\delta_j(t) = \begin{cases} f'_j(net_j(t)) E_j(t); & \text{falls } t = t_{max} \\ f'_j(net_j(t)) \left[E_j(t) + \sum_k \delta_k(t+1) w_{jk} \right]; & \text{falls } 1 \le t < t_{max} \end{cases} \tag{4-45}$$

Durch Einsetzen in Gleichung 4-41 ergibt sich für die Gewichtsänderung eines Musters über den Zeitraum $t+1$ bis t_{max}:

$$\Delta w_{ij} = -\eta \frac{\partial E_p}{\partial w_{ij}} = -\eta \sum_{t=1}^{t_{max}} \delta_j(t)\, a_j(t-1) \tag{4-46}$$

Der Algorithmus ist ausführlich in [Zel94] beschrieben.

4.4 Selbstorganisierende Karten (SOM)

Aufgrund der Weiterentwicklung der Speichertechnologie und der hieraus resultierenden Kostenreduzierung von Speichermedien haben sich in den vergangenen Jahren die Möglichkeiten Daten zu erheben, zu generieren und auch zu speichern entscheidend verbessert. Einhergehend mit dieser Tendenz wird die optimale Ausnutzung und Analyse vorhandener Datenbestände als potentielle Quelle wertvoller Informationen zur Schaffung von Wettbewerbsvorteilen im Markt unumgänglich sein. Unter dem Begriff des Knowledge Discovery in Databases (KDD) und Data Mining (DM) ist in den vergangenen Jahren eine neue Forschungsrichtung entstanden, innerhalb derer neuronale Netze und hier speziell die Gruppe selbstorganisierender Karten Anwendung findet. Hierbei wird versucht, Verfahren und Methoden zur Analyse umfangreicher Datenbestände zu entwickeln.

Die bisher vorgestellten Typen neuronaler Netze und zugehöriger Lernverfahren befassten sich ausschließlich mit überwachten Lernvorgängen. Die Gruppe der Self-Organizing-Feature-Maps (SOM) oder auch nach ihrem Entwickler Teuvo Kohonen benannten Kohonen-Netze wurden in den 80er Jahren auf Basis der Überlegungen von Rosenblatt und Malsberg entwickelt [Ros58] [Mal73]. Dabei beruht der Hauptunterschied gegenüber konventionellen Netzen darin, dass bei selbstorganisierenden Karten keine explizite Ausgabe vorgeschrieben ist und daher keine Fehlerfunktion für das Netz definiert werden kann. Sie kommen wahrscheinlich in ihrem Verhalten der Arbeitsweise des menschlichen Gehirns am nächsten, wenn man wie Kohonen davon ausgeht, dass im menschlichen Gehirn bei Aktivierung nicht nur ein Neuron reagiert, sondern ein ganzer nachbarschaftlicher Bereich konnektivistisch involviert ist.

Wie im menschlichen Gehirn für gewisse Muster immer dieselben Hirnregionen angeregt werden, sollen bei den selbstorganisierenden Karten gleichartige Muster auch in denselben Bereichen wiederzufinden sein. Diese kollektive Zusammenarbeit der Neuronen basiert auf gegenseitiger Beeinflussung, wobei auf ein Eingabemuster immer ein ganzer Bereich von benachbarten Neuronen reagiert. Damit sind diese Netze in der Lage, selbständig aus einer Datenmenge klassenzusammengehörige Strukturen von topologieerhaltenden Abbildungen hochdimensionaler Merkmalsräume in einem Ausgaberaum niedriger Dimension zu detektieren. Im Gegensatz zu den bisher vorgestellten Netztypen sind diese Karten damit in der Lage, sich durch die bloße Präsentation von Eingabedaten selbst in einer sinnvollen Weise zu organisieren und unbekannte Strukturen in einer zu analysierenden umfangreichen Datenbasis (ohne a-priori-Informationen) „selbstständig“ zu extrahieren und diese grafisch darzustellen.

Die Anwendung von SOM zur explorativen Datenanalyse, die sich in die Kategorien Clusteranalyse und Projektionsmethoden unterteilen lässt, stellt somit grundlegende Werkzeuge für Data Mining (DM) und Knowledge Discovery in Databases (KDD) dar, und erscheint in diesem Kontext als ein sehr interessantes Verfahren.

Um die Technik der Kohonen-Netze einfacher nachvollziehen zu können, wird zunächst die Netzarchitektur dargestellt und auf das Verfahren der lernenden Vektorquantisierung (LVQ) eingegangen, da sich Kohonen-Netze dann als einfache Erweiterung von LVQ verstehen lassen.

4.4.1 Prinzip und Netzarchitektur der SOM

Eine SOM ist ein äußerst übersichtliches neuronales Netz, das im Regelfall aus einem zweischichtigen„feed forward“-Netzwerk mit einer Input- und einer Output-Schicht besteht. Abbildung 4-14 zeigt exemplarisch die Architektur einer eindimensionalen Kette und einer zweidimensionalen Karte.

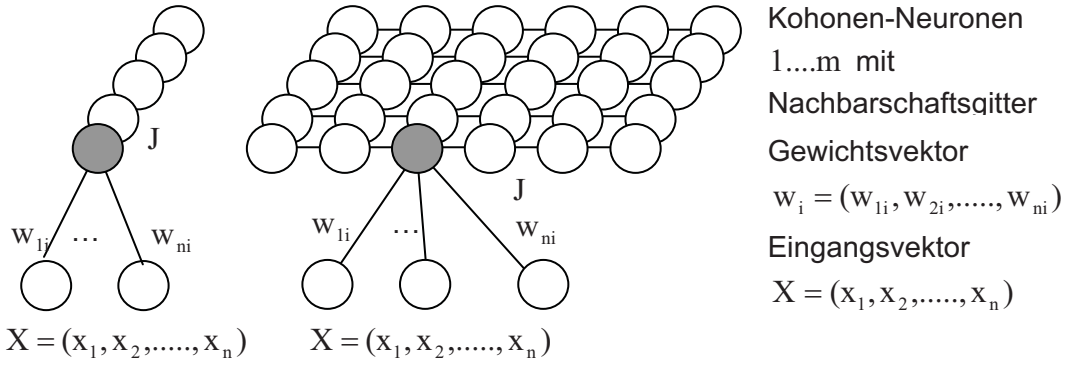


Abbildung 4-14: Netzstruktur einer Selbstorganisierenden Kette und Karte

In der praktischen Anwendung werden überwiegend zweidimensionale Gitterstrukturen, sogenannte „Karten“ eingesetzt, da diese gute Visualisierungsmöglichkeiten bieten [Pol05]. Die Inputschicht, deren Anzahl von Neuronen davon abhängig ist, welche Dimension der zu bestimmende Merkmalsraum aufweist, dient dazu, externe Informationen dem Netz zuzuführen. Dabei ist jedes Inputneuron vollständig mit den Neuronen innerhalb der Karte über Verbindungsgewichte verknüpft. Die Besonderheit der SOM-Architektur besteht darin, dass die Kohonenschicht zwischen den einzelnen Neuronen eine fest definierte Ordnungsrelation, z. B. eine Vorgänger/Nachfolger-Relation aufweist, und somit eine „räumliche“ Nachbarschaftsbeziehung zwischen diesen besteht.

Diese Beziehung sorgt dafür, dass eine „Metrik“ auf der Ausgabeschicht bestimmt werden kann, d. h. der Abstand oder die Entfernung zwischen zwei Neuronen kann ermittelt oder gemessen werden. Die fest definierte Nachbarschaftsbeziehung der Outputschicht ist eine Voraussetzung für den Topologieerhalt bei der Dimensionsreduzierung der Inputschicht. Durch die Nachbarschaftsbeziehung erfolgt bei den selbstorganisierenden Karten eine spezielle Art der Vektorquantisierung, bei der die Gewichtsvektoren der SOM im Rahmen des Trainings so angepasst werden, dass sie

den Eingabevektor räumlich auf der Karte repräsentieren. Durch die Definition der lokalen Nachbarschaft werden dann nicht nur die Gewichtsvektoren eines Neurons, sondern auch die umliegenden Vektoren beeinflusst. Das heißt, je geringer ein Neuron von der Stelle der Änderung entfernt ist, desto stärker wird dessen Gewichtsvektor beeinflusst, und je weiter dieses räumlich entfernt ist, desto geringer ist die Beeinflussung.

Das Lernverfahren selbstorganisierender Karten gehört zur Gruppe der überwachten Lernverfahren. Dem Netz werden ausschließlich Eingabevektoren zur Verfügung gestellt, zu denen keine zugehörigen Ausgabemuster existieren. Es kann daher genutzt werden, um Cluster innerhalb der Eingabevektoren zu detektieren und nach Abschluss der Trainingsphase unbekannte Eingabemuster einer Klasse zuzuordnen. Während der Trainingsphase werden die n-dimensionalen Eingabevektoren X sequentiell an die SOM-Inputschicht parallel angelegt und mit den Gewichtsvektoren $W_i = (w_{1i}, w_{2i}, ..., w_{ni})$ verglichen. Das Neuron, dessen Gewichtsvektor der Eingabe im Sinne einer der nachfolgend beschriebenen Methode am ähnlichsten ist, wird als Gewinner deklariert.

$$\left\|X - W_i\right\| = \min_{j=1..N}\left\|X - W_i\right\| \tag{4-47}$$

Dabei erfolgt ein Vergleich über das Distanzmaß zwischen den Gewichtsvektoren W_i der Kartenneuronen und den Eingangsvektoren X. Die absolute Distanz ist eines der schnellsten Verfahren [Sch91], bei dem zur Berechnung die Beträge der Differenzen der Elemente vom Eingangsvektor zum Gewichtsvektor des Kartenneurons summiert werden. Eines der am häufigsten verwendeten Verfahren ist das Skalarprodukt. Nach vorheriger Normierung des Eingangsvektors unterscheidet sich dann der Eingangs- vom Ausgangsvektor durch seine Richtung bzw. durch den Winkel zueinander. Das Neuron mit der maximalen Aktivierung, das größte Skalarprodukt, ist dann das Gewinnerneuron. Bei diesem Verfahren ist darauf zu achten, dass nach jedem Anpassungsvorgang der Gewichtsvektoren in der Lernphase diese wieder normiert werden müssen, weil durch die Anpassung die Länge nicht mehr eins, sondern geringfügig kleiner geworden ist. Ein weiteres Verfahren, welches ebenfalls sehr häufig Verwendung zur Ermittlung des Ähnlichkeitsmaßes findet, ist die euklidische Distanz. Bei diesem Verfahren ist eine explizite Normierung nicht nötig, da sie im Verfahren schon integriert ist. Hierbei wird als Gewinner das jeweilige Neuron mit der minimalsten euklidischen Distanz zwischen seinem Gewichtsvektor und dem Eingangsvektor gewählt [Kuh04].

Wenn zu einem Eingabevektor das Gewinnerneuron ermittelt wurde, kann mit der Adaption der Gewichtsvektoren auf der Kartenschicht fortgefahren werden. Das heißt, während der Lernphase wird nicht nur der Gewichtsvektor des Gewinnerneurons verändert, sondern gleichzeitig erfahren Neuronen, die sich innerhalb einer topologischen Nachbarschaft im Neuronengitter zum Gewinner befinden, ebenfalls eine Gewichtsanpassung. Dabei kann die Anpassung dieser Gewichte mit zunehmenden Iterations-

schritten in Form und Größe variieren. In der Regel beginnt man mit einem großen Nachbarschaftsradius, der mit fortschreitender Iteration abnimmt. Formal kann diese Vorgehensweise wie folgt beschrieben werden:

$$W_j(t+1) = W_j(t) + \eta(t) \cdot h_{n,j}(t) \cdot \left(X - W_j(t)\right) \tag{4-48}$$

Hierbei ist $h_{n,j}(t)$ die Nachbarschaftsfunktion, mit deren Hilfe die Anpassung der Neuronen erfolgt, und $\eta(t)$ die Lernrate, die üblicherweise mit zunehmender Iterationsschrittanzahl abnimmt. Die Nachbarschaftsfunktion, mit der die Umgebung des Gewinnerneurons beeinflusst wird, und η_t als Faktor, über den mit Zunahme der Iterationsschritte die Beeinflussung auf die Nachbarschaftsneuronen gesteuert wird, müssen algorithmisch beschrieben werden. Dabei bieten sich verschiedene alternative Funktionen an [Zel94]. Häufig verwendete Funktionen für die Beeinflussung der Nachbarschaftsneuronen sind zum Beispiel die statistischen Funktionen wie die bubble-Funktion oder die Gaußsche Glockenkurve, Zylinder- oder Kosinusfunktionen, die im Folgenden kurz vorgestellt werden. Bei der Zylinderfunktion ist der Wert der Beeinflussung abhängig vom Abstand $z = \left\| r_n - r_j \right\|$, der Lagevektoren r_n und r_j, der Neuronen n und j und in der Regel von der Iterationenzahl t.

$$h_{cylinder}(z,d) = \begin{cases} 1 & \text{falls } z < d \\ 0 & \text{sonst} \end{cases} \tag{4-49}$$

Die Zylinderfunktion ist konstant Eins für eine Nachbarschaft von Punkten um das Gewinnerneuron n mit dem Abstand d und sonst Null. Die normalisierte Gaußfunktion wird wie folgt beschrieben:

$$h_{n,j\,gauss}(t) = \frac{1}{\sigma(t)\sqrt{2\pi}} \exp\left(- \frac{\left\| r_n - r_j \right\|^2}{2\sigma^2(t)} \right) \tag{4-50}$$

Die Varianz $\sigma(t)$ gibt hier die Breite des Nachbarschaftskerns an, welche anfangs auf einen großen Wert gesetzt wird, und mit wachsender Iterationszahl abnimmt. Unter Verwendung eines Parameters d, der statt des Zeitparameters t die Größe des Nachbarschaftsradius sowie alle Konstanten enthält, kann diese Funktion auch folgendermaßen beschrieben werden (z sei die wie oben definierte Abstandsgröße $\left\| r_n - r_j \right\|$):

$$h_{gauss}(z,d) = e^{-(z/d)^2} \tag{4-51}$$

Die Kosinus-Funktion:

$$h_{cos}(z,d) = \begin{cases} \cos\left(\frac{z}{d} \frac{\pi}{2} \right) & \text{falls } z < d \\ 0 \end{cases} \tag{4-52}$$

Die Kegelfunktion:

$$h_{cone}(z,d) = \begin{cases} 1 - \frac{z}{d} & \text{falls } z < d \\ 0 \end{cases} \qquad (4\text{-}53)$$

Abbildung 4-15 zeigt den Funktionsverlauf dieser Distanzfunktionen. In der Literatur werden jedoch auch Abstandsfunktionen erwähnt, bei denen es auch zu Abstoßungsreaktionen auf weit entfernte Neuronen kommt. Ein Beispiel für dieses Verhalten ist die so genannte Mexican-Hat-Funktion. Da beim Einsatz dieser Funktion oft ein Divergieren des Netzes auftrat, wird in der Praxis darauf verzichtet. Nach Abschluss der Lernphase kann nun die geordnete Kohonenschicht analysiert werden. Da der Schritt der SOM-Datenclusterung abhängig von der Art und Dimension des jeweiligen Eingangsvektors und der Kohonenkarte ist, lässt sich leider keine verallgemeinerungsfähige Methodik beschreiben.

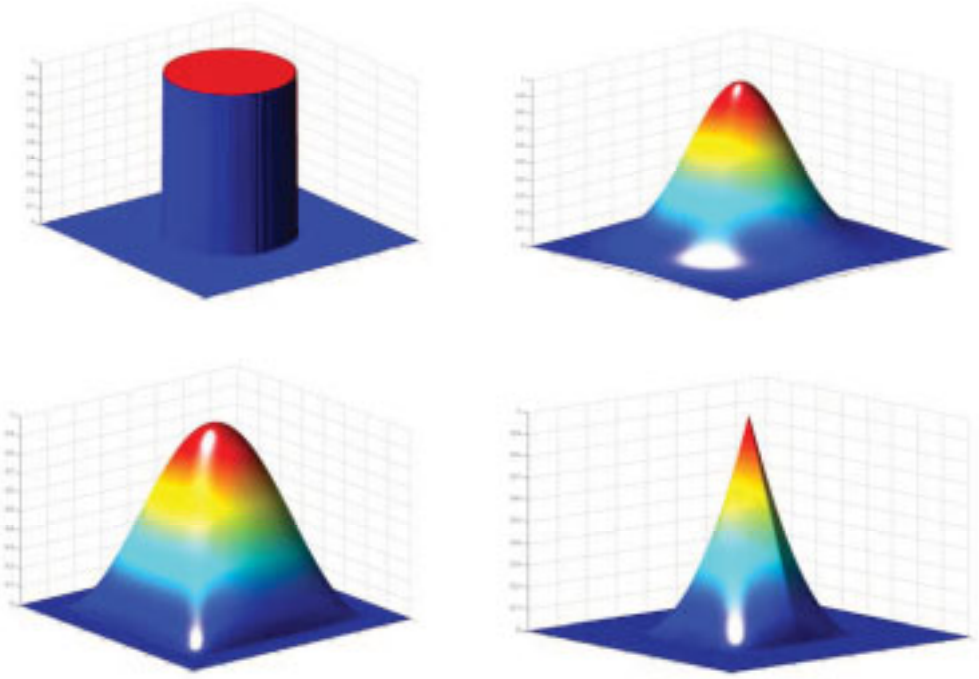


Abbildung 4-15: Distanzfunktionen ($h_{cylinder}$,h_{gauss}, h_{cos}, h_{cone})

In der Regel bietet die Kartenschicht keine Möglichkeit, erkennbare Datenstrukturen wahrzunehmen. Hierzu ist es notwendig, ein „Labeling“ der Karte vorzunehmen. Dabei werden die zuvor als Trainingsdaten verwendeten Vektoren der Karte erneut repräsentiert und die Merkmale des Eingabedatensatzes dem Gewinnerneuron zugeordnet und gekennzeichnet.

Auf Basis dieser Ergebnisse können weitere Grafen zur Visualisierung der Datenclusterung erstellt werden, auf denen beispielsweise die Entfernung (gemessene Euklidische Distanz) der Kohonenkarte oder die Anzahl der Treffer eines Gewinner-

neurons zu einem Eingangsdatensatz dargestellt werden. Die hiermit erstellten Grafen ermöglichen Strukturen innerhalb der Karten sichtbar und interpretierbar zu machen [Koh91]. Hierbei können eine Vielzahl von verschiedenen Visualisierungsverfahren und Techniken Anwendung finden. Diese werden ausführlich in [Kas98], [Kra95] und [Ult93] beschrieben.

5 Entwicklung von Lastgangprognosemodellen

Im folgenden Kapitel wird die Entwicklungsmethodik von Lastprognosemodellen erarbeitet und vorgestellt. Dabei erfolgt zunächst eine Modellierung mit Hilfe eines Zerlegungsansatzes, wobei die verbleibende Restkomponente als stochastischer Prozess aufgefasst wird. Des Weiteren werden Prognosemodelle auf Basis neuronaler Netze unterschiedlicher Netztypen und deren systematische Einbindung in Prognosesysteme entwickelt. Zusätzlich wird ein entkoppeltes Modell vorgestellt, bei dem eine Einbindung exogener Größen vorgenommen wurde. Abschließend wird mit Hilfe selbstorganisierender Karten (SOM - Self-Organizing Maps) ein Verfahren entwickelt, aus Lastdatensätzen vergangener Jahre Unregelmäßigkeiten im Lastgang selbst zu detektieren und Datenclusterungen innerhalb der Lastgangdaten zu bilden, um so eine Reduktion diskontinuierlicher Werte (Merkmale eines Tages) für die Prognoseerstellung zu erreichen.

5.1 Modellbildung und Anpassung stochastischer Prozesse

Verbesserungen von Prognoseergebnissen können erreicht werden, wenn eine Zerlegung in deterministische Anteile und Rauschen berücksichtigt wird und die Modellparameter mit feinsinnig ausgewählten Daten angepasst werden. Die optimale Parametrisierung dieser kombinierten Modelle in Abhängigkeit des betrachteten Lastkollektivs ist eine zwingende Voraussetzung, die zwar mit hohem manuellem Aufwand verbunden ist, aber für jedes Kollektiv nur einmal zu erfolgen hat. Ein Teil der Parameter muss für jede neue Prognosefindung eigens berechnet werden. Im Folgenden werden der Zerlegungsansatz, die Parametrisierung des Modells sowie die Ergebnisverbesserung beschrieben. Dem Modell zur Behandlung der Rauschanteile liegt ein zunächst univariater Prozess zu Grunde. Weitere Untersuchungen zeigen, dass die zusätzliche Verwendung exogener Größen wie Wetterdaten erfolgen kann.

5.1.1 Zerlegungsansatz und Modelle der Restkomponente

Der Lastverbrauch stellt eine typische Zeitreihe dar. Abbildung 3-1 zeigt typische Verläufe und lässt Tages- und Wochenregelmäßigkeiten erkennen. Den weiteren hier folgenden Betrachtungen liegt die Annahme zugrunde, dass aus dem bisherigen und damit bekannten Verlauf einer solchen Zeitreihe auf deren zukünftige Entwicklung geschlossen werden kann und somit eine Vorhersage möglich ist. Die Erstellung eines geeigneten einschlägigen Prognosemodells beruht wesentlich auf Methoden der Zeitreihenanalyse und im Speziellen auf der Theorie stochastischer Prozesse. Die klassische Zeitreihenanalyse betrachtet Zeitreihen als Überlagerung dreier Komponenten (vgl. 3-21). Die Modellierung des Trendanteils $g(t)$ kann mit ver-

schiedenen Ansätzen erfolgen. Neben exponentiellen Modellen wird hier ein polynomiales Modell in der Form:

$$g(t) = \sum_{i=0}^{p} a_i \cdot t^i \text{ mit Polynomgrad } p \text{ und } a_i \in IR \text{ angewendet.} \quad (5\text{-}1)$$

Die Bestimmung der Koeffizienten a_i erfolgt mit der „Methode der kleinsten Quadrate“ nach Gauß [Sch04], [Eld90], der Polynomgrad p wird möglichst gering gehalten. Üblicherweise findet ein Grad <4 Anwendung. Betrachtet man Lastgänge der Versorgungswirtschaft, so besitzen diese oft ausgeprägte periodische Elemente. Die Modellierung des zyklischen Anteils erfolgt nach Abzug des Trendanteils durch die zyklische Komponente $z(t)$ der Form:

$$z(t) = \sum_{k=1}^{q} \left[a_k \cdot \cos(2\pi\pi_k t) + b_k \cdot \sin(2\pi\pi_k t) \right] \text{ mit } a_k, b_k \in IR\,. \quad (5\text{-}2)$$

Die Auswahl einer Frequenz f_k und der zugehörigen Koeffizienten a_k, b_k erfolgt mit Hilfe des Periodogramms $I(f)$, das auch aus der diskreten Fouriertransformation gewonnen werden kann [Sch89]:

$$I(f) = \frac{1}{N}\left[\left(\sum_{t=1}^{N} (x_t - \bar{x}) \cos(2\pi\pi t) \right)^2 + \left(\sum_{t=1}^{N} (x_t - \bar{x}) \sin(2\pi\pi t) \right)^2 \right] \quad (5\text{-}3)$$

$\bar{x}$ ist der Mittelwert der gemessenen Zeitreihenwerte. Ausgewählt werden diejenigen Frequenzen f, die hervorstechend große Werte $I(f)$ liefern. Abbildung 3-6 zeigt das Periodogramm einer ganzjährigen Zeitreihe eines städtischen Verbrauchskollektivs. Nach Auswahl der Frequenzen werden wiederum mit der Methode der kleinsten Quadrate die zugehörigen Frequenzanteile bestimmt. Mit Hilfe der Trend- und zyklischen Anteile werden Prognosen zu bekannten Messwerten errechnet und beispielweise der mittlere relative Fehler ermittelt. Dabei wird die Zeitreihe in zwei Abschnitte aufgeteilt. Mit den Daten des vorderen Abschnitts werden, wie bereits erläutert, Trend- und zyklischer Anteil ermittelt. Auf deren Basis wird die Prognose für den Zeitbereich des hinteren Abschnitts erstellt. Die Prognosewerte und Messwerte des hinteren Abschnitts führen zum Fehler (MRE). Polynomgrad und Anzahl der größten Frequenzanteile werden so bestimmt, dass dieser Fehler klein wird. Mit Anwendung dieses Verfahrens auf die Lastdaten unterschiedlicher Verbrauchskollektive kann bereits ein mittlerer relativer Fehler kleiner 5 % erreicht werden.

Ein weiterer Verbesserungsschritt zur Minimierung des mittleren relativen Fehlers ist die Gruppierung der Daten in entsprechende Klassen, z. B. Werk-, Wochenend- und Sondertage. Die Art der Gruppierung ist abhängig vom jeweiligen Verbrauchskollektiv. Die Verwendung des Verfahrens auf die zuvor gruppierten Lastdaten erreicht bei der getrennten Prognose von Werktagen und Wochenenden einen Fehler < 4 %. Weitere

Reduzierungen des Prognosefehlers und damit der Restkomponente $r(t)$ sind mit den bisher betrachteten Verfahren nicht möglich.

Um eine zusätzliche Reduzierung des Prognosefehlers zu erreichen, wird nun die Restkomponente $r(t)$ genauer betrachtet und Modelle der Restkomponente generiert. Sie werden als Ausprägung eines linearen stochastischen Prozesses aufgefasst. Einen Überblick über die Systematik zur Modellierung der Restkomponente bietet Abbildung 5-1. Lediglich die Restkomponente wird hierbei nach Abzug des Trends und des zyklischen Verhaltens als stochastischer Prozess beschrieben, wahlweise als AR-, MA- oder als $ARMA$-Prozess. Rechnerische Vorteile bietet ein Prozess mit AR-Anteil, da die Prozessordnung geringer ausfällt und die dazu notwendigen Vergangenheitswerte x_{t-i} vorhanden sind. Exemplarisch wird ein $ARMA$- und ein „entkoppelter“ AR-Ansatz in den Fallstudien vorgestellt.

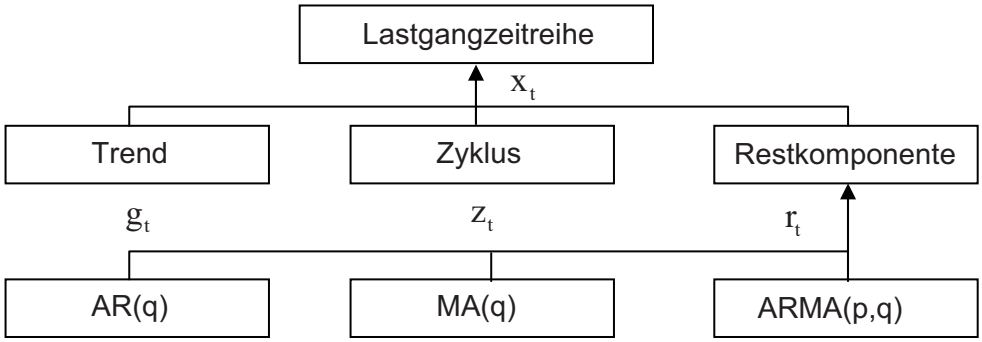


Abbildung 5-1: Modell des Zerlegungsansatzes

Die unbekannten Residuen r_t des Prozesses werden durch die Anpassung eines AR-Prozesses hoher Ordnung an die Zeitreihe $(x_t)_{t\in T}$ zuvor näherungsweise bestimmt. Die Anwendung eines $ARMA$-Modells reduziert den mittleren relativen Fehler der vorliegenden Testdaten des deterministischen Modells von 4 % auf 2,8 %. Nachteilig bei der Anwendung des $ARMA$-Ansatzes für große Parameterordnungen p und q ist das zeitaufwändige Lösen großer Gleichungssysteme [Cor99], [Kar05] sowie das mögliche Auftreten von Instabilitäten bei der Extrapolation der Zeitreihe für die Prognose. Eine weitere Möglichkeit der Modellerstellung bietet der „Entkoppelte“ Autoregressive-Prozess. Untersuchungen zeigen, dass ein Lastwert zu einer bestimmten Tageszeit eine sehr große Ähnlichkeit zum Lastwert des Vortags der entsprechenden Zeit aufweist, was durch das Korrelogramm in Abbildung 5-2 bestätigt wird.

Dieses ermöglichte die Annahme, dass sich ein Lastwert einer bestimmten Tageszeit im Wesentlichen aus den Lastwerten der entsprechenden Tageszeit der Tage zuvor vorhersagen lässt. Aus dieser Erkenntnis kann ein vereinfachtes Prognoseverfahren abgeleitet werden. Die Originalzeitreihe wird dabei in 96 einzelne Tageszeitreihen zerlegt. Die Lastwerte zu den unterschiedlichen Tageszeiten werden hierdurch „ent-

koppelt“ und somit für jeden ¼-h-Wert dieser Zeitreihe einem Prognoseverfahren unterzogen. Beim Trendpolynom beschränkt man sich auf Grad Eins. Der Restanteil wird als $AR(4)$-Prozess aufgefasst. Die Koeffizienten werden nicht durch die Methode der kleinsten Quadrate, sondern durch Lösen der Yule-Walker-Gleichung gemäß Kapitel 9.5 berechnet.

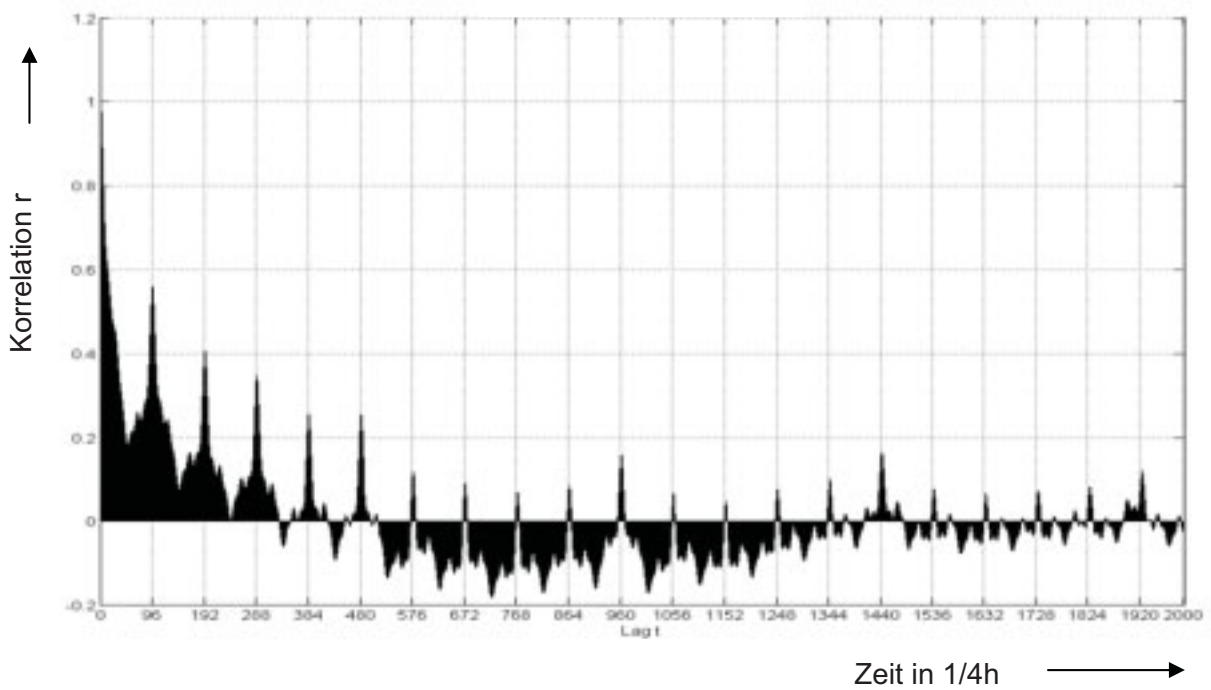


Abbildung 5-2: Autokorrelation der Wochentage eines Jahres

5.1.2 Modellbildung und Bestimmung der Parameter

Die vorgestellten Modelle verfügen über eine nicht unerhebliche Anzahl von Parametern wie z. B. Grad des Trendpolynoms, Anzahl verwendeter Amplituden für die zyklische Komponente, Anzahl der betrachteten historischen Daten usw. Die Bestimmung der Parameter ist der zentrale Schritt in der Konfiguration stochastischer Modelle. Neben mathematischen Standardverfahren, wie „Methode der kleinsten Quadrate“ können analog zum Zeitbereich die Parameter auch im Frequenzraum bestimmt werden [Schl89]. Es stellt sich die Frage, welches der linearen Prozessmodelle optimale Ergebnisse liefert und wie die Wahl der Parameter vorzunehmen ist.

Entsprechend dem Ansatz von Box–Jenkins können durch die Verwendung der Autokorrelation und der partiellen Autokorrelation die Ordnungen p und q des Modells bestimmt werden. Die Verwendung der partiellen Autokorrelation ist jedoch nicht unkritisch, da sie für größere Modelle ungeeignet ist und die resultierende Korrelationsmatrix nicht notwendigerweise positiv definit sind [Spo08]. Die partielle Autokorrelation liefert aber Rückschlüsse über die Ordnung des AR-Anteils bei einem $ARMA/ARIMA$-Prozess, während die Ordnung des MA-Anteils durch die Autokorrelation angedeutet

wird. Diese Herangehensweise liefert letztlich nur Schätzwerte der möglichen Ordnung. Da das Ziel der Parametrisierung eine optimale Prognose ist, soll hier ein wissensbasierter Ansatz zur Bestimmung der Modellordnungen gewählt werden.

Autoregressive Prozesse

Die Autokorrelationsfunktion eines linearen stochastischen stationären Prozesses konvergiert sicher gegen Null für $t \to \infty$. Die Länge des Intervalls, auf der die Autokorrelationsfunktion verschieden von Null ist, hängt von der Ordnung des Prozesses ab. Da im Anwendungsfall die Lastprognose eines Tages aus der Vorhersage von 96-1/4-h-Werten besteht und diese charakteristische Auffälligkeiten aufzeigt, die über das ganze Intervall eines Tages (siehe Abbildung 3-1) verteilt sind, werden kürzere Prozesse nicht in der Lage sein, Prozessmodelle zu liefern, deren Aussagekraft in ausreichender Länge vorhält, um dieses Verhalten nachzubilden. Für Modelle, die gute Prognosen liefern sollen, empfiehlt sich daher, eine Ordnung von mindestens 96 AR-Parametern anzuwenden. Für einen $AR(q)$-Prozess der Form $X_t = \alpha_1(X_{t-1} - \mu) + .. \ .. + \alpha_p(X_{t-p} - \mu)$ müssen die Parameter α, μ und e_t bestimmt werden. Zur Vereinfachung wird e_t auf Null gesetzt, da der Erwartungswert $E(e_t) = 0 \Rightarrow e_t \approx N(0;\sigma^2)$ Null ist. Nach der Bereinigung von Trend und Zyklen ist der Mittelwert in der Regel $\mu = 0$, ansonsten kann er mit $\mu = \bar{x}$ geschätzt werden. Danach bleiben die Parameter $\alpha_1,...,\alpha_p$ übrig, die mit der „Methode der kleinsten Quadrate“ geschätzt werden können. Eine weitere einfache Möglichkeit zur Bestimmung der Autokorrelationsfunktion bietet die Verwendung der Yule-Walker-Gleichung (Kapitel 9.5).

Moving-Average-Prozesse

Während die Bestimmung der Parameter eines AR-Prozesses relativ einfach über die Yule-Walker-Gleichnung oder über eine Regression ermittelt werden konnte, ist die Parametrisierung eines MA-Prozesses deutlich schwieriger. Das erste Problem besteht darin, dass der zu filternde Rauschprozess $e_t \approx N(0;\sigma^2)$ in der Regel unbekannt ist. Über die AR-Darstellung des MA-Prozesses kann das zugehörige Fehlerrauschen bestimmt werden, indem eine $AR(q)$ großer Ordnung definiert und die Residuen des MA-Prozesses berechnet werden. Für kleine Modelle können die Parameter bestimmt werden, indem die empirische der theoretischen Autokorrelation r_τ gleichgesetzt und dann das Gleichungssystem nach den Parametern aufgelöst wird. Eine weitere Möglichkeit, die MA-Parameter zu bestimmen, ist der Einsatz der „Methode der kleinsten Quadrate“ nach Wilson. Dazu wird ein lineares Ausgleichspolynom ($(y = X\beta + e \Rightarrow \beta = (\beta_0, \beta_1)$ verwendet und unter Anwendung der Korrelation werden die Parameter geschätzt [Schl87]. Die Parameterbestimmung von MA- Prozessen wird hier nicht explizit betrachtet, da diese im Rahmen der Arbeit zusammen mit denen der Anpassung von $ARMA$-/ $ARIMA$- Prozessen erfolgt.

ARMA/ARIMA-Prozesse

Die Methoden zur Bestimmung von $ARMA$-Prozessen gelten analog für $ARIMA$-Prozesse, ausgehend von der differenzierten Zeitreihe. Das Verfahren zur Parameterbestimmung wird als „Durbin Methode“ bezeichnet und baut auf dem Ansatz der „Methode der kleinsten Quadrate“ auf. Dabei wird das Modell in folgender Form aufgefasst: $X_t = \alpha_1 \cdot X_{t-1} + . + \alpha_q \cdot X_{t-p} + e_t - \beta_1 \cdot e_{t-1} - . - \beta_q \cdot e_{t-q}$. Die unbekannten Residuen werden durch die Anpassung eines AR-Prozesses großer Ordnung an die Zeitreihe approximiert. Die Vorgehensweise der Parameterbestimmung ist im Anhang 9.5 (Yule-Walker-Gleichung) beschrieben.

Die Bestimmung des Maßes, wie gut der Prozess den tatsächlichen Verlauf abbildet, wird als Güte bezeichnet. Die Güte der Anpassung Q kann nicht mit der Güte der Prognose gleichgesetzt werden. Es kann von einer guten Prognose ausgegangen werden, wenn das Gütekriterium Q möglichst minimal ist. Als Erklärung hierfür wird angegeben, dass Beobachtungen in der Vergangenheit in gewisser Weise künftigen Werten ähnlich sind, ihnen aber nie exakt entsprechen werden; geringe Abweichungen sind daher durchaus zulässig. AR-Prozesse oder MA-Prozesse können über ihre Güte ($Q(p) = \sigma_e^2 / \sigma_x^2$) diagnostiziert werden. Für $ARMA$-/$ARIMA$-Prozesse gibt es mehrere Möglichkeiten, die Auswahl der Ordnungen zu verifizieren [Kar05]:

- Vergleich von Kenngrößen der Reihe mit denen des Prozesses.
- Vergleich der Kenngrößen der Residualreihe mit denen eines White-Noise-Prozesses.
- Vergleich mit übergeordneten Modellen (Overfitting).

5.2 Entwicklung eines entkoppelten Modells mit exogenen Größen

Der Einfluss exogener Größen bei der Erstellung von Lastprognosen kann in Abhängigkeit des Kollektivs zur Verbesserung der Prognosequalität führen. Zur Vorhersage des elektrischen Lastverhaltens wird auch hierbei ein hierarchisch (komponentenbasiertes) strukturiertes Modell verwendet. Analog des zuvor beschriebenen Vorgehens werden dazu die Daten zunächst mit deterministischen Methoden, wie der Fourieranalyse und der Regression, untersucht und abgebildet. Die Differenz der aus dem deterministischen Modell erzeugten Daten und den Originaldaten (dem Residuum) wird als linearer stochastischer Prozess aufgefasst, wobei sich eine modifizierte Form eines AR-Prozesses, der „entkoppelte“ AR-Ansatz als optimal erwiesen hat. Um den Einfluss der exogenen Größen auf den AR-Prozess abzubilden, wird dieser wie folgt exemplarisch für zwei exogene Größen erweitert:

$$x_t = \sum_{i=1}^{p} a_i \cdot x_{t-i} + \sum_{j=0}^{p-1} b_j \cdot y_{t-j} + \sum_{l=0}^{p-1} c_l \cdot v_{t-l} \quad \text{mit } a_i, b_i, c_i \in IR\,, \tag{5-4}$$

wobei y_t und v_t in diesem Fall die exogenen Zeitreihen, Temperatur und Gobalstrahlung sind, es können aber auch beliebige Zeitreihen sein. Ebenfalls ist aus der Formel ersichtlich, dass für die Prognose eines Lastwerts zukünftige exogene Werte verwendet werden müssen. In diesem Fall dient dazu die jeweilige Wetterprognose. Die Koeffizientenmatrix A hat dadurch die Form:

$$A = \begin{pmatrix} x_{t-1} & . & x_{t-p} & y_t & . & y_{t-p-1} & v_t & . & v_{t-p-1} \\ x_{t-2} & . & x_{t-p-1} & y_{t-1} & . & y_{t-p} & v_{t-1} & . & v_{t-p} \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ x_{t-p-q} & . & x_{t-2p-1} & y_{t-p-1} & . & y_{t-2p-2} & v_{t-p-1} & . & v_{t-2p-2} \\ . & . & . & . & . & . & . & . & . \end{pmatrix} \quad (5\text{-}5)$$

wobei $A \circ \vec{x} = d$ mit $\vec{d} = \left[x_{t,} x_{t,-1,}, ..., x_{t,-p}\right]^T$, und $\vec{x} = \left[a_{1,}..a_p, b_0..b_{p-1}, c_0..c_{p-1}\right]^T$ mit $\vec{x} \in IR^{3p}$, $\vec{d} \in IR^n$, $A = IR^{nx3p}$. Die Koeffizienten a_i, b_i und c_i werden durch die „Methode der kleinsten Quadrate“, d. h. durch Lösen der Normalengleichung, bestimmt: $A^T A \circ \vec{x} = A^T \cdot \vec{d}$, mit $\vec{x} = (A^T A)^{-1} A^T \circ \vec{d}$.

Das „Entkoppelte Modell mit exogenen Größen“ (PARX Parallel Auto Regressive eXogenous input) basiert auf der Grundidee des entkoppelten AR-Ansatzes und der Ausnutzung der Periodizität der Uhrzeiten der Tage und den daraus resultierenden Eigenschaften. Bei der Betrachtung der Autokorrelation der elektrischen Last, die in viertelstündlicher Auflösung gemessen wird, ist zu erkennen, dass selbst nach Elimination der Trend- und Saisonkomponente die Zeitpunkte, die äquidistant um 96 Einheiten verschoben sind, eine sehr große Ähnlichkeit zur aktuellen Messung aufweisen.

Wie Abbildung 5-2 zeigt, hat ein Punkt auf der Zeitachse einen sehr großen linearen Zusammenhang mit den Messungen, die zuvor zur selben Uhrzeit, z. B. am Vortag, aufgenommen wurden. Dieses lässt sich als eine Zerlegung der Zeitreihe in 96 einzelne Zeitreihen deuten. Die Gesamtzeitreihe wird durch eine Matrix dargestellt, in der Form, dass die Zeilen den Tagen und jede Spalte einer festen Uhrzeit entsprechen. Für jeden Zeitpunkt kann man somit auf einen Vektor zugreifen, dessen Elemente die Messungen darstellen, die alle zur selben Uhrzeit aufgenommen worden sind. Vorteile dieser Darstellungsform sind erhöhte Verarbeitungsgeschwindigkeit, Genauigkeit und Stabilität bei der Prognose, da anstelle eines großen AR-Prozesses, der bei einer Voraussage von zwei Tagen folglich 192 Werte erzeugen müsste, jetzt jeweils 96 kleine AR-Prozesse nur zwei Werte pro Prozess prognostizieren müssen.

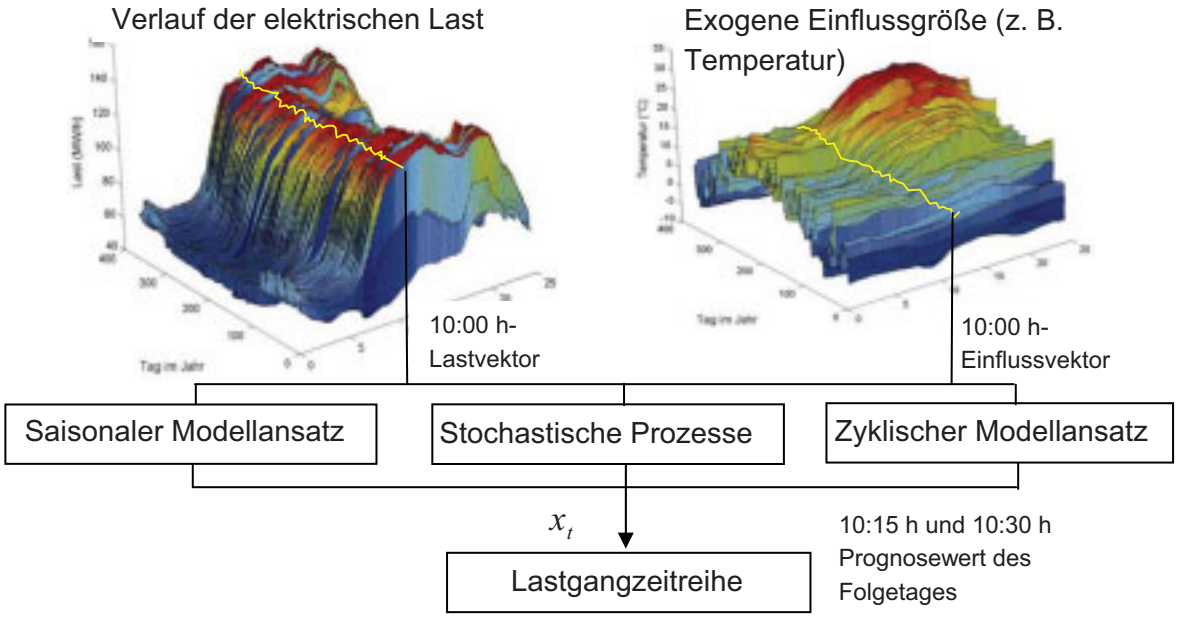


Abbildung 5-3: Struktur des Prognosemodells

Zudem kann der Bereich der verwendeten Daten kleiner als bei konventionellen Prognoseansätzen gehalten werden, da weniger historische Daten pro Prozess benötigt werden, was die Flexibilität des Modells erhöht. Allgemein besitzt der entkoppelte AR-Prozess folgende Struktur:

$$x_t = \sum_{i=1}^{p} a_i \cdot x_{t-i\cdot k} + \sum_{j=0}^{p-1} b_j \cdot y_{t-j\cdot k} + \sum_{l=0}^{p-1} c_l \cdot v_{t-l\cdot k} \qquad (5\text{-}6)$$

$a_i,\ b_i,\ c_i \in \mathrm{IR}$ und $k \in \mathrm{IN}$. Im Fall einer viertelstündlichen Auflösung der Messwerte ist k = 96, für stündliche Auflösung k = 24. Abbildung 5-3 verdeutlicht den Aufbau eines Prognosemodells für entkoppelte AR-Prozesse in Verbindung mit exogenen Größen.

5.3 Prognosen von Zeitreihen mit neuronalen Netzen

5.3.1 Grundlagen bei der Prognoseerstellung mit neuronalen Netzen

Werden die empirischen aufeinander folgenden Variablen einer Zeitreihe $x_1, x_2, ..., x_n$ gemessen, so kann dies als Grundlage für die Prognose zu einem Zeitpunkt $T+1$ Anwendung finden. Besteht zwischen den Größen ein funktionaler Zusammenhang, so lässt sich dieser mit Hilfe eines Backpropagation-Netzes beschreiben [Roj93]. Bei der Prognoseerstellung wird ein Abschnitt der Zeitreihe $x_{t_1-n}, x_{t_2-n}, ..., x_{t-n}$ aus einem zusammenhängenden Segment von gemessenen Variablen dem Netz zu Trainings-

zwecken angeboten und eine optimale Anpassung des Trainingsausgangswertes x_t während der Lernphase vorgenommen. Das Training wird dann für alle Kombinationen von n aufeinander folgenden Segmenten der Zeitreihe für Eingangs- und Ausgangszeitpunkte wiederholt. Nach Abschluss der Lernphase können dann die zuletzt gemessenen und bekannten Werte dem Netz präsentiert werden und eine Prognose für den Zeitpunkt $T+1$ erfolgen. Die hier geschilderte Vorgehensweise entspricht dem Verfahren des Autoregressionsmodelles (Gleichung 3-43).

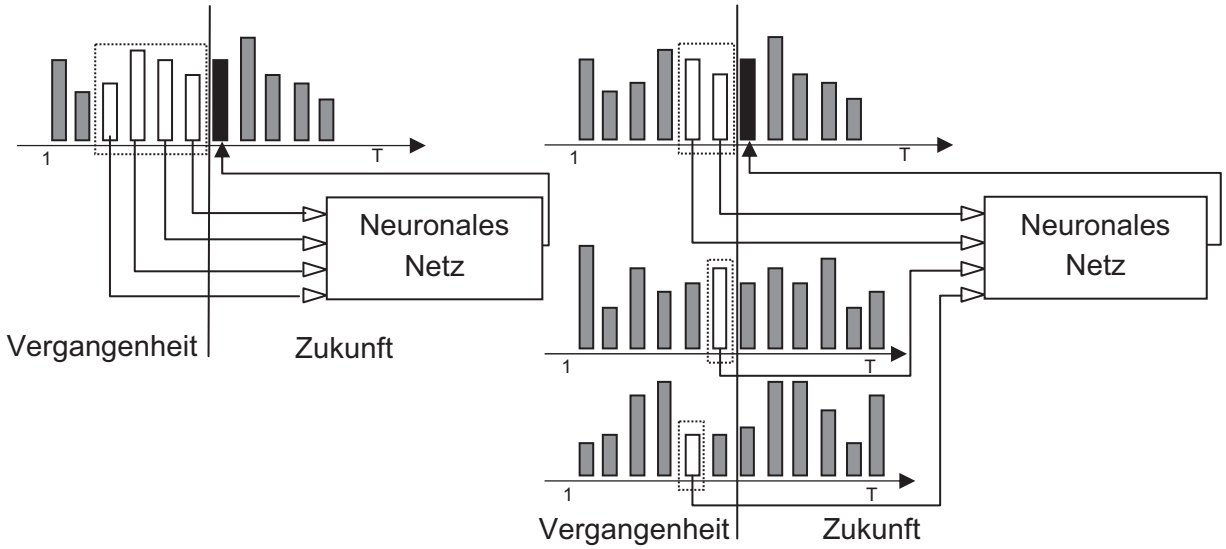


Abbildung 5-4: Lastprognose mit neuronalen Netzen

Viele dieser Modelle benutzen eine lineare Approximation und unterscheiden sich nur in der Anzahl und der Verwendung stochastischer Variablen und ihrer funktionalen Abhängigkeiten. Rein theoretisch sollten Backpropagation-Netze bessere Prognoseresultate erzielen können, da sie auch in der Lage sind, nichtlineare Zusammenhänge zu modellieren [Gro93]. Der grundsätzliche Aufbau der Eingangsgrößen einer Netzstruktur kann dabei unterschiedlich gestaltet werden. Im Vergleich zu univariaten, autoregressiven oder auch horizontalen Verfahren wird im Fall mehrerer Einflussgrößen eine vertikale Prognose angewendet, motiviert durch die lineare Regression. Bei der Lastprognose hingegen wird im Regelfall eine Mischform aus horizontaler und vertikaler Prognose angewendet, wie in Abbildung 5-4 dargestellt ist [Thi04].

Hierbei fließen neben den autoregressiven Lastgangdaten selbst auch exogene Einflussgrößen ein, die in unterschiedlicher zeitlicher Tiefe in dem Prognosemodell Berücksichtigung finden. Die Auswahl der Eingangsgrößen muss vor der Prognose bekannt sein und ist in Abhängigkeit der jeweiligen Aufgabenstellung anzupassen. In der Praxis jedoch ist die Erstellung von Prognosen ein komplexer Vorgang, der ein systematisches Vorgehen erfordert und der im Folgenden näher beschrieben wird. Die allgemeine Bezeichnung „Prognosesystem“ bezieht sich auf die ganzheitliche Erfassung der Prognoseproblematik. Ein Prognosesystem integriert dabei drei wesent-

liche Phasen, die sich in Datenaufbereitung, Lern- und Testphase und die Prognosephase bis hin zur Aktualisierung des Modells einteilen lassen.

5.3.2 Problemstellung und Entwicklungsmethodik

Obwohl die Prognose mit künstlichen neuronalen Netzen und anderen universellen nichtlinearen Funktionsapproximationen, deren genauere Approximation im Vergleich mit den üblichen Regressionsverfahren einen breiten Einsatz in der Praxis und gute Ergebnisse erwarten lassen, befinden sich derartige Prognosesysteme bisher nur in eingeschränktem Maße in der industriellen Anwendung im Einsatz [Huf99]. Probleme im Zusammenhang mit zur Lastprognose eingesetzten künstlichen neuronalen Netzen ergeben sich durch offene Fragen insbesondere in den folgenden Bereichen [Huf97]:

Systemarchitektur
- Anzahl der verwendeten Netze.
- Verknüpfung dieser Netze.

Datenauswahl/-aufbereitung
- Identifikation geeigneter Einflussgrößen.
- Normierung/Codierung der Ein- und Ausgangsgrößen.
- Auswahl der Lerndaten, d. h. der Vergangenheitsdaten, die zur Modellanpassung verwendet werden.

Netzstruktur
- Anzahl Neuronen und Verbindungen.
- Anzahl verdeckter Schichten.
- Verknüpfung der Neuronen.

Modellanpassung
- Wahl eines geeigneten Lernverfahrens.
- Übertrainieren, d. h. das Rauschen in den Lerndaten wird angepasst, die Generalisierung des zugrunde liegenden funktionalen Zusammenhangs verschlechtert sich.

Schlechte Modellanpassung durch Abbruch des Lernverfahrens in lokalem Minimum der Fehlerfunktion.

In der Literatur sind verschiedene Ansätze zu finden, die ein systematisches Vorgehen beschreiben. Es existiert jedoch bis heute keine einheitliche Methode, die unabhängig von der Problemstellung anwendbar ist [Zha03]. Auch ein allgemein gültiges Vorgehen, nach dem aus speziellen Situationen eine Handlungsanweisung für die Modellierung eines optimalen Prognose-Modells abgeleitet werden könnte, hat sich bis heute nicht durchsetzen können. Stattdessen werden die einzelnen Methoden oft nahezu willkürlich ausgewählt und die Parameter der jeweiligen Methoden in der Regel jedes Mal neu durch einen mehr oder weniger aufwändigen „Trial and Error“-Prozess bestimmt. Der im Rahmen der Arbeit verwendete Ansatz ist in Abbildung 5-5 dargestellt und umfasst für die Erstellung eines Prognosemodells die Phasen der Datenaufbereitung sowie die Lern- und Testphase. Sind die für eine hinreichend genaue Modellierung notwendigen

Einflussgrößen detektiert und Teilbereiche identifiziert, in denen aufgrund des unterschiedlichen Verhaltens eine getrennte Modellierung erfolgen soll, müssen die Daten für die Modellerstellung des Netzes vorbereitet werden. Der Phase der Datenauf- und -nachbereitung kommt bei neuronalen Netzen eine besondere Bedeutung zu, da sie die Trainingszeiten und die Güte des Ergebnisses beeinflussen.

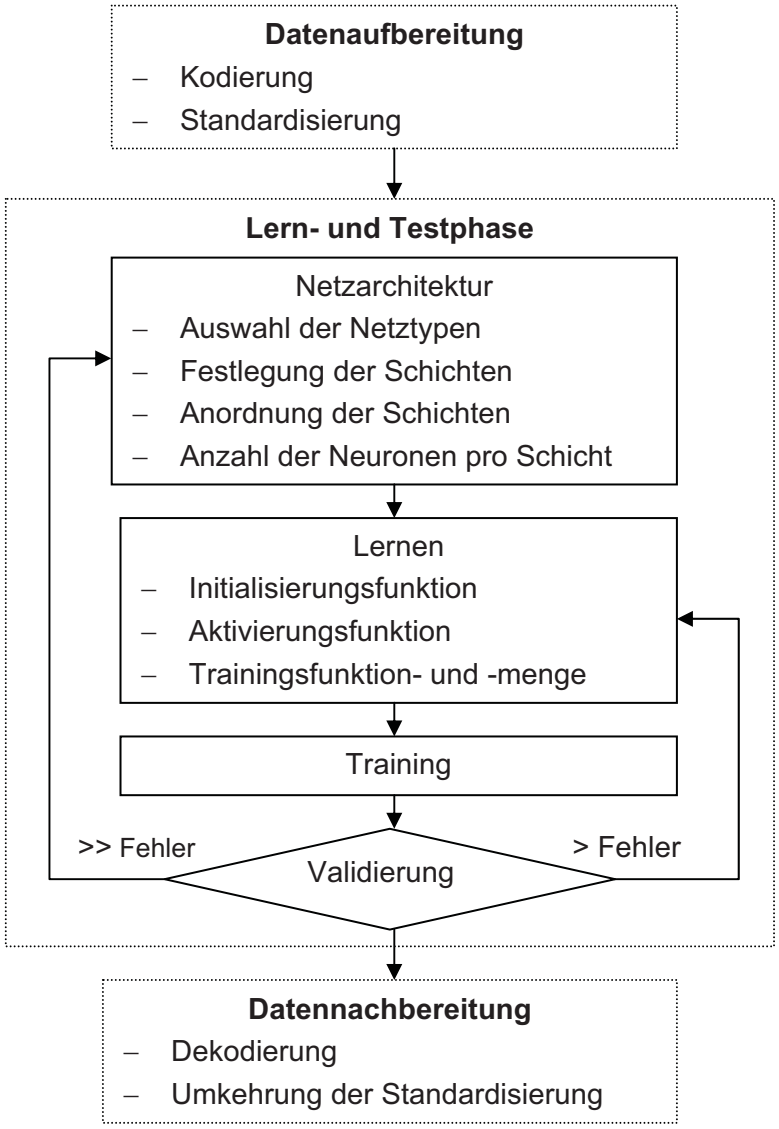


Abbildung 5-5: Entwicklungssystematik

Da neuronale Netze „strukturentdeckende Verfahren“ sind, kann eine geeignete Auswahl der Lerndaten unterschiedlicher Eigenschaften, wie z. B. ein stabiles Verhalten bei einer umfassenden festen Datenbasis von einem oder mehreren Jahren, oder eine schnelle Anpassung an Veränderungen bei einer „wandernden“ Datenbasis, die beispielsweise nur den Zeitraum der letzten sechs Wochen vor der Prognose nutzt, zur Folge haben. Die kombinierte Verwendung von Daten versucht darüber hinaus die Vorteile dieser beiden Ansätze zu verbinden [Gei01]. Üblicherweise werden die Eingangsdaten für ein Netz standardisiert, d. h. die Daten werden auf einen Bereich zwischen 1 und –1 skaliert, um sie an die Ausgabewerte der inneren Aktivierungs-

funktionen der Neuronen anzupassen. Dieses Vorgehen hat den Vorteil, dass das Signal, das auf die Aktivierungsfunktionen gegeben wird, optimal ausgewertet werden kann, denn der Punkt mit der größten Steigung der verwendeten Aktivierungsfunktionen liegt bei 0. Sind die Eingangssignale weit von diesem Arbeitspunkt entfernt, so können diese nicht mehr exakt unterschieden werden, d. h. sie besitzen ähnliche Ausgangssignale. Um nach Verarbeitung der Daten im Netz die Ergebnisse der Ausgabewerte vergleichbar mit den Beobachtungswerten zu machen, muss abschließend im Schritt der Datennachbereitung eine Rücktransformation der Ausgabewerte erfolgen.

Zu der dann folgenden Verarbeitung im neuronalen Netz muss im Arbeitsschritt Netzarchitektur festgelegt werden, wie die unterschiedlichen Netze in der Systemarchitektur zu verknüpfen und wie die einzelnen Netze selbst zu strukturieren sind. Bei der Auswahl der Netztypen (vgl. Kap. 4.2 und folgende) stehen den eigentlichen Schichten bei Jordan- und Elman-Netzen Kontextneuronen, die als Gedächtnis fungieren, zur Verfügung. Die Schichten werden so mit Kontextneuronen verbunden, dass im Netz ein Gedächtnis für Werte der letzten Rechenperiode entsteht. Im Rahmen dieser Arbeit hat sich gezeigt, dass sich mit hierarchischen Elman- und auch Jordan-Netztypen im Verhältnis zu vorwärtsgerichteten Netzen bessere Ergebnisse erzielen lassen.

Neben der Festlegung des Netztypen muss u. a. die Festlegung der Anzahl der Schichten, die Anordnung dieser sowie die Bestimmung der Anzahl der Neuronen pro Schicht erfolgen. Es existieren viele Vorschläge zur Bestimmung der Zahl der Neuronen pro Schicht. Kaastra und Boyd [Kaa96] liefern eine Übersicht zur möglichen Bestimmung. Diese Ansätze beziehen sich auf die Anzahl der Eingangs- und Ausgangsneuronen. Andere Ansätze beziehen die Anzahl der Neuronen auf die zur Verfügung stehenden Trainingsdaten. Der Ansatz von Kolmogorov legt die Obergrenze der Neuronenanzahl fest. Angewendet auf ein neuronales Netz bedeutet dies, dass jede kontinuierliche Zuordnung $\mathrm{y(x)}$ von d Eingangsvariablen x_i auf eine Ausgangsvariable y genau durch ein dreischichtiges neuronales Netz mit $\mathrm{d(2d+1)}$ Eingangsneuronen in der ersten und $(2\mathrm{d}+1)$ Neuronen in der zweiten verdeckten Schicht abgebildet werden kann [Sto00]. Im Anwendungsfall der Lastprognose hat sich gezeigt, dass die Anzahl der inneren Neuronen in Abhängigkeit zu den Eingangsgrößen zwischen 8 - 30 lag, auf die im Laufe des Kapitels „Fallstudien“ detaillierter eingegangen wird.

Nachdem der Netztyp und die Anzahl der Schichten und Neuronen bestimmt wurde, müssen nun noch Initialisierungs- und Aktivierungsfunktion sowie die Trainingsfunktion, d. h. der Lernalgorithmus und die zur Lernphase verwendeten Daten, ermittelt werden. Die Initialisierungsfunktion setzt die Neuronen vor Beginn der Lernphase in einen Anfangszustand. Die Startwerte für vorwärtsgerichtete Verbindungen wurden zufällig zwischen den Werten [0; 1] gewählt, die Verbindungen der Elman-Netze wurden auf den Wert Null gesetzt und Kontextneuronen mit dem Wert 0,5 initiiert. Änderungen

dieser Eingangsparameter haben gezeigt, dass sie keinen Einfluss auf die Güte des Lernvorganges nehmen. Die Auswahl der Aktivierungsfunktion der Neuronen beeinflusst die Reaktion auf das Eingangssignal. Die Untersuchungen erfolgten standardmäßig mit einer logarithmischen Aktivierungsfunktion.

Die eingesetzte Trainingsfunktion, also der Lernalgorithmus, nimmt den größten Einfluss auf den Trainingserfolg und damit auf die Güte der Prognose. Im Kapital 4.2 wurden neben dem Backpropagation- und Feedforward-Alogrithmus diverse Alternativen und Weiterentwicklungen vorgestellt. Im Rahmen der Untersuchungen konnte kein Unterschied in der Effektivität und im Prognosefehler der neuronalen Netze festgestellt werden. Auch bei langen Trainingszeiten, d. h. einem Durchlauf von mehr als 30.000 Zyklen, konnten keine Anzeichen für Fehlerplateaus oder Oszillationen beobachtet werden. Lediglich die Trainingsdauer unterschied sich merklich, sodass beispielsweise das sehr einfache Gradientenverfahren durch lange Laufzeiten auffiel und zudem höhere Fehlerquoten im Vergleich zu anderen Algorithmen aufwiesen.

Nach Auswahl aller zuvor genannten Parameter erfolgt nun das Training des Netzes, bei dem aus den Zeitreihen Inputmuster aufgebaut werden, gegen die das Netz ein Outputmuster trainiert. Die gemessenen Zeitreihen der Modellphase wurden hierzu in eine Menge von Trainings- und Validierungsdaten aufgeteilt. Für die Untersuchungen hat sich gezeigt, dass ein Verhältnis von 80 : 20 zu guten Ergebnissen führt. Des Weiteren sollte gerade bei der Lastprognose darauf geachtet werden, dass bei der gewählten Einteilung jeweils nur ganze Tage, also immer Einheiten von 96 Datensätzen, den Datengruppen zugeordnet werden. Die Zuordnung der ganzen Tage kann dabei zufällig erfolgen.

Die Beendigung des Lernvorganges kann anhand dreier Kriterien erfolgen: Erstens, wenn die Anzahl der vorgegebenen Lernschritte erreicht wurde, oder zweitens, wenn der Abbildungsfehler die zuvor definierten Konvergenzbedingungen erfüllt oder drittens, es zum Abbruch des Lernvorganges führt, weil das Netz Gefahr läuft, übertrainiert zu werden. Die vorgenannten Bedingungen werden nach jeder Lernepoche mit den Testdaten geprüft. Nach Beendigung der Lernphase wird mit Hilfe von Validierungsdaten entschieden, ob durch ein zielgerichtetes Anpassen der vorhergehenden Schritte weitere Verbesserungen erreicht werden können. Bei Bedarf sind dabei nicht nur die Schritte „Netzarchitektur“ und „Lernen“ erneut zu durchlaufen, sondern auch der Schritt „Datenaufbereitung“, wenn das Validierungsergebnis es nahelegt, die Einflussgrößen und deren Aufbereitung zu verbessern.

Abschließend erfolgt die Datennachbereitung, bei der es zu einer Dekodierung und zu einer Umkehrung der Standardisierung der Daten kommt. Ist auf diese Weise ein Modell entstanden, sollte abschließend noch eine Strategie entwickelt werden, wie innerhalb des Modells aktuelle Entwicklungen auf das Lastverhalten berücksichtigt werden können, um zu vermeiden, dass alle zuvor durchgeführten Schritte der

Entwicklungsphase erneut durchlaufen werden müssen. Hierbei könnten die zuvor genannten Vorgehensweisen entsprechend angewendet und ab dem Entwicklungsschritt „Lernen“ ein verändertes Lastverhalten berücksichtigt werden.

5.4 Mustererkennung in Lastgangdaten mit SOM

Die Analyse historischer Daten und deren Clusterung bilden die Grundlage einer systematischen Entwicklung hin zum Lastprognosesystem. Da die Einflussgrößen auf die elektrische Last in starkem Zusammenhang mit Ereignissen stehen, ist eine Datenclusterung zur Bestimmung des Tagtyps notwendig. Sowohl die Identifizierung der Sondertage als auch die Einteilung der Lastmuster in typische Gruppen wurden im Rahmen einer Lastmusteranalyse mit Hilfe von SOM bei den Untersuchungen mit Datensätzen verschiedener Kollektive durchgeführt.

5.4.1 Entwicklung einer zweidimensionalen Netzstruktur zur Analyse von Lastgangdaten

Eine SOM ist im Regelfall ein sehr einfach konzipiertes zweischichtiges Netz, das aus einer Input- und einer Kohonenschicht besteht. Im Rahmen des entwickelten Verfahrens wurde die Kohonenschicht um eine weitere aktorische Output-Schicht ergänzt. Diese dient in besonderer Weise der Interpretation der von der Kohonenschicht davon unabhängig erlernten Tagtyporganisation. Abbildung 5-6 veranschaulicht die Architektur des eingesetzten Netzes.

Das implementierte Netzwerk wurde aus drei Schichten (Output-, Kohonen- und Inputlayer) aufgebaut. Input- und Kohonenlayer entsprechen denen einer Self-Organizing-Map. Ergänzend wurde die SOM um den Outputlayer erweitert, der mit dem Begriff ‚motorische Karte’ beschrieben werden kann. Jedes Neuron im Kohonenlayer hat sieben Ausgänge zum Lernen der motorischen Karte, da die Neuronen einen 7-Bit-Code ausgeben. Das heißt, jedes Neuron im Outputlayer hat eine Anzahl von Eingängen entsprechend der Anzahl der Neuronen in der sensorischen Karte. Jeder der in Abbildung 5-6 dargestellten Linien repräsentiert einen Gewichtsvektor. Im Anwendungsfall der Lastdatenanalyse hat der Eingangsvektor die Dimension von 96, um so den üblichen Merkmalsraum der viertelstündlich gemessenen Lastgangdaten eines Tages abzubilden. Die Inputschicht ist vollständig mit den Neuronen der „lernenden“ Kohonenschicht verbunden. Diese wurde als zweidimensionales Raster ausgeprägt, wodurch zusätzlich zwischen den einzelnen Neuronen eine Nachbarschaftsbeziehung hinsichtlich der Lage berücksichtigt wird. Jedes Neuron der Kohonenschicht ist mit den sieben Neuronen der Outputschicht verbunden, die auch als motorische Karten bezeichnet wird [Kin94] und den Eingangsdatensatz einem Wochentag zugeordnet.

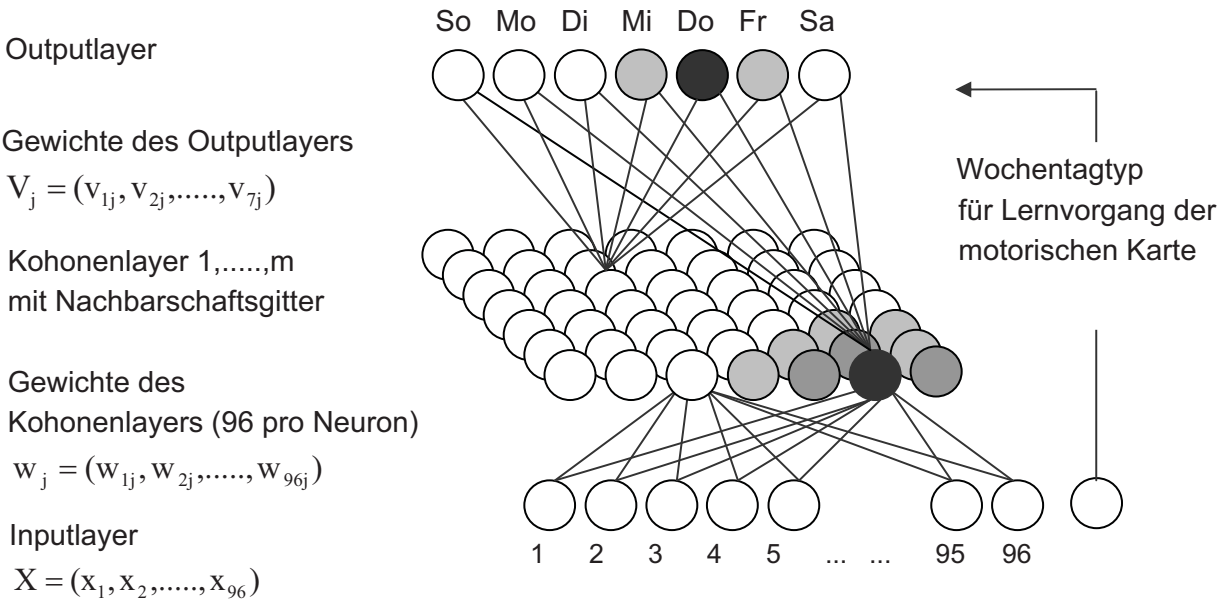


Abbildung 5-6: Netzstruktur einer zweidimensionalen selbstorganisierenden Karte

Das Lernverfahren des aktorischen Outputlayers ist im Gegensatz zum Lernen der Kohonenschicht ein überwachtes Lernen. Bei jedem Iterationsschritt werden die Gewichte des Outputlayers mitgelernt. Das Lernen erfolgt mit Unterweisung (Supervised Learning), da der Ausgabevektor der Neuronen im Kohonenlayer vorgegeben wird, und somit das „gewünschte“ Ergebnis der Eingangsgewichte des Outputlayers. Der Ausgabevektor hängt dabei ausschließlich vom Wochentag des Lerndatensatzes ab. Der Wochentag wird aus dem Datum vom Datensatz ermittelt. Alle Neuronen in der Kohonenkarte geben den entsprechenden Vektor aus. Der Ausgang 1 von jedem Neuron der Kohonenkarte ist mit demselben Neuron auf dem Outputlayer verbunden. Dadurch kann einem Neuron auf dem Outputlayer eindeutig der Wochentag zugeordnet werden. Innerhalb des Radius ‚Sigma’ werden die Eingangsgewichte des Outputlayers mit Eins berechnet. Das heißt, es werden nicht nur die sieben Eingänge der Outputneuronen, die mit dem Gewinnerneuron der Kohonenkarte verbunden sind, adaptiert, sondern alle Eingangsgewichte des Outputlayers, die zu den Neuronen in einem bestimmten Radius um das Gewinnerneuron liegen.

5.4.2 Lernverfahren und Datenanalyse mit SOM

Vor dem Beginn der Lernphase erfolgt zunächst eine Normierung aller Eingangswerte. Um eine inhomogene räumliche Verteilung auf der Kohnenkarte zu Beginn des Lernprozesses zu vermeiden, werden vor dem Lernen der Karte die Eingangsgewichte der Neuronen im Outputlayer und die Eingangsgewichte der Neuronen im Kohonenlayer Gewichtsvektoren $w_j = (w_{1j}, w_{2j}, ..., w_{nj})$ zufällig im Bereich zwischen Null und Eins initialisiert. Die Neuronen im Kohonenlayer sind in Gitterform angeordnet, mit einem

Abstand von Eins zwischen den direkten Nachbarn. Beim Lernverfahren wird der 96-dimensionale Eingangsvektor $X = (x_1, x_2, ..., x_{96})$, der die viertelstündlichen Lastgangdaten eines Tag repräsentiert, mit den Gewichtsvektoren des Kohonenlayers mit Hilfe einer beliebigen Metrik verglichen. Dabei wird das Neuron „n“, dessen Gewichtsvektor W_n dem Eingangsvektor X am ähnlichsten ist, als Gewinnerneuron des Iterationsschrittes detektiert. Mit Hilfe der Gleichung (4-14) wird das Gewinnerneuron n bestimmt: Im folgenden Schritt erfolgt eine Adaption der Gewichte des Kohonenlayers. Hierbei erfolgt nicht nur die Anpassung des eigentlichen Gewinnerneurons, sondern alle sich in topologischer Nachbarschaft befindlichen Gitterneuronen erfahren eine Gewichtsanpassung, deren Beeinflussung mit der Anzahl der Iterationsschritte kleiner wird. Im Detail sieht das Lernverfahren zur Anpassung der sensorischen Karte, kombiniert mit der Anpassung der motorischen Ausgabeschicht, wie folgt aus [Rit90]:

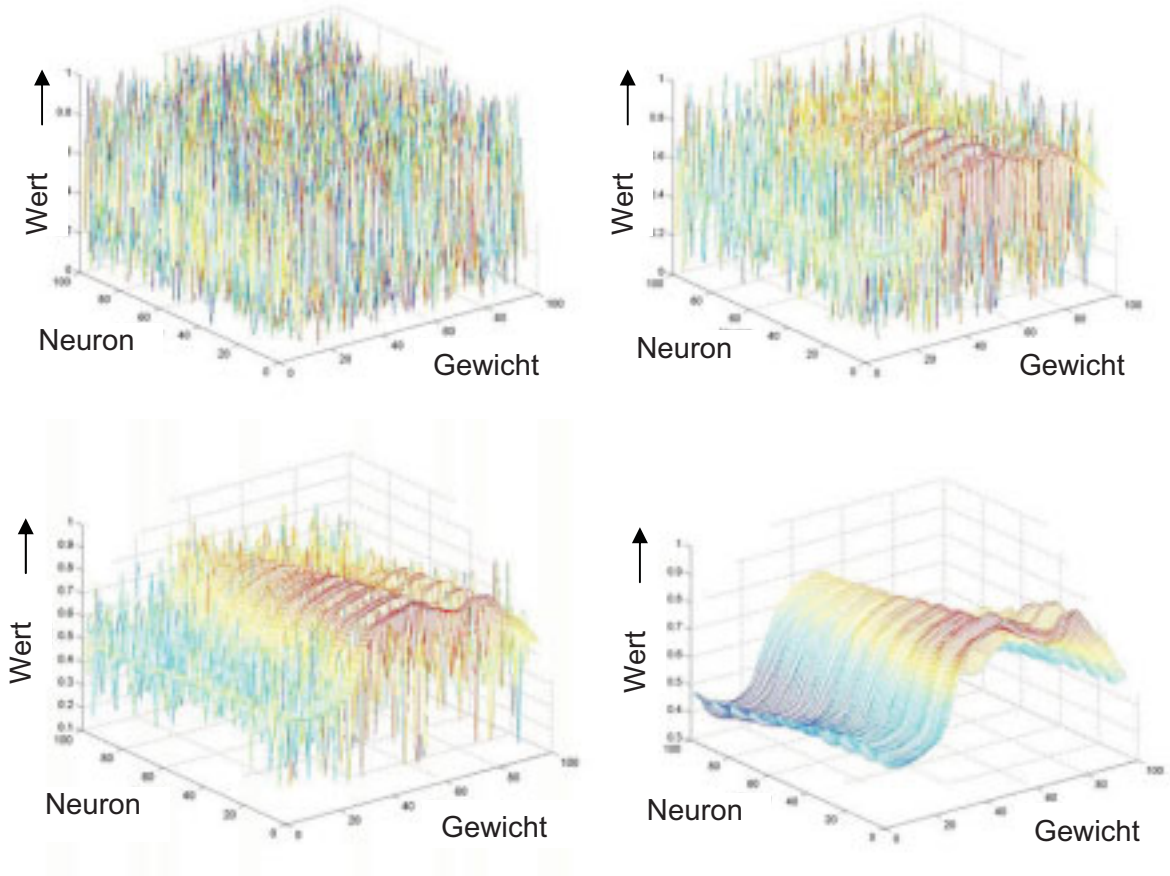


Abbildung 5-7: Gewichtsvektoren des Kohonenlayers nach 5, 50, 100 und 500 Iterationsschritten

1. Wähle zufällig alle Gewichte W_k und V_k.
2. Gib einen Eingabevektor $X = (X_1, \ldots, X_{96})$ vor (sensorisches Signal). X wird zufällig aus den vorhandenen Datensätzen ausgewählt.
3. Suche das Gewinnerneuron n gemäß Gleichung (4-47).

4. Adaptiere alle Gewichte W_k der Kohonenschicht gemäß Gleichung (4-48) bezogen auf das Gewinnerneuron n.
5. Berechne alle Gewichte V_k der motorischen Schicht gemäß:
   $V_k(t+1) = V_k(t) + \eta'(t) \cdot h'_{n,k}(t) \cdot (Y - V_k(t))$.
   Dabei ist $\eta'(t)$ ein weiterer Lernfaktor, der mit wachsender Iterationszahl t gegen 0 geführt wird; $h'_{n,k}(t)$ beschreibt eine weitere Nachbarschaftsfunktion gleichen Typs wie $h_{n,k}(t)$, jedoch mit eigener Varianzführung $\sigma'(t)$. Y ist die zu einem Lastverlauf X zugehörige siebendimensionale Tagbeschreibung.
6. Erhöhe t um eins und mache weiter bei Schritt 2.

Das Lernverfahren des aktorischen Outputlayers ist im Gegensatz zum Lernen der Kohonenschicht ein überwachtes Lernen. Wie sich im Verlauf des Lernens die Gewichtsvektoren einer zweidimensionalen Karte der Größe 10 x 10 entwickeln, dokumentiert die Abbildung 5-7. Dabei sind die 100 Gewichtsvektoren der Karte bei unterschiedlichen Iterationsschritten dargestellt. Nach dem Abschluss des Lernvorgangs der SOM können nun mit Hilfe der SOM-Karten Datenanalysen erfolgen. Eine der einfachen Möglichkeiten diese durchzuführen, kann durch erneute Präsentation der zuvor zum Lernen der Karte verwendeten Eingangsvektoren erfolgen. Mit Hilfe von Filterfunktionen können die einzelnen Wochentagetypen getrennt voneinander an die SOM angelegt werden und mit externen Tools, wie z. B. Matlab, übersichtlich dargestellt werden.

Abbildung 5-8 zeigt die Trefferanzahl der räumlichen Verteilung von Gewinnerneuronen des Wochentages eines Jahres auf der Kohonenkarte mit der Dimension von 10 x 10 Neuronen. Bei ersten Begutachtungen der Trefferanzahl von Gewinnerneuronen eines Wochentages fällt auf, dass sich die Tage vom Wochenende offensichtlich homogener innerhalb eines Areals auf der Karte abbilden als Wochentage. Der Bereich, in dem sich Wochenenden entfalten, begrenzt sich im Beispiel fast nur auf das linke obere Viertel der Karte. Differenziert man innerhalb der Wochenenden Sonntage und Samstage voneinander, lässt sich zeigen, dass auch zwischen den beiden Tagtypen auf der Karte deutliche räumliche Abgrenzung festzustellen sind. Bei den Wochentagen ist eine strikte Abgrenzung in eigenständige Gebiete nicht so offensichtlich. Diese Tage breiten sich zwar mehrheitlich auf einem Teil der Karte aus, die von den Wochenendtagen nicht belegt werden, es kommt jedoch zu Überschneidungen der räumlichen Abbildung.

Eine weiterführende Analysemöglichkeit besteht darin, die bereits während des Lernvorgangs verwendeten Lastgangdaten erneut, erweitert um die zugehörigen über das Datum referenzierten Merkmale des jeweiligen Tages, den Karten zuzuführen. Hierzu wurden zusätzliche Filterfunktionen entwickelt, um die Karten für Analysezwecke zu nutzen. Die mit Hilfe des Filters ermittelten Lastgangtage werden der sensorischen Karte zugeführt und, ähnlich, wie bereits beim Lernvorgang beschrieben, über eine Metrik ein Gewinnerneuron der sensorischen und auch motorischen Karte ermittelt. Mit

Hilfe einer graphischen Ausgabe der sensorischen Karte wie auch der motorischen Karte lassen sich nun Datencluster entsprechend der Filtereinstellung detektieren. Der schematische Ablauf des Analysevorgangs ist in Abbildung 5-9 dargestellt.

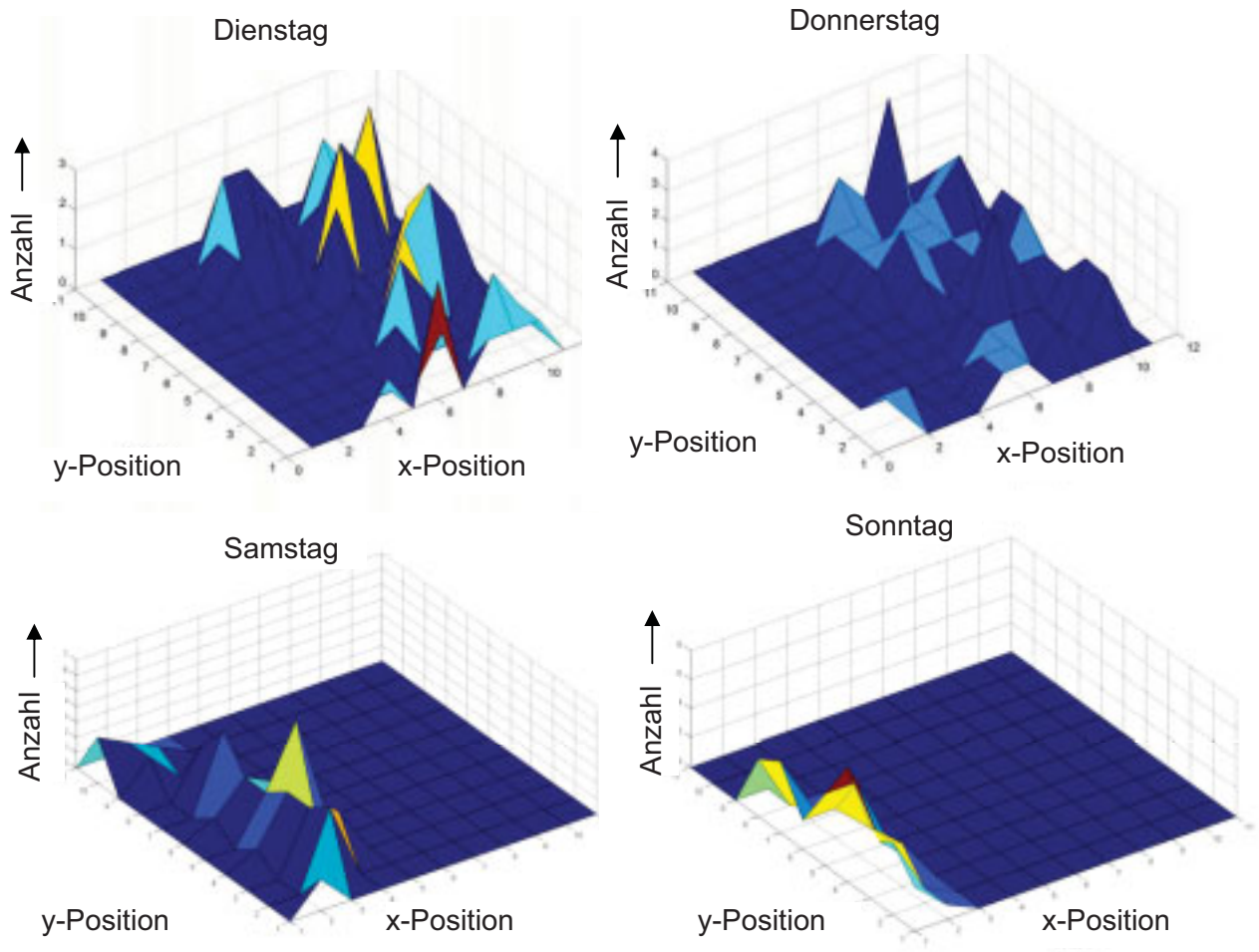


Abbildung 5-8: Entfaltung von Dienstag, Donnerstag, Samstage und Sonntage auf der Kohonenkarte

Im Rahmen der Untersuchungen wurden Programme entwickelt [Kuh04] [Wel05], mit deren Hilfe zweidimensionale Kohonenkarten mit variabler Größe und Outputlayer erstellt und zugehörige Merkmale von Tagtypen analysiert werden können. Mehrere Karten können als Set gespeichert und geladen werden. Die Programme ermöglichen die Einstellung aller Variablen jeweils getrennt für die Kohonenkarte und den Outputlayer. Zusätzlich kann durch eine Protokollfunktion dokumentiert werden, mit welchen Parametern die Karten gelernt wurden. Weiterhin ist eine optische Auswertung der Kohonen- und des Outputlayers möglich.

Mit Hilfe eines Filters können bestimmte Ereignisse, Wochentage oder Monate auf der Kohonenkarte angezeigt werden. Es stehen verschiedenste Ansichten und Filtereinstellungen für die Datenanalyse zur Verfügung. Die Programme sind als Multi-

Document-Anwendung programmiert und wurden in der Entwicklungsumgebung Microsoft Visual Studio .NET mit Programmiersprache C# erstellt. Durch Exportfunktionen können die Ergebnisse in anderen Programmen weiter verarbeitet werden.

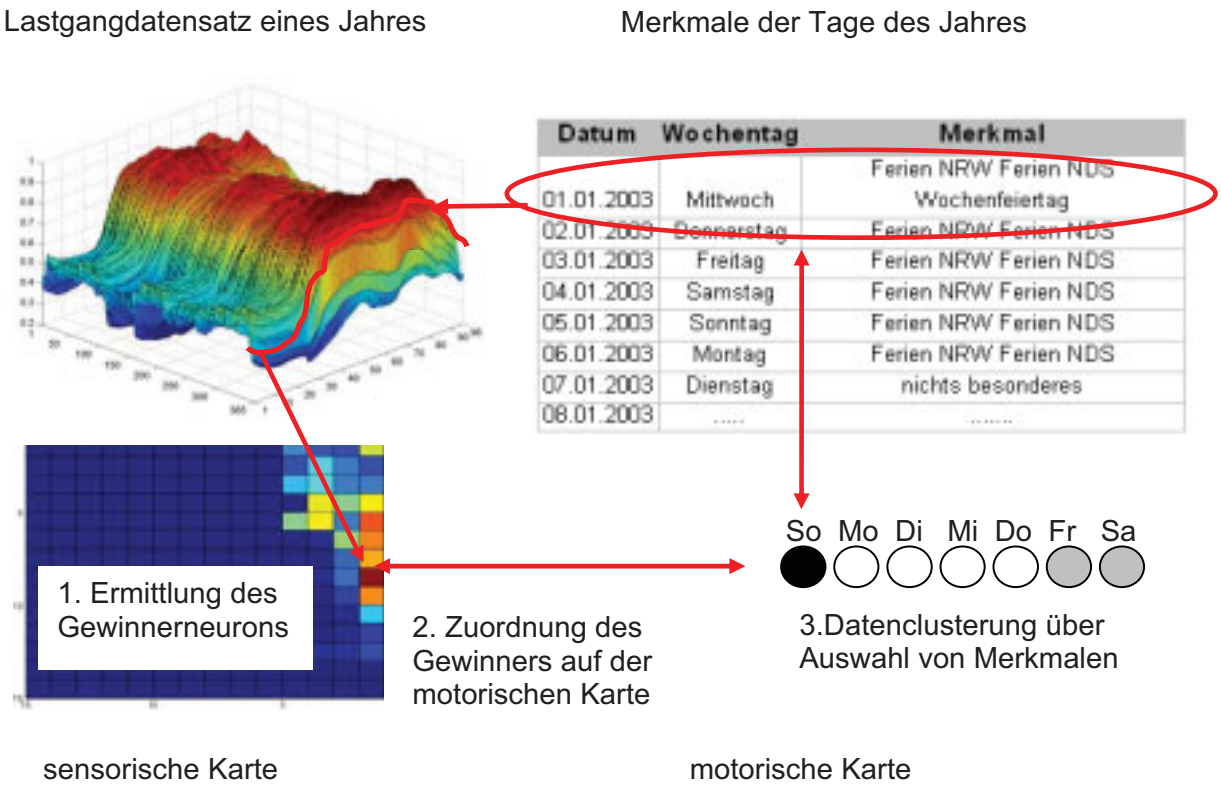


Abbildung 5-9: Schematischer Ablauf der Datenanalyse mit SOM

6 Fallstudien

6.1 Untersuchungsdaten und Vorgehensweise

Die Gesamtlast eines Versorgungsnetzes wird durch verschiedene Faktoren beeinflusst. Diese Faktoren sind im jeweiligen EVU unterschiedlich und stark von der Zusammensetzung der Verbraucherkollektive abhängig. Im Rahmen der folgenden Untersuchungen standen Lastgangdaten von fünf Jahren im ¼-h-Zeitraster zur Verfügung. Aufgrund der unterschiedlichen Netzgebiete lassen sich die Kollektive in folgende Klassen gruppieren:

- ländliches Netzgebiet
- Mischung aus städtischem und ländlichem Netzgebiet
- städtisches Netzgebiet

Auf Basis bestehender Betriebsdatenbanken, die verschiedene Verbrauchs- und Prozessdaten aus den Jahren 1998 bis 2003 enthielten, wurden für die weitere Analyse die Daten im Rahmen von [Mie03] analysiert. Diese Rohdaten wurden im Datascreening zunächst auf Datenfehler, z. B. Lücken oder Ausreißer, untersucht.

Tagtyp	Wochentag	Ereignis	Zeitzone
nichts besonderes	Montag	nichts besonderes	MEZ
Ferien Nds.	Dienstag	Volksfest (Frühjahr)	MESZ
Ferien NRW	Mittwoch	Stadtfest	EZ/SZ
Ferien NRW Ferien Nds.	Donnerstag	Volksfest (Herbst)	
Ferien NRW Brückentag	Freitag	Weihnachtsmarkt	
Ferien Nds. Brückentag	Samstag	WM-Spiel	
Ferien NRW Ferien Nds. Brückentag	Sonntag		
Ferien NRW Ferien Nds. Wochenfeiertag			
Ferien Nds. Wochenfeiertag			
Ferien NRW Wochenfeiertag			
Brückentag			
Wochenfeiertag			

Tabelle 6-1: Tagesmerkmale

Da auch Datenaufzeichnungen aus automatischen Systemen fehlerhafte Werte enthalten können, wurden die hierbei als fehlerhaft identifizierten Daten entweder in Absprache mit den Anwendern korrigiert oder von der weiteren Behandlung im Rahmen der Lastganganalyse ausgeschlossen. Die Identifikation von Fehlern wurde mit Hilfe linearer Filter unter Festlegung von Grenzwerten durchgeführt. Lag dabei ein Wert oberhalb eines festgelegten Grenzwertes, wurde dieser als Ausreißer identifiziert und ausgefiltert. Dabei wurde der Grenzwert anhand einer Häufigkeitsverteilung aller Abweichungen festgelegt. Die gefilterten Werte außerhalb des Grenzwertes wurden dabei, wenn möglich, durch Ersatzwerte ergänzt. Zur Verarbeitung der Lastgangdaten innerhalb unterschiedlicher Programme wurden die 96 Lastwerte eines Tages normiert und in entsprechende Datenformate zur weiteren Verarbeitung konvertiert. Die auf-

bereiteten Daten bilden die Grundlage für die folgenden Untersuchungen. Zusätzlich erfolgten Aufzeichnungen von Merkmalen der Kalendertage sowie lokaler und globaler Ereignisse der Kollektive. Die für die nachfolgenden Untersuchungen genutzten Merkmale sind in Tabelle 6.1 dargestellt.

Als weitere Datensätze und Eingangswerte für Lastprognosen standen Wetterbeobachtungs- wie auch Prognosewerte zur Verfügung. Die Bereitstellung von Wettermess- und Prognosewerten erfolgt überwiegend innerhalb der deutschen Energiewirtschaft durch den Deutschen Wetterdienst (DWD). Dieser stellt Wettervorhersagen in vielfältiger Form zur Verfügung. Im Rahmen der Untersuchungen wurden die Parameter täglich ab 06:00 Uhr auf dem FTP-Server in stündlicher Auslösung als 72-Stunden- und in dreistündlicher Auflösung als 120-Stunden-Kurzfristprognose zur Verfügung gestellt. Eine Übersicht der im Rahmen der Untersuchungen verwendeten Punkt-Termin-Prognosen ist in Tabelle 6.2 dargestellt. Für die weiterführenden Untersuchungen wurden Wetter- und Lastgangdaten von fünf zurückliegenden Jahren verwendet. Da energiewirtschaftlich Lastgangdaten ¼-stündlich aufgezeichnet und bilanziert werden, wurden die vorliegenden Wetterdaten linear interpoliert, so dass für jeden Lastgangwert ein Wetterdatum in gleicher zeitlicher Auflösung vorlag.

Wetterelement	Beobachtungs-werte	Prognose-werte	Maß-einheit	Abkürzung
Temperatur 2 m Höhe	X	X	° C	T2M
Temperatur Boden	X	X	° C	TB
Windrichtung	X	X	°	WR
Windgeschwindigkeit	X	X	km/h	WG
Luftdruck	X	X	mhPa	LD
Luftfeuchtigkeit	X	---	%	LF
Bedeckungsgrad	X	X	8/8	BD
Globalstrahlung	X	X	J/cm^2	GSS
Niederschlag	X	---	mm	NS

Tabelle 6-2: Übersicht der Wettermess- und Prognosewerte

Um die Güte der verwendeten 72-h- wie auch 120-h- Wetterprognosen mit den Beobachtungswerten beurteilen zu können, wurden statistische Analysen durchgeführt. Hierfür wurden Prognosewerte mit den Beobachtungswerten eines Zeitraumes von drei Jahren verglichen und analysiert. Da die Wettermesspunkte nicht immer exakt den aktuellen Zustand aufgrund von örtlichen Schwankungen innerhalb dieses Gebietes aufzeichnen, sind grundsätzlich Differenzen durch die Aufzeichnungssystematik bedingt. Des Weiteren erfolgt die Vorhersage jedes einzelnen Elementes auf der Basis unterschiedlicher Methoden. Im Ergebnis lässt sich folgendes zusammenfassen: Die relative Trefferwahrscheinlichkeit ist stark vom jeweils zu prognostizierenden Wetterelement abhängig. Grundsätzlich liegt die Güte der 72-h- Prognose über der von 120-h-Vorhersagen. Die durchschnittliche Gesamtabweichung aller Vorhersagewerte der 72-h

kann mit 5,2 %, die der 120-h-Vorhersage mit 16,6 % angegeben werden. Eine ausführliche Analyse erfolgte in [Ste04]. Im Rahmen der Untersuchungen wurde eine Vielzahl unterschiedlicher Prognoseverfahren und -methoden miteinander in Bezug auf die Güte und Qualität verglichen und bewertet. Da die Ermittlung der Prognosefehler von Zeitreihen der in Kapitel 3.4.2 beschriebenen Kenngrößen mit hohem manuellem Aufwand verbunden ist, wurde in einer Diplomarbeit [Bil03] ein Programm in der Entwicklungsumgebung von .NET erstellt. Das Programm hat eine Schnittstelle, um beliebige Prognose-Dynamic Link Library (DLL) zu integrieren, wurde modular aufgebaut und verfügt über eine graphische Oberfläche. Die Prognosealgorithmen werden in Form von DLL an den Programmkern angebunden. Damit ist sichergestellt, dass unabhängig von der verwendeten Programmiersprache zur Erstellung der Algorithmen eine Anbindung möglich ist. Somit können Prognosemethoden wie neuronale Netze, Fuzzy-Logic-Konzepte oder Hybridsysteme innerhalb einer Umgebung miteinander verglichen werden. Weiterhin können über Importschnittstellen dem Programmkern CSV-Dateien mit Lern- und Testdatensätzen von Lastgangzeitreihen zur Verfügung gestellt werden. Nach dem Abschluss der manuellen Konfiguration aller nötigen Prognoseparameter testet das Programm die Bibliotheken mit den Datensätzen und die erhaltenen Lastprognosen werden mit einer ausführlichen Fehlerrechnung bewertet.

6.2 Lastgang- und Wetterdatenanalyse

Dieser Abschnitt befasst sich mit der Untersuchung von Last- und Wetterdaten. Dabei sollen grundlegende Eigenschaften der Daten ermittelt und darauf aufbauend Gemeinsamkeiten zwischen ihnen detektiert werden. Mit diesen Informationen soll dann im nächsten Schritt eine Auswahl an Wetterelementen getroffen werden, die als Einflussgrößen für die Lastprognose von Relevanz sind. Das Kapitel ist so aufgebaut, dass im ersten Teil eine detaillierte Untersuchung der Lastgangdaten erfolgt. Im zweiten Abschnitt werden die Wetterelemente und ihre Beziehung zueinander analysiert. Des Weiteren erfolgen Untersuchungen, ob diese Elemente jeweils einzeln für sich betrachtet werden müssen oder ob eine Gruppierung möglicher Einflussgrößen mit ähnlichen Eigenschaften möglich ist. Im letzten Abschnitt werden schließlich Zusammenhänge zwischen den Wetterelementen und dem Lastverlauf herausgearbeitet. Die ausführlichen Untersuchungen erfolgten im Rahmen der Diplomarbeit [Fre05].

6.2.1 Lastgang

Der Lastgang in der Darstellungsform einer Zeitreihe spiegelt die Höhe der Last in Abhängigkeit von der Zeit wider. Die Abbildung 6-1 zeigt die Lastverläufe einzelner Tage einer Woche, die in viertelstündlichen Abständen gemessen werden. An dem Beispielverlauf wird weiterhin ersichtlich, dass arbeitsfreie Tage, wie in diesem Fall das

Wochenende, einen von den Werktagen (Montag bis Freitag) abweichenden Verlauf haben. Die Werktage scheinen dagegen nur geringe Abweichungen voneinander aufzuweisen. In den folgenden Unterabschnitten wird der Lastgang mit geeigneten Methoden auf seine wichtigsten Eigenschaften hin untersucht.

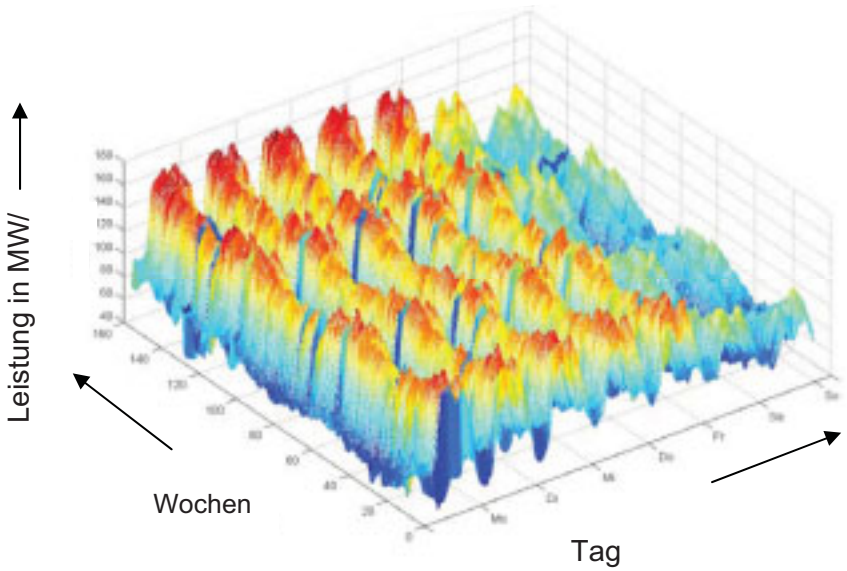


Abbildung 6-1: Lastgang in Wochenrasterdarstellung

6.2.2 Empirische Momente und Stationarität

Die empirischen Momente der Lastgangzeitreihen sind exemplarisch in Tabelle 6.3 dargestellt. Im Folgenden wurde der Lastgang auf Mittelwertstationarität und Varianzstationarität untersucht. Für die Überprüfung einer Zeitreihe auf Mittelwertstationarität steht der zwei-Stichproben-t-Test zur Verfügung. Die Erfüllung der Varianzstationaritätsbedingung wurde mit dem $\chi 2$-Test ermittelt. Die Berechnung der Freiheitsgrade und Prüfgrößen erfolgte in der Diplomarbeit von [Kar05]. Deshalb wird hier nur kurz auf den Testablauf und die Testergebnisse eingegangen. Zwei-Stichproben-t-Test und $\chi 2$-Test haben einen ähnlichen Ablauf, bei dem an zufälligen Stellen aus dem Lastgang Stichproben mit einer Länge von 96 Werten entnommen werden. Diese Länge entspricht bei viertelstündlichen Messungen einem Tag. Exemplarisch sind die Ergebnisse des Mittelwertstationaritätstests sowie des Varianzstationaritätstests (im Anhang 9.7) dargestellt. Um die Aussagekraft zu erhöhen, wurden für jede Testreihe 20 verschiedene Stichproben entnommen und überprüft. Die Prüfgrößen wurden in der Tabelle direkt in p-Werte umgerechnet, die die Wahrscheinlichkeit angeben, mit der eine falsche Annahme vorliegt, falls das Ergebnis für gültig erklärt wird. Zusätzlich liegt eine Auswertung der Tests für die Irrtumswahrscheinlichkeit $\alpha = 0,05$ vor. Bei den Testergebnissen des t-Tests, die ungleich Null sind, wird gefolgert, dass $\bar{x}_t \neq x_t$, ist. Ähnliches gilt für den $\chi 2$-Test. Ein Ergebnis von -1 bedeutet, dass $s_t^2 < s^2$ ist, und 1

bedeutet, dass $s_t^2 > s^2$ ist. Wie aus den Testergebnissen hervorgeht, gibt es innerhalb des Lastgangs des Öfteren Bereiche, für die keine Stationarität gegeben ist. Die Zeitreihe ist daher insgesamt in ihrem vorliegenden Zustand als instationär zu bewerten.

Mittelwert $\bar{x}$ in MW	Standardabweichung s in MW	Varianz s^2 in MW^2
87,86	25,69	660,06

Tabelle 6-3: Empirische Momente der Lastgangzeitreihe

6.2.3 Trend- und Periodenanteil

Trend- und Periodenanteile ermöglichen, das mittlere Niveau einer Lastgangzeitreihe zu ermitteln. Im Folgenden werden exemplarisch die Lastgangdaten eines Untersuchungsjahres analysiert. Der Lastgang wird dabei als Zusammensetzung aus dem Trendanteil m_t, dem Periodenanteil z_t und einer Restkomponente r_t gemäß der Gleichung (3-23) verstanden. Auf die Restkomponente wird im nachfolgenden Kapitel noch näher eingegangen.

Die Verwendung von Polynomen zur Modellierung der Trendfunktion ist nicht frei von Nachteilen, denn eine globale Gültigkeit des verwendeten Modells wird hierbei vorausgesetzt, lokale Eigenheiten werden nur schwach gewichtet und alle Beobachtungen gehen gleich stark ein. Zusätzlich laufen polynomiale Trends schnell gegen $\pm\infty$, was zur Folge hat, dass der prognostizierbare Zeitraum begrenzt ist. Es stellt sich daher die Frage nach dem optimalen Grad des verwendeten Polynoms. Aufgrund der periodischen Struktur der Zeitreihe würde die Steigerung des Polynomgrades zu einer weiteren Minimierung des Fehlers führen, aber auch zur Approximation von zyklischen Eigenschaften. Das Trendmodell ist dazu jedoch nicht geeignet, da für eine hinreichende Abbildung ein Polynom mit sehr hohem Grad benötigt wird. In der Regel verwendet man Polynome vom Grad 3 oder 4 zur Approximation eines Trends. Der Trendanteil des Lastgangs zweier Jahre wird exemplarisch mit Hilfe der linearen oder polynomialen Regression auf lineare oder polynomiale Eigenschaften geprüft. Dabei wird die „Methode der kleinsten Quadrate“ angewendet. Das Ergebnis dieser Auswertung ist ein Trendpolynom der Form: $p(x_t) = a_0 + a_1 x + \cdots + a_j x_m$. Die Koeffizienten a_0 bis a_j werden unter Anwendung der Methode der kleinsten Quadrate so berechnet, dass eine bestmögliche Annäherung des Trendpolynoms an die Zeitreihe erreicht wird.

Die Abbildung 6-2 zeigt die Verläufe der Trendpolynome. Dabei wird ersichtlich, dass der hier gezeigte Lastverlauf über das Jahr einem geringfügig positiven linearen Trend folgt. Es ist nicht davon auszugehen, dass der Lastgang einem quadratischen Trend folgt.

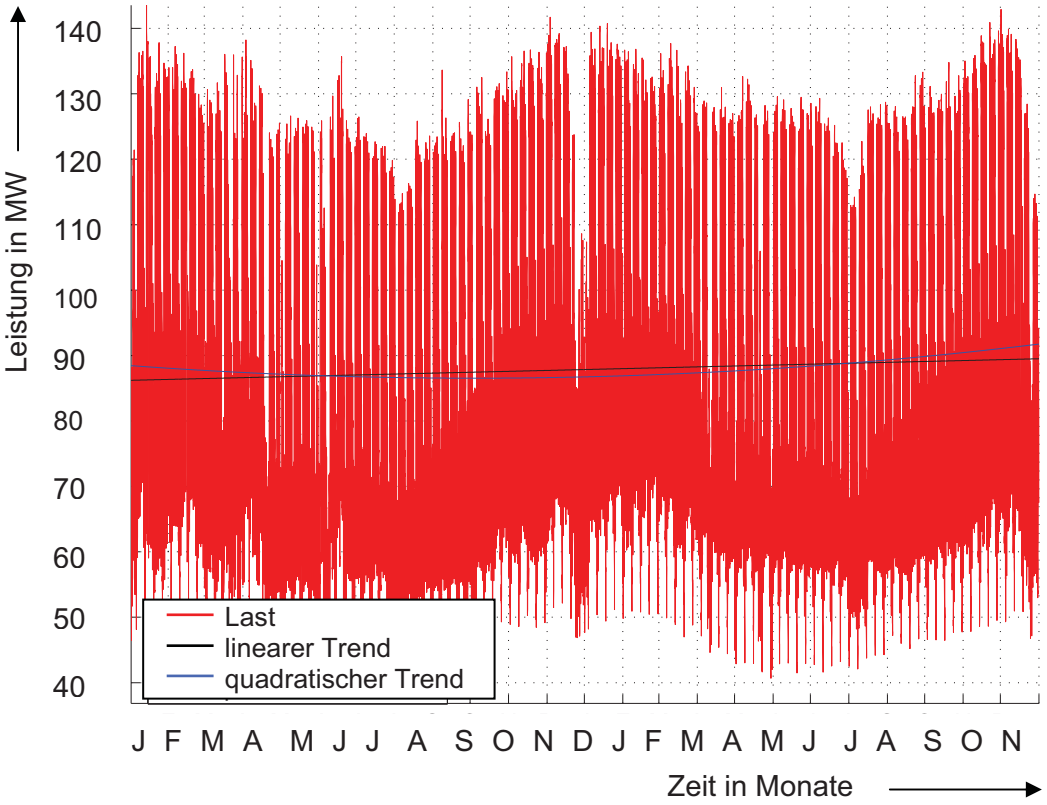


Abbildung 6-2: Verlauf des linearen und des quadratischen Trends

Die Koeffizienten der beiden untersuchten Trendfunktionen sind in Tabelle 6-4 aufgeführt.

p(x)	x	x^1	x^2
linear	8,6215E+01	4,6906E-05	0
quadratisch	8,8457E+01	-1,4480E-04	2,7000E-09

Tabelle 6-4: Koeffizienten des linearen und des quadratischen Trends

Der Periodenanteil des Lastgangs kann mit der Methode der Autokorrelation und der Auswertung des zum Lastgang gehörigen Periodogramms ermittelt werden. Die Autokorrelationsfunktion ist dabei eine Normierung der Autokovarianzfunktion gemäß Gleichung (3-20) und (3-19). Mit der Autokovarianzfunktion lassen sich um τ-Zeiteinheiten gegeneinander versetzte Beobachtungen innerhalb derselben Zeitreihe auf lineare Abhängigkeiten hin untersuchen. $\overline{x}_{(1)}$ und $\overline{x}_{(2)}$ sind die arithmetischen Mittelwerte der jeweiligen Komponente. Ist die Zeitreihe stationär, so kann $\overline{x}_{(1)} = \overline{x}_{(2)} = \overline{x}$ gesetzt werden [Ger03].

Der Lastgang erfüllt ohne Bereinigung um Trend- und Periodenanteile nicht die Bedingungen der schwachen Stationarität; daher dürfen die Mittelwerte in der Autokovarianzfunktion nicht gleichgesetzt werden. Weitere Informationen über Periodizitäten können durch das Erstellen eines Periodogramms der Lastzeitreihe gewonnen werden.

Zur Ermittlung der Jahresperiode des Lastgangs wird der Lastverlauf über zwei Jahre für die Berechnung verwendet, wobei die Daten hierfür zuvor um den linearen Trend- und Mittelwertanteil $(x_t - \bar{x})$ bereinigt wurden (siehe Abbildung 3-6). Das Periodogramm ist als zweiseitiges Spektrum eines städtischen Kollektivs dargestellt. Die ursprünglichen Werte der Koeffizienten im 2. Quadranten können über die Addition mit der Länge T = 70.080 (für die Jahre 2001 - 2002) der Zeitreihe mit den negativen Koeffizienten rekonstruiert werden. Die Analyse der Zeitreihe zeigt deutliche „Peaks“ in bestimmten Bereichen, die es zu identifizieren gilt. Wie aus der Analyse der empirischen Momente bekannt ist, besitzen die Tage (mit jeweils 96 Messpunkten) einen periodischen Verlauf. Bei 96 Messungen pro Tag müsste die Schwingungskomponente $m = T/k = 96$ sein, wenn diese Amplitude den Einfluss der Tage auf die Zeitreihe repräsentiert.

Amplitudenwert	Fourierkoeffizient	m	Beschreibung
10.405e+006	730	96	Tag
1,5968e+005	104	672	Woche
15,69e+005	1460	48	halber Tag
13,009e+005	2	35040	Jahr

Tabelle 6-5: Identifizierung der Koeffizienten

Die Koeffizienten der Woche nehmen keinen ganzzahligen Wert an (siehe Tabelle 6-5). Dieses lässt sich durch die endliche Stichprobe der betrachteten Zeitreihe erklären da die Folge der Wochentage an den Rändern der Stichprobe nicht mehr gegeben ist. Eine erhöhte Auflösung zeigt auch, dass sich neben den sehr deutlichen Schwingungskomponenten wie Tag, Woche, Jahr und halber Tag weitere Peaks lokalisieren lassen (1/2 Woche, 1/3 Woche, …), mit den Koeffizienten $k \approx m \cdot k_{Woche} + 1;\ m = 2,3,4.....;$ und $k_{Woche} = 104$. Nähere Untersuchungen zeigten, dass es sich um Oberschwingungen der Woche handelt. Die Rekonstruktion der Zeitreihe für die Jahre 2001 und 2002 unter Verwendung der Informationen des Periodogramms, d. h. unter Berücksichtigung der Oberschwingungen, führt zu einer deutlichen Absenkung des MSE- Fehlers. In Tabelle 6-6 sind diese Zusammenhänge dargestellt. Die Koeffizienten 209 und 626 können als halbe Woche bzw. als 1/6 Woche interpretiert werden.

Fourierkoeffizienten	MSE
730, 104, 1460, 2	255,25 MW^2
730, 104, 1460, 2, 209	223,35 MW^2
730,104,1460,2,209,626	207,15 MW^2

Tabelle 6-6: MSE- Fehler bei der Rekonstruktion von zwei Jahren

6.2.4 Tagtypbestimmung

Der Wochenverlauf der Last, wie in Abbildung 6-1 und Abbildung 3-1 dargestellt, lässt im ersten Schritt vermuten, dass eine Woche in Bezug auf die Ähnlichkeit der Lastverläufe untereinander in Werktage, Samstage und Sonntage eingeteilt werden kann. Dass diese Einteilung für alle Wochen eines Jahres sinnvoll erscheint, verdeutlicht die Abbildung 6-3.

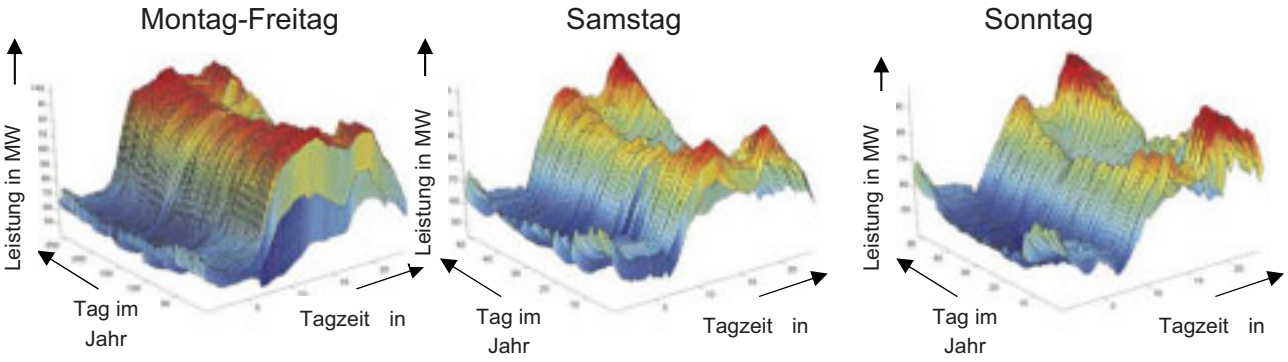


Abbildung 6-3: Tageslastgangverläufe nach Wochentaggruppen

Die Wochentagsgruppen weisen in Abhängigkeit der Jahreszeit verschiedene Tageslastverläufe auf. Im Folgenden wird daher die Einteilung in die Wochentagsgruppen mit rechnerischen Methoden belegt, wobei zuerst die Lastzeitreihen in einzelne Gruppen mit jeweils einem Wochentagstypus geteilt werden. Im folgenden Schritt werden diese Zeitreihen miteinander (kreuz)korreliert und lineare Abhängigkeiten detektiert. Anders als bei der Autokorrelation wird jedoch untersucht, ob lineare Zusammenhänge zwischen zwei verschiedenen Zeitreihen auftreten. Der Korrelationskoeffizient r_{xy} beschränkt sich auf einen Wertebereich von [1, -1], dabei bedeutet ein Wert von 1, dass beide Zeitreihen vollständig linear korrelieren, wohingegen -1 vollständige Antikorrelation bedeutet. Werte um 0 bedeuten, dass keine linearen Zusammenhänge bestehen. Je nach Anwendungsgebiet finden sich in der Literatur unterschiedliche Bewertungen des Wertebereiches in Bezug auf die Korrelationsstärke. Für die Bewertungen der Lastgangzeitreihen wurden die Definitionen von [Gna04] angewendet:

- |r| = 0 kein Zusammenhang
- |r| <= 0.2 sehr schwacher Zusammenhang
- 0.2 < |r| <= 0.5 schwacher Zusammenhang
- 0.5 < |r| <= 0.7 mittlerer Zusammenhang
- 0.7 < |r| <= 0.9 starker Zusammenhang
- 0.9 < |r| <= 1 sehr starker Zusammenhang

Ab welchem Grenzwert ein Korrelationskoeffizient als signifikant unterschiedlich von 0 angesehen werden kann, wurde mit Hilfe eines zweiseitigen t-Tests ermittelt, dessen Ablauf im Anhang 9.3 beschrieben wird. Auf der Grundlage der vorgestellten Methoden

wird die Kreuzkorrelation der einzelnen Wochentagszeitreihen durchgeführt. Dabei erfolgt eine Analyse aller Wochentage und die Gruppe der Wochentage wird um Feier- und Brückentage bereinigt. Die Tabelle 6-7 zeigt diese Ergebnisse.

Wochentage mit Feiertagen							
	Mo	Di	Mi	Do	Fr	Sa	So
Mo	1	0,9633	0,9581	0,9028	0,9519	0,8722	0,7245
Di	0,9633	1	0,9956	0,9353	0,9507	0,8908	0,7182
Mi	0,9581	0,9956	1	0,9348	0,9506	0,8877	0,7158
Do	0,9028	0,9953	0,9348	1	0,9474	0,8635	0,6889
Fr	0,9516	0,9507	0,9506	0,9474	1	0,8807	0,6802
Sa	0,8722	0,8908	0,8877	0,8635	0,8807	1	0,8585
So	0,7245	0,7182	0,7128	0,6889	0,6802	0,8585	1

Wochentage ohne Feiertage							
	Mo	Di	Mi	Do	Fr	Sa	So
Mo	1	0,9880	0,9888	0,9923	0,9865	0,8986	0,6558
Di	0,9880	1	0,9958	0,9912	0,9805	0,907	0,6406
Mi	0,9888	0,9958	1	0,993	0,9825	0,9042	0,6395
Do	0,9923	0,9912	0,993	1	0,987	0,9031	0,6462
Fr	0,9865	0,9805	0,9825	0,987	1	0,8938	0,6193
Sa	0,8986	0,9070	0,9042	0,9031	0,8938	1	0,8134
So	0,6558	0,6406	0,6395	0,6462	0,6193	0,8134	1

Tabelle 6-7: Korrelationsmatrix der Wochentage

Die Tabelle 6-8 zeigt für verschiedene Signifikanzniveaus die Intervalle, in denen der Korrelationskoeffizient liegen muss, damit die Nullhypothese H_0 angenommen wird.

Der Vergleich der in den beiden Korrelationsmatrizen enthalten Koeffizienten mit den zugehörigen Signifikanzgrenzen zeigt, dass die Korrelationskoeffizienten für jede Irrtumswahrscheinlichkeit α weit im Ablehnungsbereich der Nullhypothese H_0 liegen und damit als verschieden von 0 anzusehen sind. Ein Vergleich zwischen den Signifikanzgrenzen für die Zeitreihen mit und ohne Feiertage verdeutlicht, dass die Signifikanzintervalle größer werden, je kürzer die verglichenen Zeitreihen sind und je geringer die direkt damit verbundene Anzahl der Freiheitsgrade ist. Die Matrix der vollständigen Zeitreihen zeigt, dass Werktage sehr gut miteinander korrelieren. Trotz des allgemein hohen Korrelationsniveaus zwischen den Werktagen gibt es noch Abweichungen, am stärksten ausgeprägt beim Donnerstag. Auch der Samstag hat eine große Ähnlichkeit zu den Wochentagen. Der Sonntag korreliert am wenigsten mit den anderen Wochentagen. Die Ursache für den unterschiedlichen Grad an Korrelation zwischen den Werktagen und der Ähnlichkeit des Samstages mit ihnen sind die enthaltenen Feiertage in den einzelnen Zeitreihen. Feiertage weichen in ihrem Tagesverlauf stark von Werktagen ab. Dadurch wirken sich auf Werktage fallende Feiertage bei der Korrelation von Werktagszeitreihen negativ auf den Betrag des Korrelationskoeffizienten aus.

Wochentage mit Feiertagen - Freiheitsgrade : 4990								
α	0,2000	0,1000	0,0500	0,0100	0,0050	0,0010	0,0050	0,0010
Interval (+/-)	0,1810	0,0232	0,0277	0,0365	0,0397	0,0465	0,0492	0,0550
$t_{df;\,1-(\alpha/2)}$	1,2800	1,6400	1,9600	2,5800	2,8100	3,2900	3,4800	3,8900

Wochentage ohne Feiertage - Freiheitsgrade : 4414								
α	0,2000	0,1000	0,0500	0,0100	0,0050	0,0010	0,0050	0,0010
Interval (+/-)	0,0193	0,0247	0,0295	0,0388	0,0423	0,0495	0,0523	0,0585
$t_{df;\,1-(\alpha/2)}$	1,2800	1,6400	1,9600	2,5800	2,8100	3,2900	3,4800	3,8900

Tabelle 6-8: Signifikanzgrenzen zu den Koeffizienten der Korrelationsmatrix Tabelle 6-7

Die Korrelationsmatrix in der Tabelle 6-7 zeigt, dass die Grenzen zwischen den Gruppen ”Werktage“, ”Samstage“ und ” Sonntage“ durch Bereinigung um die Feiertage deutlicher wurden. Die Korrelationskoeffizienten zwischen den Werktagen liegen jetzt ausnahmslos auf einem hohen Niveau sehr nahe 1. Die Samstage sind hinsichtlich der Höhe der Korrelationskoeffizienten nach wie vor den Werktagen ähnlicher als den Sonntagen, aber dadurch, dass jetzt alle Werktage auf das gleiche Niveau gehoben wurden, wird die Einordnung der Samstage in eine eigene Gruppe sinnvoll. Die Abbildung 6-4 zeigt die Autokorrelationsfunktionen der drei Wochentags-gruppenzeitreihen. Die Einteilung in Wochentagsgruppen hat auch bewirkt, dass die Wochenperiode in den einzelnen Zeitreihen nicht mehr festzustellen ist (vergleiche hierzu Abbildung 3-2) sondern nur noch die Tagesperiode.

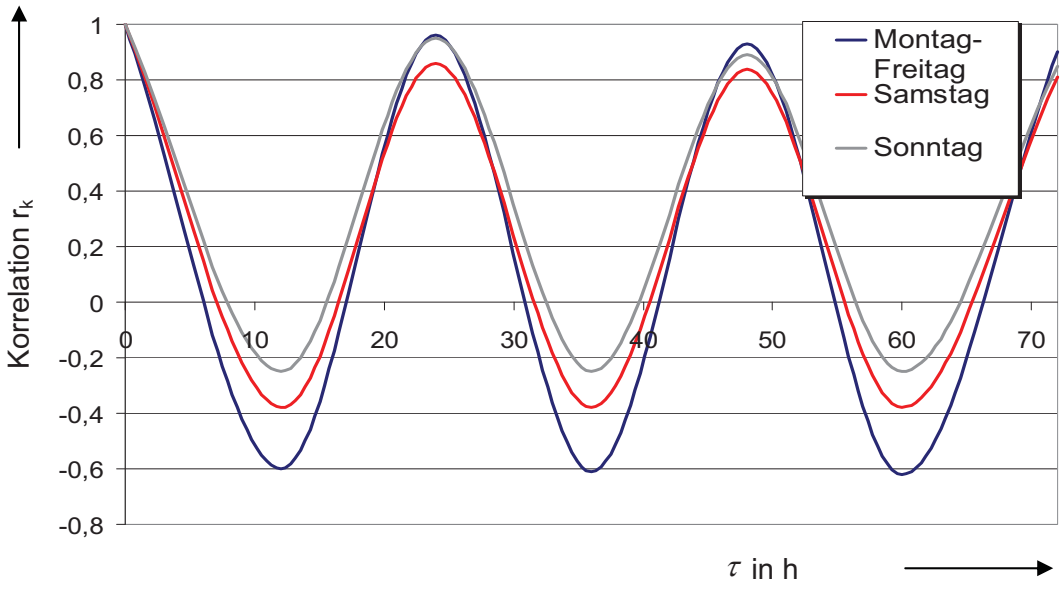


Abbildung 6-4: Autokorrelation der Wochentagtypen

Eine letzte Überprüfung wird an dieser Stelle noch mit Hilfe der vorgestellten Konfidenzintervalle durchgeführt. Sollten die Konfidenzintervalle der einzelnen Koeffizienten einer Gruppe so groß sein, dass sie sich mit den Konfidenzintervallen der

Korrelationskoeffizienten einer anderen Gruppe überschneiden, so wäre die Gruppeneinteilung in Frage zu stellen. Die Tabelle 6-9 zeigt die Ober- und Untergrenze jedes Korrelationskoeffizienten für die um die Feiertage bereinigten Wochentagszeitreihen. Diese zeigen, dass die Intervallgrenzen so dicht bei dem jeweiligen Korrelationskoeffizienten liegen, dass es zu keinen Überschneidungen kommt und die Gruppeneinteilung somit nicht in Frage gestellt werden muss.

Obergrenzen der Konfidenzintervalle							
	Mo	Di	Mi	Do	Fr	Sa	So
Mo	1	0,9887	0,9894	0,9927	0,9872	0,9041	0,6723
Di	0,9887	1	0,9960	0,9917	0,9816	0,9121	0,6577
Mi	0,9887	0,9960	1	0,9934	0,9835	0,9095	0,6566
Do	0,9894	0,9917	0,9934	1	0,9877	0,9084	0,6630
Fr	0,9927	0,9816	0,9835	0,9887	1	0,8995	0,6371
Sa	0,9041	0,9121	0,9095	0,9084	0,8995	1	0,8231
So	0,6723	0,6577	0,6566	0,6630	0,6371	0,8231	1

Untergrenzen der Konfidenzintervalle							
	Mo	Di	Mi	Do	Fr	Sa	So
Mo	1	0,9873	0,9881	0,9918	0,9856	0,8928	0,6387
Di	0,9873	1	0,9955	0,9907	0,9793	0,9016	0,6229
Mi	0,9881	0,9955	1	0,9926	0,9814	0,8987	0,6217
Do	0,9918	0,9907	0,9926	1	0,9862	0,8975	0,6287
Fr	0,9856	0,9793	0,9814	0,9862	1	0,8877	0,6008
Sa	0,8928	0,9016	0,8987	0,8975	0,8877	1	0,8032
So	0,6387	0,6227	0,6217	0,6287	0,6008	0,8032	1

Tabelle 6-9: Konfidenzintervallgrenzen zu den Koeffizienten der Korrelationsmatrix ohne Feiertage aus Tabelle 6-7

Eine weitere Möglichkeit, die Zeitreihen auf Gemeinsamkeiten und Unterschiede hin zu untersuchen, besteht darin, die empirischen Momente Mittelwert, Standardabweichung und Varianz zu ermitteln und zu vergleichen. Die Auswertung der empirischen Momente der Wochentagszeitreihen in Tabelle 6-10 bestätigt die schon durch die Kreuzkorrelation ermittelten Unterschiede zwischen den drei Gruppen „Werktage“, „Samstage“ und „Sonntage“. Die Werktage weisen dabei den größten Mittelwert auf und haben die größte Standardabweichung. Die Samstage liegen beim Vergleich der empirischen Momente im Gegensatz zu den Ergebnissen der Kreuzkorrelation näher an den Sonntagen als an den Werktagen. Die Sonntage bilden die Gruppe mit dem geringsten arithmetischen Mittelwert und der geringsten Standardabweichung. Wie auch schon bei der Kreuzkorrelation zuvor bewirkt ein Bereinigen der Wochentagsreihen um die Feiertage, dass die Gruppierungen deutlicher hervorgehoben werden.

Wochentage mit Feiertagen							
	Mo	Di	Mi	Do	Fr	Sa	So
$\bar{x}$ in MW	95,75	98,20	98,46	97,27	93,95	74,74	65,27
s in MW	27,16	25,36	25,04	25,29	23,29	14,32	12,88
s^2 in MW^2	737,41	643,31	627,21	639,70	567,89	205,12	166,00

Wochentage ohne Feiertage							
	Mo	Di	Mi	Do	Fr	Sa	So
$\bar{x}$ in MW	97,12	98,40	98,46	98,55	95,66	77,20	65,52
s in MW	26,70	25,36	25,04	24,85	23,57	14,28	13,00
s^2 in MW^2	713,12	643,31	627,21	617,47	555,43	203,93	168,91

Tabelle 6-10: Mittelwert $(\bar{x})$, Standardabweichung (s) und Varianz (s^2) der Wochentagstypzeitreihen

Die empirischen Momente der Feiertage werden in Tabelle 6-11 gezeigt. Mit Blick auf den arithmetischen Mittelwert ähneln sie den Sonntagen, die Standardabweichung liegt zwischen den Sonntagen und den Samstagen.

Feiertage	
$\bar{x}$ in MW	66,42
s in MW	13,89
s^2 in MW^2	190,34

Tabelle 6-11: Mittelwert $(\bar{x})$, Standardabweichung (s) und Varianz (s^2) der Feiertage

6.2.5 Wetterelemente

Im Folgenden werden die einzelnen Wetterelemente und deren Beziehungen zueinander analysiert. Detaillierte Untersuchungen hierzu, wie beispielsweise die Ermittlung der empirischen Momente, Trendpolynome und die Spektralanalyse, erfolgten im Rahmen der Diplomarbeit [Fre05]. Da nur die Daten über einen Zeitraum von drei Jahren zur Verfügung standen, lassen sich keine Aussagen treffen, ob die linearen Trendverläufe nicht auf Periodizitäten mit einer Periode größer als die des Betrachtungszeitraumes zurückzuführen sind. Zum Nachweis der Periodizitäten wurden die Periodogramme der Wetterelemente miteinander verglichen (siehe Anhang). Dabei fällt auf, dass die Beobachtungswerte der Bodentemperatur, der Temperatur in zwei Metern Höhe und die Globalstrahlung ein Frequenzspektrum besitzen, in dem es einige wenige ausgeprägte Frequenzen gibt. Die Frequenzspektren von Bedeckungsgrad, Windgeschwindigkeit, Windrichtung und Luftdruck weisen dagegen eine Vielzahl von Frequenzen mit fast gleichen Amplituden auf. Zwischen den Temperaturverläufen und der Globalstrahlung lassen sich beim Vergleich der größten enthaltenen Perioden weitere Gemeinsamkeiten finden, da diese Wetterelemente von einer Jahres- und Tagesperiode geprägt sind. Dabei lassen sich Unterschiede in der Reihenfolge der Perioden nachweisen. Während bei den Temperaturen die Jahresperiode die größere Amplitude aufweist, ist bei der Globalstrahlung die Tagesperiode gegenüber der Jahresperiode dominanter. Die Analyse der Wetterdaten auf Periodizität und Trendzyklen führte zu folgenden Ergebnissen:

- Die Verläufe von Bodentemperatur und Temperatur in zwei Metern Höhe weisen die größten Ähnlichkeiten auf, gefolgt von der Globalstrahlung.

– Die anderen Wetterelemente Bedeckungsgrad, Windrichtung und Luftdruck können bei der Betrachtung der Hauptschwingungen nicht gruppiert werden und müssen als relativ unabhängig voneinander betrachtet werden.

	TB	T2M	GSS	BD	WG	WR	LD
TB	1	0,9727	0,6158	-0,0070	0,1003	0,0541	-0,0771
T2M	0,9727	1	0,4856	-0,0543	0,0771	0,0301	-0,0851
GSS	0,6158	0,4856	1	-0,1022	0,0742	0,0263	0,0762
BD	-0,0070	-0,0543	-0,1022	1	0,1883	0,2384	-0,2698
WG	0,1003	0,0771	0,0742	0,1883	1	0,1922	-0,3222
WR	0,0541	0,0301	0,0263	0,2384	0,1922	1	-0,0824
LD	-0,0077	-0,0851	0,0762	-0,2698	-0,3222	-0,0824	1

Tabelle 6-12: Korrelation zwischen den Wetterelementen über einen Zeitraum von drei Jahren (Abkürzungen siehe Tabelle 6-2)

Die Korrelationsanalyse (Tabelle 6-12) zwischen den einzelnen Wetterelementen bestätigt im Wesentlichen die nach der Analyse der Periodizitäten aufgestellten Behauptungen. Am besten korrelieren die Verläufe der Bodentemperatur und der Temperatur in zwei Metern Höhe miteinander. Den nächst höheren Korrelationskoeffizienten erreicht die Globalstrahlung im Vergleich mit den genannten Größen. Zwischen Globalstrahlung und Bodentemperatur liegt der Korrelationskoeffizient bei 0,6158 und zwischen Globalstrahlung und Temperatur in zwei Metern Höhe bei 0,4856. Der Korrelationskoeffizient zwischen Windgeschwindigkeit und den Temperaturen oder der Globalstrahlung zeigt, dass kaum lineare Abhängigkeiten bestehen. Die Windgeschwindigkeit hat höhere Korrelationskoeffizienten zu anderen Wetterelementen wie zum Beispiel Luftdruck. Die verbleibenden Wetterelemente Bedeckungsgrad, Luftdruck und Windrichtung zeigen in Bezug auf die Temperaturen und die Globalstrahlung ebenfalls keine hohen linearen Abhängigkeiten. Betrachtet man weiterhin die Abhängigkeit zwischen diesen Elementen, finden sich keine weiteren Korrelationen, die betragsmäßig größer 0,3 liegen. Es gibt folglich nur schwache Zusammenhänge.

6.2.6 Zusammenhang zwischen Wettereinflussgrößen und Lastgang

Im Allgemeinen wird davon ausgegangen, dass exogene Größen wie Temperatur, Globalstrahlung, Bedeckungsgrad usw. einen erheblichen Einfluss auf das Verbrauchsverhalten eines Verbrauchskollektivs haben. Um zu entscheiden, wann und ob die Einbeziehung von Wetterdaten in die Prognosemodelle sinnvoll ist, muss untersucht werden, inwiefern ein signifikanter Zusammenhang zwischen der zu prognostizierenden Größe und einer oder mehrerer Wettercharakteristika besteht. Des Weiteren soll die Art des Kollektivs in Beziehung zu den exogenen Größen gesetzt werden. Ein erster Vergleich der Periodogramme des Lastverlaufes in Abbildung 3-6 mit den einzelnen Wetterelementen (Anhang) zeigt, dass das Frequenzspektrum des Lastverlaufes die größten Übereinstimmungen mit den Frequenzspektren der Temperaturen

und der Globalstrahlung aufweist. Der Lastverlauf wird analog zu diesen Wetterelementen auch nur von einigen wenigen Perioden wie der Jahres- und der Tagesperiode geprägt. Es ist davon auszugehen, dass der Lastverlauf von diesen Wetterelementen beeinflusst wird. Folgende Perioden des Lastverlaufes spiegeln die Wettereinflüsse wider; die Jahresperiode, da im Sommer weniger Last abgenommen wird als im Winter. Die Tagesperiode und die Halbtagesperiode, die Tageslastverläufe der Wochentagsgruppen weisen je nach Jahreszeit unterschiedlich stark ausgeprägte Nachmittags- und Abendspitzen auf. Die weitere Analyse erfolgte mittels Korrelationsbetrachtungen. Der Einfluss von Feiertagen und einigen besonderen Ereignissen auf den zu prognostizierenden Tageslastverlauf ist evident; der Einfluss von Wettergrößen mit statistischen Mitteln nur sehr schwer nachzuweisen. Am deutlichsten wird ein derartiger Einfluss bei den Lastdaten eines eher ländlichen Kollektivs, wie bereits andere Untersuchungen gezeigt haben. Dabei wurden sowohl unterschiedliche Untersuchungszeiträume (Jahr, Monat oder Tag) als auch unterschiedliche Tagtypen zu Grunde gelegt. Tabelle 6-13 enthält Korrelationswerte über einen Zeitraum von einem Jahr.

	TB	T2M	GSS	BD	WG	WR	LD
komplett	0,0795	-0,0126	0,3182	0,1308	0,1548	-0.0106	0,0085
Mo – Fr	0,1273	0,0149	0,4181	0,1671	0,1918	0,0486	0,0025
Sa	-0,0497	-0,1530	0,2853	0,2661	0,1607	0,1104	0,0403
So	-0,1639	-0,1959	0,0989	0,0697	0,0947	-0,0362	0,0425

Tabelle 6-13: Korrelationsbetrachtung zwischen Lastgang und Wetterelementen bei einer Signifikanzgrenze (α =0,05)

Wie sich bei der Betrachtung der Tabelle zeigt, sind diese Einflussgrößen in den meisten Fällen nur schwach mit dem Lastwert korreliert. Auffällig ist die insgesamt schlechte Korrelation der Temperaturen mit dem Lastgang. Begründen lässt sich dies mit den Hauptperioden der Temperaturverläufe und des Lastverlaufes, da die Verläufe der Zeitreihen sehr stark von einer Jahres- und einer Tagesperiode geprägt sind. Während die Tagesschwingung der Temperatur aber annähernd phasengleich mit der Last verläuft, verläuft die Jahresschwingung um circa 180° phasenverschoben. Dieser Umstand führt dazu, dass sich die positive Tageskorrelation gegen die negative Jahreskorrelation aufhebt und die Gesamtkorrelation geringer wird. Bei der Globalstrahlung wird dieser Effekt trotz gleicher Periodenverhältnisse zur Last nicht so deutlich, da bei ihr ebenso wie bei der Last die Tagesperiode gegenüber der Jahresperiode stärker ausgeprägt ist. Hieraus wird ersichtlich, dass es für Korrelationsanalysen entscheidend ist, zunächst die grundlegenden Eigenschaften der beteiligten Zeitreihen zu identifizieren. Des Weiteren muss vor der Analyse festgelegt werden, über welches Zeitintervall lineare Zusammenhänge ermittelt werden sollen. Je nach Festlegung müssen für den Fall, dass es bei den beteiligten Zeitreihen sich gegenseitig aufhebende Faktoren gibt, jene Faktoren eliminiert werden, die das Ergebnis der gewünschten Gemeinsamkeitsbetrachtung verfälschen. Dies bedeutet, dass die Daten-

reihen vor der Korrelationsanalyse um die jahreszeitlichen Einflüsse bereinigt werden müssen. In diesem Betrachtungsfall wurden die Daten um die Jahresperiodizität und den linearen Trend über das Jahr bereinigt. Abbildung 6-5 zeigt diesen Schritt am Beispiel des Lastverlaufes und der Bodentemperatur für die Werktagszeitreihen exemplarisch. Im Ergebnis ist zu sehen, dass die Temperatur und Last besser korrelieren, da der antikorrelierende jahreszeitliche Effekt eliminiert wurde. Eine hinreichend starke Korrelation und damit ein hinreichend starker linearer Zusammenhang zwischen Wettergrößen und dem Lastverbrauch liegen nur unter bestimmten Gegebenheiten vor.

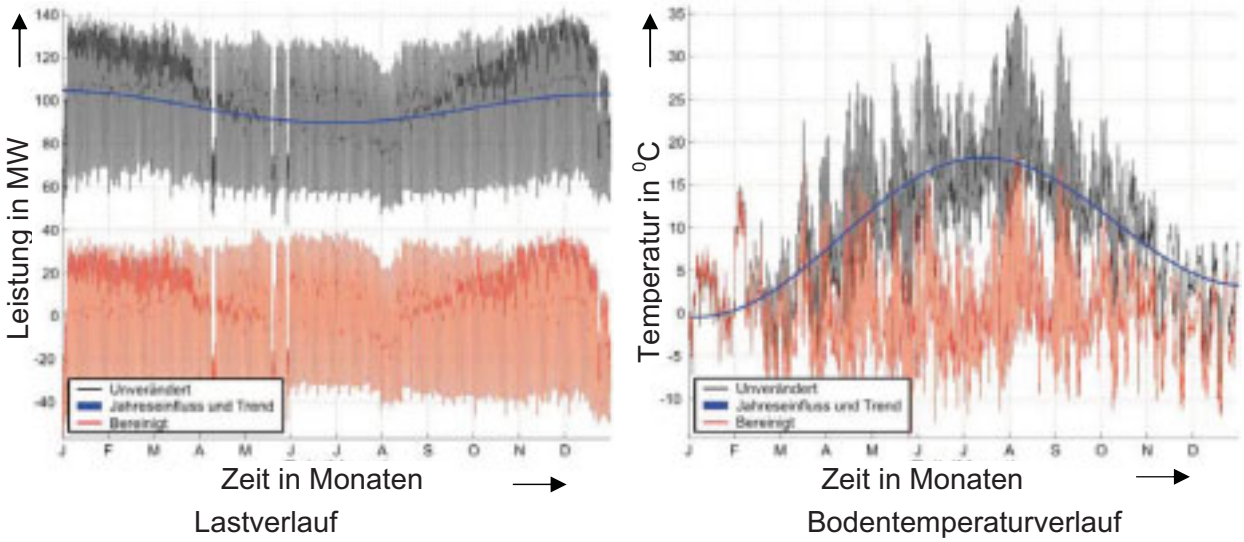


Abbildung 6-5: Jahreseinfluss in den Werktagszeitreihen

In der Tabelle 6-14 sind die Korrelationsergebnisse nach der vorgenannten Bereinigung für alle zur Verfügung stehenden Wetterelemente aufgeführt. Die Anwendung verschiedener Verfahren verdeutlicht, dass ein Großteil der Korrelation zwischen den hier betrachteten Wetterelementen und der Last in keinem signifikanten Zusammenhang steht. Die Integration exogener Größen in konventionelle Prognosemodelle für elektrische Last erscheint daher eher nicht sinnvoll. Eine Verbesserung der Ergebnisse könnte durch Anwendung der zeitverschobenen Korrelation erreicht werden.

	TB	T2M	GSS	BD	WG	WR	LD
komplett	0,3373	0,2295	0,4198	0,0999	0,1305	-0,0012	0,0067
Mo. – Fr.	0,4132	0,2734	0,5244	0,1411	0,1692	0,0581	-0,0071
Sa	0,3560	0,2230	0,4890	0,1856	0,0980	0,1508	0,0982
So	0,2209	0,2037	0,2824	-0,0013	0,0546	-0,0130	0,0733

Tabelle 6-14: Korrelationsbetrachtung zwischen Lastgang und Wetterelementen nach Bereinigung der Trend- und Jahreseinflüsse (α =0,05)

Die zeitverschobene Korrelation (Gleichung 6-1) zeigt die zeitlichen Zusammenhänge, die auch aus der Autokorrelation ersichtlich werden.

$$r(\tau) = \frac{c_{xy}}{s_x s_y} = \frac{\sum_{i=1}^{n-\tau} (x_i - \bar{x})(y_{i+\tau} - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}} \qquad (6\text{-}1)$$

Es werden jedoch zeitliche lineare Abhängigkeiten zwischen unterschiedlichen Zeitreihen (x_i, y_i) untersucht. Die Korrelationsfunktion steht in Abhängigkeit zur Größe der Zeitverschiebung τ und ist in Gleichung 6-1 dargestellt.

Die Abbildung 6-6 zeigt exemplarisch die Korrelationsfunktionen der Temperatur in zwei Metern Höhe und der Globalstrahlung mit der Last bis zu einer maximalen Zeitverschiebung von zwölf Stunden. Dabei wurden ausschließlich Zeitreihen mit Wochentagen, bereinigt um die Jahresperiode, verwendet. Bei der Lastzeitreihe wurden zusätzlich noch die Wochen- und Halbwochenperioden extrahiert. Die rote Kurve in den Abbildungen zeigt den Verlauf der Korrelationskoeffizienten bei einem Vergleich zwischen den Lastzeitreihenwerten x_i und den Wetterzeitreihenwerten $y_{(i+\tau)}$. Bei der blauen Kurve werden die Wetterelemente x_i mit den Lastzeitreihenwerten $y_{(i+\tau)}$ verglichen. Die schwarze Kurve zeigt zum Vergleich noch einmal die Autokorrelation der Last.

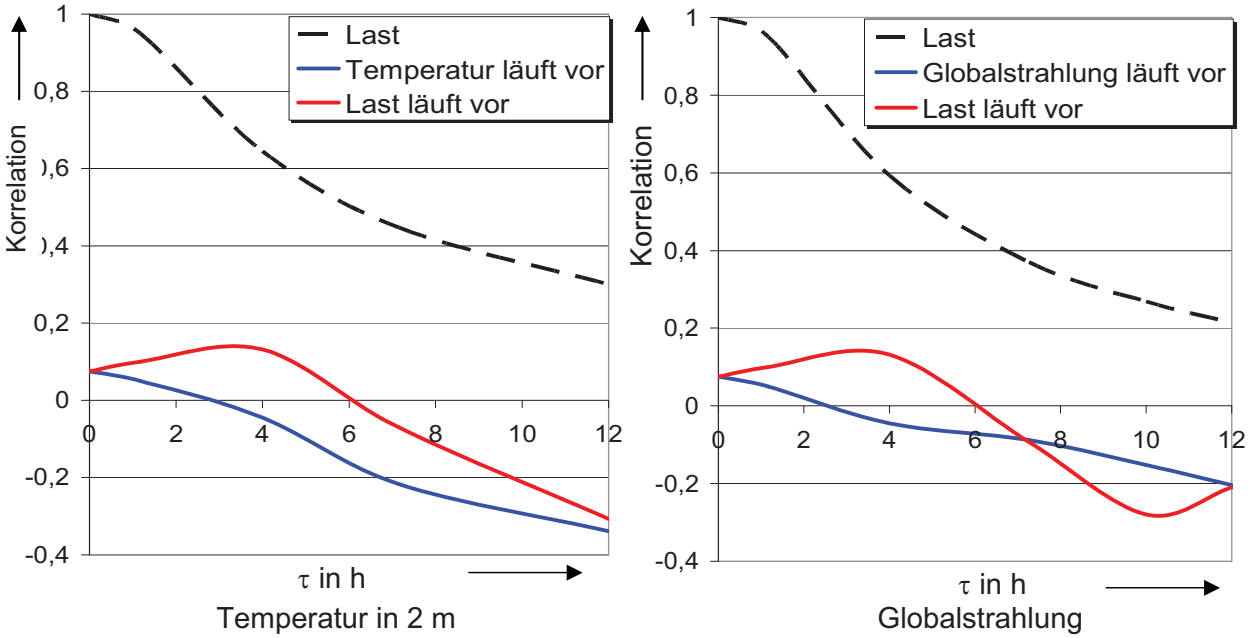


Abbildung 6-6: Korrelationsfunktion des Vergleichs zwischen Wetterelementen und elektrischer Last

Die ersten Maxima der roten Kurven und damit die größten positiven Korrelationen liegen bei einem τ von ca. drei Stunden bei der Globalstrahlung, und bei einem τ von ca. drei Stunden bei der Temperatur in zwei Metern Höhe. Dies bedeutet, dass die Wetterelemente der Last um zwei bzw. drei Stunden zeitlich vorauseilen. Das hier dargestellte Verhalten konnte bei allen Untersuchungsdaten, d. h. unabhängig vom Kollektiv, beobachtet werden. Betrachtet man die Korrelationen zu jeweils konstanten Zeitpunkten, ergibt sich für jede Uhrzeit der lineare Zusammenhang zwischen der exogenen Größe und dem Lastprofil. Abbildung 6-7 zeigt exemplarisch die Korrelation von Bodentemperatur, Temperatur in 2 Metern Höhe und Globalstrahlung. Wie aus den zuvor durchgeführten Untersuchungen ersichtlich, ist die Korrelation für den überwiegenden Teil des Tages eher schwach. Eine qualitative Verbesserung der Prognosemodelle wäre rein zufällig und lässt sich nicht in einen systematischen Zusammenhang mit den verwendeten exogenen Größen bringen.

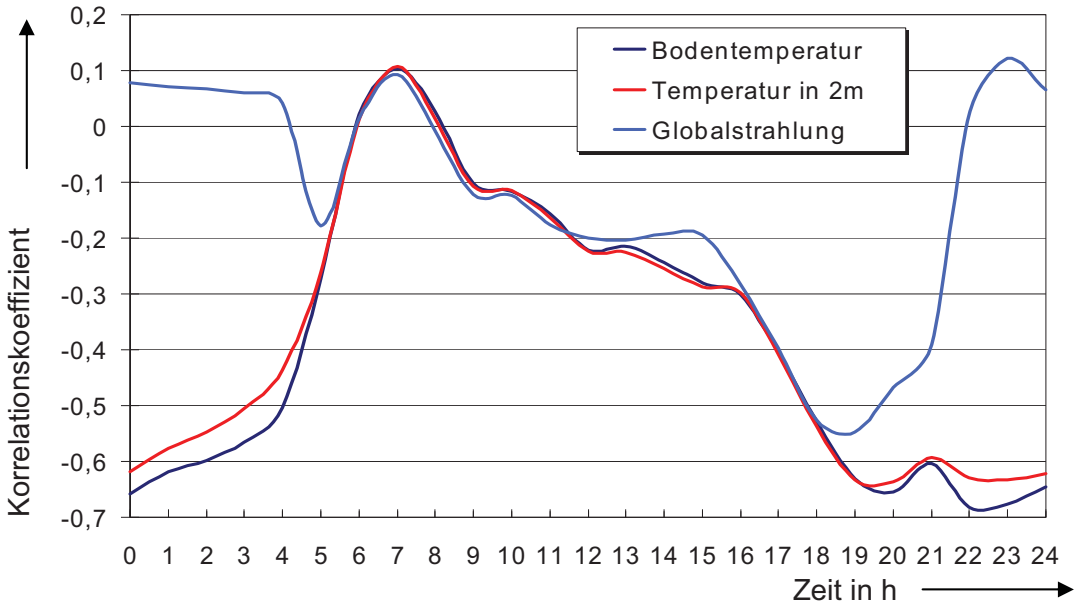


Abbildung 6-7: Korrelation zwischen Last und exogenen Größen an konstanten Zeitpunkten

In den Zeitintervallen (00:00 Uhr - 03:00 Uhr und 18:00 Uhr- 00:00 Uhr) eines Tages, die ein annähernd ausreichendes Korrelationsniveau erreichen, könnte die Integration von exogenen Größen eine Verbesserung der Prognose bewirken. Die untersuchten Lastprofile neigen in der Regel in den Mittagsstunden, also genau in den Zeitintervallen, die eine geringe Korrelation mit den exogenen Größen besitzen, zu eher zufälligem Verhalten. Diese Zeitintervalle sind aber gleichzeitig die Bereiche, in denen eine Verbesserung der bereits bestehenden Modelle energiewirtschaftlich nötig wäre, während sich die Morgen- und Abendstunden, in denen die Korrelation ein signifikantes Niveau erreicht und die Integration von exogenen Größen Erfolg versprechend ist, bereits mit relativ einfachen Mitteln (Trend & Saisonkomponente) sehr gut abbilden lassen. Die Zusammenfassung aller durchgeführten Untersuchungen zeigt, dass die Berücksichtigung von Wetterelementen bei der Lastprognose nur bedingt zu einer

Verbesserung der Ergebnisse führt. Deutlich wird aber auch, dass die Zusammenhänge zwischen Lastverlauf und Wettereinflüssen zuvor sehr genau analysiert und mit dem Prognosemodell abgestimmt werden müssen. Die hier durchgeführten Untersuchungen der Zusammenhänge zwischen Wetterelementen und elektrischer Last lassen sich daher wie folgt zusammenfassen:

- Die Bereinigung von Trendeinflüssen und jahreszeitlichen Einflüssen führt zur Verbesserung der Korrelation zwischen Temperatur, Globalstrahlung und Last.
- Eine Korrelation zwischen Lastgang und typischen Wettergrößen (Temperatur und Globalstrahlung) ist zwar schwach erkennbar, für eine Anwendung im Rahmen der Prognose der elektrischen Last jedoch nicht ausreichend.
- Durch die Verwendung zusätzlicher Wettergrößen wie Luftfeuchtigkeit, Windgeschwindigkeit usw. lassen sich ebenfalls keine positiven Effekte auf den Korrelationskoeffizienten und damit auf das Prognoseergebnis erzielen.
- Aus der Betrachtung der Korrelationen zu jeweils konstanten Zeitpunkten ergibt sich für jeden Viertelstundenwert ein linearer Zusammenhang zwischen den exogenen Größen und dem Lastprofil.

Bei Verwendung eines hierarchischen (komponentenbasierten) Prognosemodells [Kar06] ist der Einfluss exogener Größen bereits durch Komponenten wie Trend und Saison ausreichend abgebildet. Die Einbeziehung exogener Wettergrößen zur Vorhersage eines Lastganges mit hoher Abhängigkeit von Wettergrößen, wie zum Beispiel der des Erdgasbedarfs, ist hingegen aufgrund der stärkeren Korrelation sinnvoll und kommt zur Anwendung.

6.3 Analyse von Einflüssen auf Tagtypen mit Hilfe SOM

Die zur Verfügung stehenden Merkmale zur Beschreibung eines Tages entsprechen den in Kapitel 6.1 erfassten Eigenschaften. Die maximale Anzahl von Variationsmöglichkeiten der Merkmale ist aufgrund regionaler Unterschiede der Kollektive möglich, aber auch auf die Lage der Fest- und Feiertage im Jahr zurückzuführen. Die Tabelle 6-15 zeigt die im jeweiligen Jahr tatsächlich aufgetretene Anzahl der Variationsmöglichkeiten der aufgezeichneten Merkmale des Tages.

Jahr	Anzahl von Variationsmöglichkeiten
1998	182
1999	172
2000	195
2001	194
2002	185

Tabelle 6-15: Anzahl der Merkmalsvariationen in Abhängigkeit des Kalenderjahres

Um einen Ergebnisvergleich zwischen den unterschiedlichen Kollektiven zu erhalten, wurden alle SOM-Karten vor dem Lernvorgang mit einheitlichen, zufällig initiierten Startvektoren beaufschlagt. Für die Untersuchungen wurden zweidimensionale Karten der Größe 15x15 (225 Neuronen) gewählt.

Mit größeren Anordnungen hat sich gezeigt, dass mit hoher Iterationszahl sich jeder Tag auf einem einzelnen Neuron abbildet. Karten mit kleineren Dimensionen lassen jedoch keine hinreichend genaue Clusterung zu [Wel05] [Kuh04].Die Festlegung des Startwertes $\sigma=\sigma$ (1) orientiert sich stark an der Größe der Karte und wurde für die Untersuchungen auf fünf Längeneinheiten festgelegt. Mit diesem Wert wurde gewährleistet, dass am Anfang der Lernphase ein Großteil der Karte sich im Radius um das Gewinnerneuron befindet. Startwerte größer als fünf Längeneinheiten sind aufgrund der Kartendiagonale von ca. 20 Längeneinheiten ungeeignet, zu kleine σ-Werte sorgen für Teilclusterbildung. Da 365 verschiedene Datensätze eines Jahres zur Verfügung standen und jeder Datensatz mehrfach dem Netz präsentiert werden sollte, wurde die Anzahl der Iterationsschritte auf 12.000 festgelegt. Die im Folgenden dargestellten Ergebnisse zeigen exemplarisch die motorische Karte der unterschiedlichen Kollektive. Die verschiedenen Farben stehen für den jeweiligen Wochentag, der von der motorischen Karte zum jeweiligen Lastgang geclustert wurde.

Der Outputlayer des ländlichen Versorgungskollektivs ist in Abbildung 6-8 dargestellt. Es ist zu erkennen, dass die Sonntage, Samstage und Freitage relativ große Cluster im oberen Kartenrand bilden. Der überwiegende Anteil der Freitage bildet sich im rechten oberen Kartenrand der motorischen Karte ab und befindet sich ebenfalls innerhalb eines Datencluster auf der Karte.

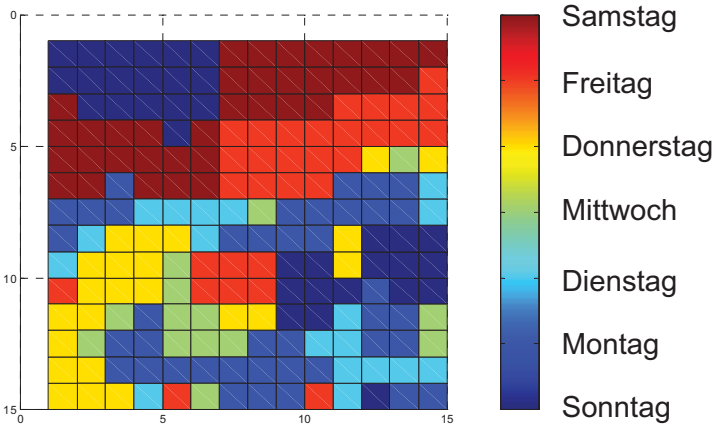


Abbildung 6-8: Outputlayer eines ländlichen Versorgungskollektivs

Auf der Karte des Mischgebietes (Abbildung 6-9) ist eine ähnliche Clusterung der Freitage, Samstage und Sonntage zu beobachten. Der Tagtyp Sonntag ist im Vergleich zum ländlichen Versorgungskollektiv weniger zwischen den normalen Werktagen zu

erkennen, sondern zeichnet sich deutlich in einem Cluster ab. Wie auch im rein ländlichen Kollektiv zu beobachten ist, lässt sich durch weitere Filterung der Samstage eine jahreszeitliche Unterteilung der Samstagscluster (April - September und Oktober - März) feststellen.

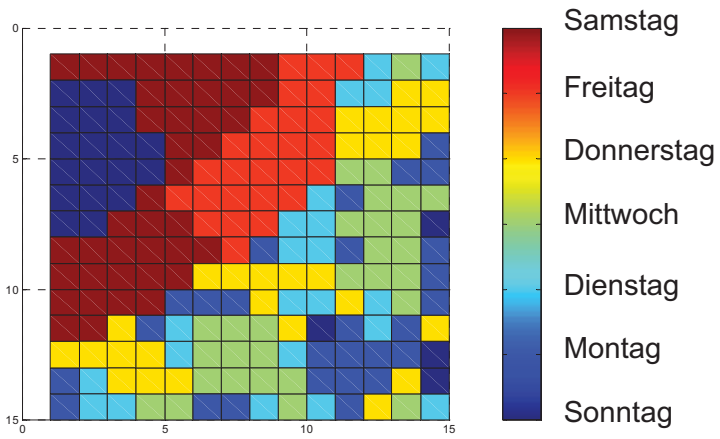


Abbildung 6-9: Outputlayer in städtischem und ländlichem Gebiet

In Abbildung 6-10 ist der Outputlayer eines rein städtischen Versorgungskollektivs dargestellt. Die beiden Wochenendtage sind hierbei in einem Cluster im unteren Kartenrand zusammengefasst, die sich durch Zuhilfenahme weiterer Filterfunktionen jahreszeitlich unterteilen lassen. Genau wie bei beiden anderen Beispielen hat sich ein relativ großer Freitagscluster gebildet. Es liegen auch hier fast alle Freitage der Monate April bis September in einem Kartenbereich.

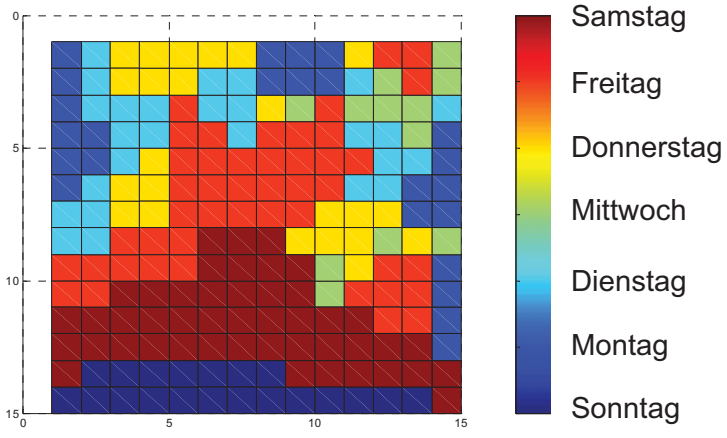


Abbildung 6-10: Outputlayer eines städtischen Versorgungskollektivs

Die Untersuchungen zeigen, dass in Abhängigkeit des Verbraucherkollektives sich unterschiedliche Datenclusterungen mit Hilfe der SOM detektieren lassen. Die Wochentage bilden, mit Ausnahme des Tagtyps Freitag, kein einheitliches Datencluster, was den Rückschluss erlaubt, die Wochentage Montag bis Donnerstag als einen einzigen Tagtyp im Rahmen von Lastprognosen zu behandeln. Des Weiteren werden Wochen-

feiertage, die auf die Tage Montag bis Donnerstag fallen, dem Datencluster des Sonntages zugeordnet. Ein Einfluss lokaler wie auch globaler Ereignisse innerhalb der Wochentage Montag bis Donnerstag konnte nicht nachgewiesen werden. Weiterhin ließ sich der Einfluss von Schulferien im Rahmen der Clusteranalyse nicht nachweisen.

Die hier exemplarisch aufgeführte Darstellung der Ergebnisse des Outputlayers ist aufgrund der zufällig initiierten Gewichtsvektoren der Karten nicht wiederholbar, sodass die Lage der Cluster auf der Karte selbst zufällig ist. Nur die Lage der Neuronen relativ zueinander ist entscheidend und führt bei wiederholter Durchführung der Untersuchungen zu gleichen Interpretationen. Das entwickelte Modell, das mit Hilfe eines neuronalen Netzes (selbstorganisierende Karten/SOM) ein Verfahren zur Verfügung stellt, aus gegebenen Lastdatensätzen der Vergangenheit Unregelmäßigkeiten im Lastgang selbst zu detektieren, zeigt auch in der praktischen Anwendung verwendbare Resultate. Dabei wurde die Eigenschaft der Kohonenkarte genutzt, ähnlich gelagerte Lastgänge auf nah beieinander liegende Neuronen abzubilden. Das neuronale Netz wurde mit den Tageslastgängen angelernt. Die Ergebnisse lassen sich wie folgt zusammenfassen:

Der Freitag zeigt, unabhängig vom Lastkollektiv, ein eindeutiges Cluster, das in Abhängigkeit des untersuchten Verbrauchkollektivs jahreszeitlich weiter unterteilt werden kann. Hierbei empfiehlt es sich, Freitage in zwei weitere Datencluster, mittlere Jahreshälfte von April bis September und den verbleibenden Zeitraum, separat im Rahmen von Lastprognosen zu behandeln. Freitage, die einem Feiertag folgen, sogenannte Brückentage, werden im Rahmen der Datenanalyse mit Hilfe der SOM dem Samstagscluster und Feiertage die auf den Wochentag „Freitag“ fallen, dem Sonntagcluster zugeordnet.

Die Samstage und Sonntage des Wochenendes bilden jeweils eigene Klassen. Auch hierbei erscheint in Abhängigkeit des Verbrauchskollektives eine weitere Unterteilung in Teilcluster sinnvoll. Für den Tagtyp Samstag bilden sich unabhängig des Netzgebietes jahreszeitliche Teilcluster aus. Eine weitere Unterteilung der Lastgangdaten der Netzgebiete in die Teilmengen April bis September sowie Oktober bis März lässt sich mit Hilfe der SOM-Analyse nachweisen. Lokale Ereignisse, wie beispielsweise Volksfeste und Großveranstaltungen, konnten nicht detektiert werden. Der Einfluss globaler Ereignisse wird mit Hilfe der Analyse ebenfalls nicht aufgezeigt und erlaubt keine weitere Clusterung. Sonntage bilden, unabhängig vom Kollektiv ein eigenes Datencluster, eine jahreszeitliche Unterteilung ist nicht sinnvoll. Eine Zusammenfassung aller Ergebnisse ist in Tabelle 6-16 dargestellt.

In einigen der bereits heute im Einsatz befindlichen Prognosesysteme erfolgt eine Unterteilung der Tagtypen in die Klassen Sonn- und Feiertage sowie Wochentage. In verschiedenen Stellen der Literatur [Huf00], [Als95] wird ebenfalls eine ähnliche Differenzierung empfohlen. Die Untersuchungen mit dem hier entwickelten Modell

haben jedoch gezeigt, dass in Abhängigkeit des Versorgungskollektivs eine weitere für den Anwender verbesserte Unterteilung der Tagtypen erzielt werden kann. Der Einfluss lokaler und globaler Ereignisse führt mit dieser Untersuchungsmethode nicht zur Bildung eigener Tagtypen.

Wochentag	Tagtyp	ländliches Netzgebiet	Mischung aus ländlichem und städtischem Netzgebiet	städtisches Netzgebiet
Montag-Donnerstag	Arbeitstag	Montage-Donnerstage		
	Wochenfeiertag	Sonntage		
Freitag	Arbeitstag	Freitage der Monate April-September		Freitage
	Arbeitstag	Freitage der Monate Oktober-März		
	Brückentag	Samstage		
	Wochenfeiertag	Sonntage		
Samstag	Normaltag	Samstage der Monate April-September		Samstage
	Normaltag	Samstage der Monate Oktober-März		
Sonntag	Normaltag	Sonntage		

Tabelle 6-16: Klassifizierung der Wochentage in Abhängigkeit unterschiedlicher Netzgebiete

6.4 Prognosen

6.4.1 Prognose mit klassischen Modellen

Nach der Parametrisierung der Trend- und zyklischen Komponenten kann die Zeitreihe durch die Verbindung der Komponenten approximiert werden, wobei vorausgesetzt wird, dass das Modell auch für die Zukunft Gültigkeit besitzt. Die dann folgende Prognose von Werten erfolgt durch die Verwendung eines dem Prognosezeitraum entsprechenden Zeitbereiches, für den die Komponenten die Zeitreihe extrapolieren. Dabei wird die Zeitreihe wie folgt beschrieben:

$x_t = f(t) = g_t + z_t + e_t$ mit den Elementen

$$g_t = \sum_{i=1}^{p} a_i \cdot t^i \text{ , glatte Komponente; } p \text{ =Polynomgrad} \qquad (6\text{-}2)$$

$$z_t = \sum_{k}^{q} \left[a_k \cdot \cos(2\pi \cdot f_k t) + b_k \cdot \sin(2\pi \cdot f_k t) \right] \text{ ; } f_k \text{ = ausgewählte Frequenzen}$$

e_t = Fehler

Als Maßeinheit für die Anpassung des Komponentenmodells an die Zeitreihe wird als Gütekriterium der MSE-Fehler, MRE-Fehler und $RMSE$-Fehler verwendet. Als Trendfunktion wird ein Polynom vom Grad eins verwendet. Aus der FFT werden jeweils die größten Amplituden im Modell verwendet. Der prognostizierte Zeitraum wird auf zwei

Tage (192 Vorhersagewerte) gemäß den in Kapitel 2.5 formulierten Anforderungen gesetzt. Als Datenbasis werden die Lastgangdaten aus zwei zurückliegenden Jahren, zunächst bereinigt um die Wochenenden und Feiertage gemäß den Ergebnissen der SOM-Analyse, verwendet. Benötigt wird die Prognose eines Tages und zur Berechnung des Fehlers müssen 96 Werte erfasst werden. Die Fehler werden für verschiedene Parametrisierungen der Amplitude untersucht, wobei die Größe der Stichprobe konstant ist. Die Umsetzung erfolgt dabei nach folgender Vorgehensweise: Die Zeitreihe wird von einem Zeitpunkt „Startpunkt“ an das Modell übergeben, die Anzahl der betrachteten Vergangenheitswerte ist von hieraus der Spalte „Modelldaten“ zu entnehmen. Die parametrisierten Komponenten extrapolieren den Verlauf der Zeitreihe. Das Modell wird dann mit den Modelldaten der Zeitreihe „Modellanpassung“ und mit den zu prognostizierenden Tagen „Prognosegüte“ verglichen und die Fehler bestimmt. Die Prognoseergebnisse mit den jeweiligen Parametern gemäß Tabelle 6-17 zeigen, dass der kleinste Prognosefehler mit durchschnittlich neun Amplituden (den neun größten Amplituden) erzielt wurde. Mit Hilfe zusätzlicher Untersuchungen wurde dann die Datenbasis verfeinert:

Tage		Modellanpassung		Prognosegüte			
Startpunkt k*96	Modelldaten k*96	MSE MW^2	RMSE MW	MSE MW^2	RMSE MW	MRE %	Amplituden
320	254	33,63	5,80	12,89	3,59	4,04	6
320	**254**	**29,36**	**5,42**	**8,08**	**2,84**	**3,11**	**9**
320	254	26,27	5,12	7,03	2,65	2,97	12
340	254	32,81	5,73	11,30	3,36	4,17	6
340	254	27,91	5,28	13,73	3,70	5,15	9
340	254	25,11	5,01	12,58	3,55	4,88	12
360	254	32,18	5,67	11,84	3,44	3,59	6
360	254	27,69	5,26	10,00	3,16	3,96	9
360	254	24,84	4,98	7,85	2,80	3,46	12
380	254	32,04	5,66	6,58	2,56	3,61	6
380	254	27,35	5,23	6,87	2,62	3,46	9
380	254	24,80	4,98	8,03	2,83	3,90	12
400	254	31,93	5,65	15,93	3,99	5,56	6
400	254	27,83	5,27	11,09	3,33	4,68	9
400	254	25,14	5,01	10,84	3,29	4,90	12
420	254	32,56	5,71	13,74	3,71	4,98	6
420	254	28,08	5,30	11,48	3,39	4.10	9
420	254	25,43	5,04	10,67	3,27	4,30	12
440	254	32,30	5,68	20,29	4,50	4,36	6
440	254	27,75	5,27	19,03	4,36	5,29	9
440	254	25,12	5,01	17,96	4,24	5,00	12

Tabelle 6-17: Parametrisierungen des klassischen Modells für Wochentage

Die Ergebnisse der Tabelle 6-18 zeigen einen MRE-Prognosefehler von 3,25 %. Abbildung 6-11 zeigt exemplarisch den Vergleich zwischen Prognoseergebnis und Originaldaten.

Tage		Modellanpassung		Prognosegüte			
Startpunkt k*96	Modelldaten k*96	MSE MW^2	RMSE MW	MSE MW^2	RMSE MW	MRE %	Ampli-tuden
260	254	27,19	5,21	12,09	3,48	3,33	9
290	254	28,47	5,34	24,69	4,97	5,20	9
330	**254**	**28,40**	**5,33**	**7,502**	**2,74**	**3,25**	**9**
350	254	28,09	5,30	15,08	3,88	5,31	9
460	254	27,61	5,25	12,92	3,59	4,02	9

Tabelle 6-18: Parametrisierungen des klassischen Modells für Wochentage

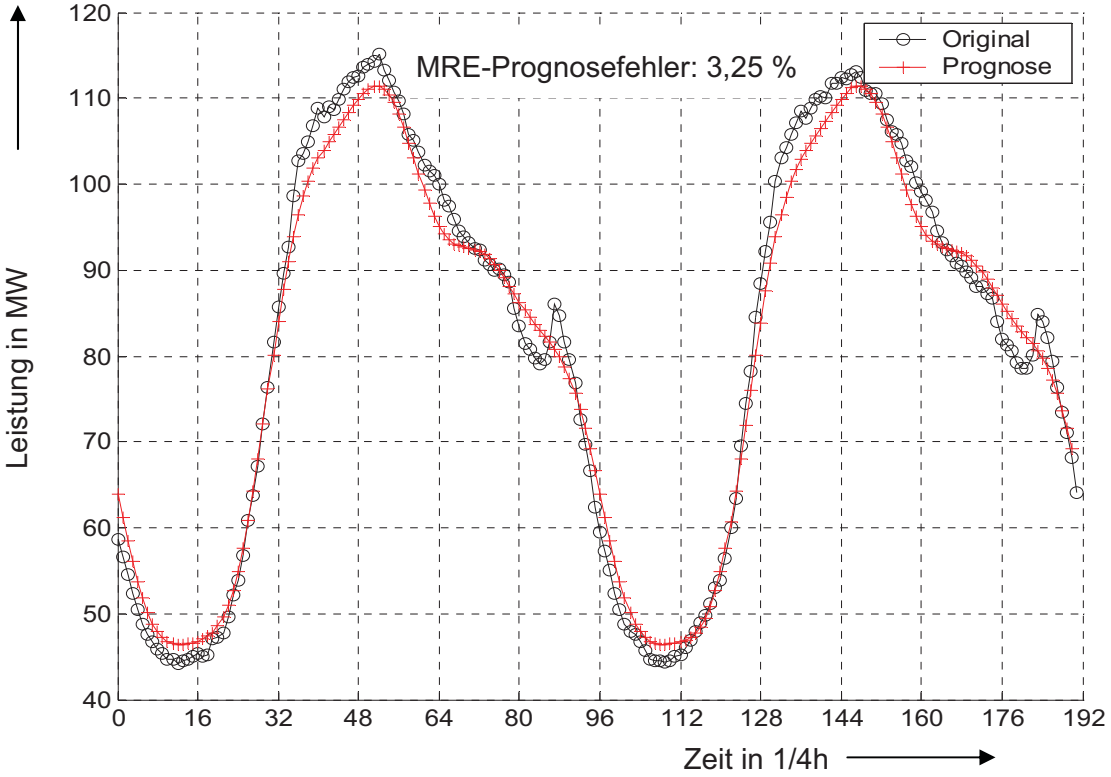


Abbildung 6-11: Prognose eines Wochentages bereinigt um Feiertage

Es besteht die Möglichkeit, die Parameter weiter zu optimieren, z. B. durch Reduktion des betrachteten Zeitintervalls der Vergangenheitswerte oder durch Verwendung anderer Trendpolynome. Selbst bei einer sehr einfachen Parametrisierung konnte damit gezeigt werden, dass die klassischen Modelle bereits hinreichend gute Prognoseergebnisse liefern. Abschließend erfolgt die Untersuchung der Wochenenden (Tabelle 6 -19) nach dem gleichen vorgenannten Schema, dabei wurden die Daten ebenfalls um Feiertage bereinigt.

Tage		Modellanpassung		Prognosegüte			
Startpunkt k*96	Modelldaten k*96	MSE MW^2	RMSE MW	MSE MW^2	RMSE MW	MRE %	Ampli-tuden
50	6	2,26	1,50	5,99	2,45	4,19	9
70	**6**	**1,73**	**1,32**	**7,97**	**2,82**	**3,77**	**9**
90	6	4,66	2,16	5,28	2,30	2,94	9
130	6	8,49	2,91	8,58	2,93	3,97	9
150	6	1,92	1,38	2,71	1,65	3,03	9
170	6	2,38	1,54	2,80	1,67	2,38	9

Tabelle 6-19: Prognose der Wochenenden

Der minimale MRE-Prognosefehler der Wochenenden beträgt 3,77 % und ist exemplarisch für ein Wochenende in Abbildung 6-12 dargestellt.

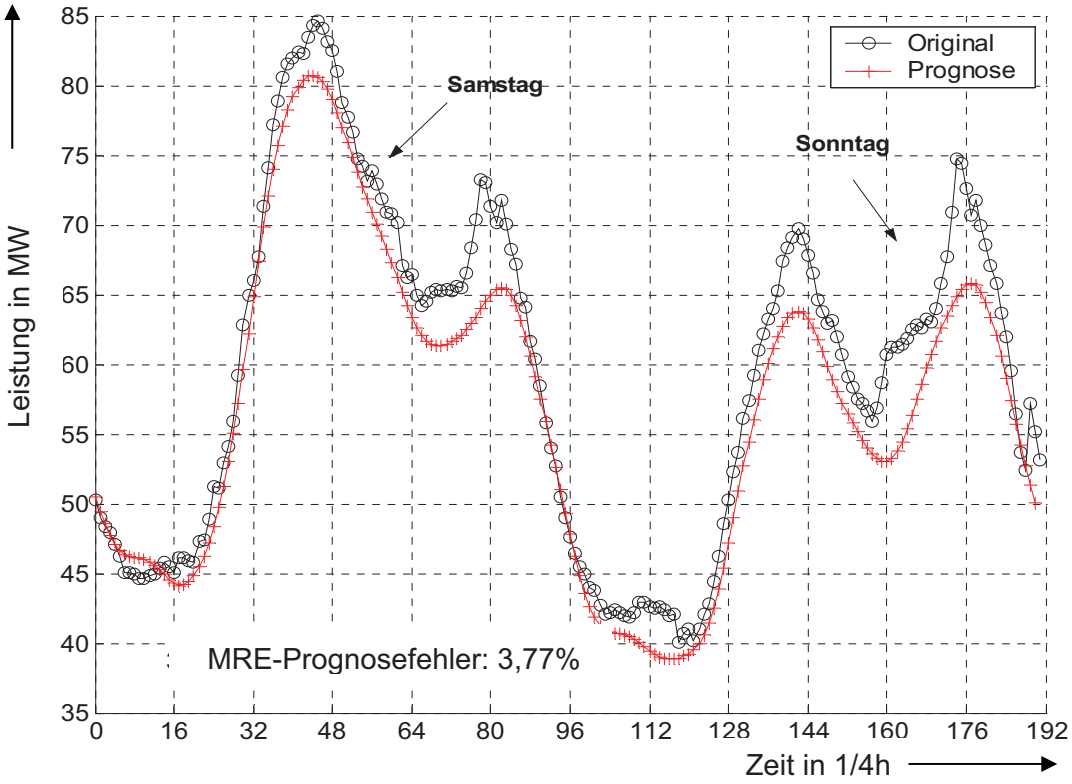


Abbildung 6-12: Prognose der Wochenenden

Die Parametrisierungen entsprechen denen der Werktage, die Anzahl der einbezogenen Vergangenheitswerte wurde jedoch reduziert. Anzumerken ist, dass durch die Variation von Amplitudenanzahl und Größe des Intervalls der betrachteten Vergangenheitswerte eine eher interpolative oder approximative Anpassung an die Originaldaten erreicht wird. Diese Eigenschaft ist mit dem Over- oder Underfitting bei neuronalen Netzen vergleichbar. Die Ergebnisse lassen sich wie folgt zusammenfassen:

- Die Zeitreihe unterliegt hauptsächlich Schwingungen, die sich als Tag, halber Tag, Woche und Jahr charakterisieren lassen.
- Für eine gute Prognose ist die Verwendung von mehr Amplituden notwendig als im Periodogramm charakterisiert werden, da die betrachtete Datenmenge der Prognose nicht äquivalent zu der des Periodogramms ist.
- Für gute Prognosen sollten die Vergangenheitswerte auf kurze Intervalle beschränkt bleiben.
- Die Aufteilung der Daten in Wochenende und Werktag sowie die Bereinigung um Feiertage verbessert die Aussagekraft der Spektren der untergeordneten Zeitreihen im Vergleich zur Ursprungszeitreihe.
- Die klassischen Modelle erlauben eine relativ gute Prognose für Werk- und Wochentage, selbst bei naiver Auswahl der Parameter.
- Die Güte der Prognose hängt stark von den zu verarbeitenden Daten ab, insbesondere von der Anzahl der Vergangenheitswerte und Amplituden.
- Die Anpassung des Modells an die Originaldaten ist kein zuverlässiges Kriterium für eine optimale Prognose. -Ein kleinerer Fehler in der Anpassung eines Modells ist keine Garantie für eine bessere Prognose im Vergleich zu einer anderen Parametrisierung.- Die Minimierung des Anpassungsfehlers ist jedoch ein guter Richtwert (Tabelle 6-17).

6.4.2 Prognose durch Kombination von deterministischen und stochastischen Ansätzen

Als Datenbasis der Untersuchungen dienen die um Feiertage bereinigten Jahre gemäß Kapitel 6.1, differenziert nach Werktagen und Wochenenden. Es werden im Folgenden mit AR-, $ARMA$-, $ARIMA$-, bzw. PAR-Modellen mit verschiedenen Konfigurationen und auch unter Einbeziehung exogener Größen Prognosen durchgeführt. Ziel ist die Festsetzung eines MSE-Fehlers pro Modell. Das erstellte Prognosemodell ist dabei eine Kombination von deterministischen und stochastischen Ansätzen (Kapitel 5.1.1). Ein Trendpolynom vom Grad vier, in Verbindung mit den 12 größten Amplituden, fungiert als untergelagertes deterministisches Modell. Die Residuen werden als stochastischer Prozess aufgefasst und mit Hilfe der Prozessmodelle verarbeitet. Die Güte Q (Kapitel 5.1.2) der Anpassung und der Prognose werden über die vorgenannten Fehlergrößen gemessen. Die Extrapolation künftiger Werte mit Hilfe der stochastischen Prozesse geht direkt aus ihren Definitionsgleichungen hervor. Das Vorgehen ist prinzipiell identisch mit dem des deterministischen Modells, hier wird jedoch auf Basis der bekannten und prognostizierten Parameter gearbeitet.

Die Variationen der unterschiedlichen p-Parameter des $AR(p)$-Prognosemodells sind exemplarisch in Tabelle 6-20 dargestellt, dabei beschreibt die Spalte „Startpunkt“ den Punkt, von dem aus die Anzahl der „Modelldaten“ verwendet werden. Die Prognose mit

dem geringsten Fehler wird mit jeweils 98 p-Parametern erreicht, wobei die Ergebnisse der Modelldiagnose ebenfalls bestätigt wurden. Exemplarisch ist in Abbildung 6-13 das Prognoseergebnis zweier Wochentage im Vergleich zu den Originaldaten dargestellt.

Tage		Anzahl	Prognosegüte			Modellanpassung	
Startpunkt k*96	Modelldaten k*96	p-Parameter	RMSE MW	MSE MW^2	MRE %	RMSE MW	Q
100	90	4	3,08	9,47	3,18	0,35	0,0105
100	90	8	3,09	9,55	3,22	0,36	0,0112
100	90	64	3,18	10,10	3,42	0,45	0,0171
100	90	96	2,92	8,53	3,32	0,53	0,0237
100	**90**	**98**	**2,31**	**5,32**	**2,36**	**0,67**	**0,0371**
140	90	4	3,16	9,99	3,08	0,30	0,0095
140	90	8	3,16	9,99	3,09	0,32	0,0103
140	90	64	3,15	9,92	3,06	0,36	0,0134
140	90	96	2,76	7,62	2,80	0,42	0,0183
140	**90**	**98**	**2,66**	**7,08**	**2,67**	**0,55**	**0,0304**
180	90	4	3,27	10,72	4,03	0,27	0,0154
180	90	8	3,27	10,70	3,97	0,29	0,0183
180	90	64	3,16	10,00	3,79	0,34	0,0249
180	90	96	2,41	5,83	2,76	0,42	0,0369

Tabelle 6-20: AR -Prognosemodell verschiedener Ordnung für Wochentage

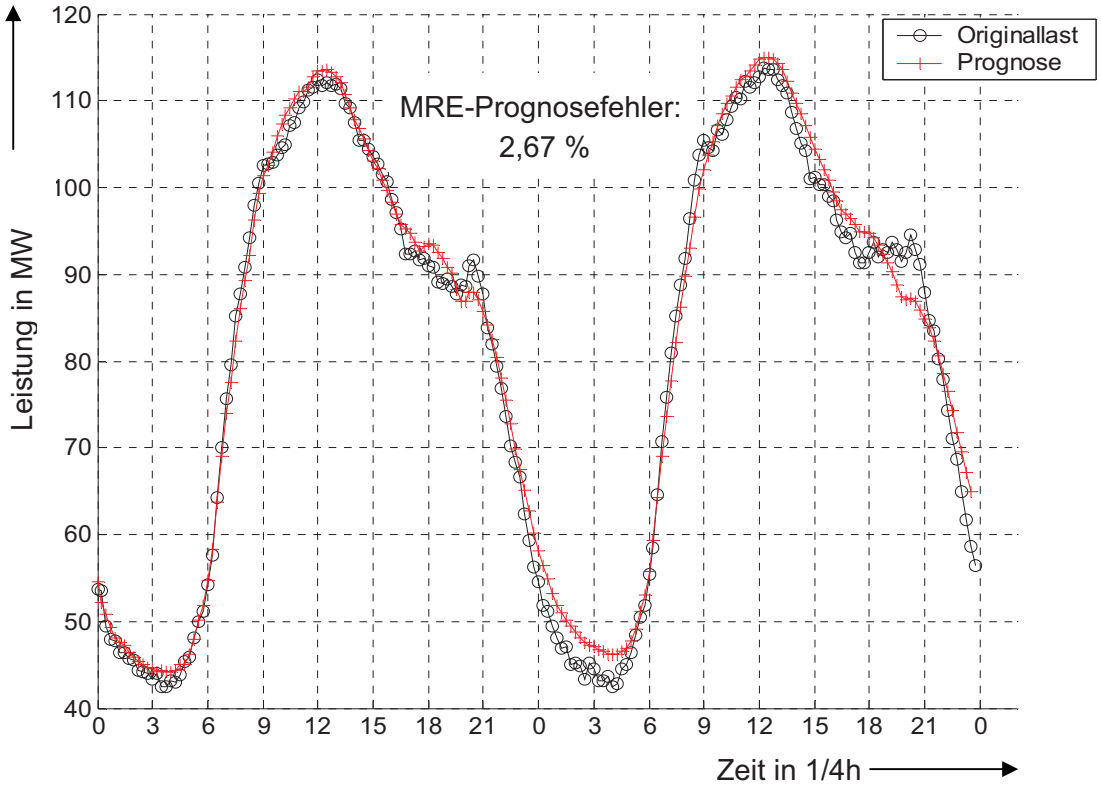


Abbildung 6-13: Ergebnis des AR -Prognosemodells für Wochentage

Für weitere Testmessungen mit den optimal dimensionierten Parametern konnte ein durchschnittlicher MSE-Prognosefehler von 2,58 % ermittelt werden. Die Ergebnisse können durch eine Vergrößerung oder Verringerung der Anzahl der berücksichtigten Messpunkte weiter optimiert werden. Hierzu wurden weitere Versuche durchgeführt und Auszüge hiervon sind in Tabelle 6-21 dargestellt. Durch die Variation (Vergrößerung bzw. Verkleinerung) des Datenbereichs ist nur eine geringfügige Verbesserung festzustellen, zudem zeigen Modelle bei kleineren Datenbereichen (< 40 Tage) ein instationäres Verhalten.

Tage		Anzahl	Prognosegüte			Modellanpassung	
Startpunkt k*96	Modelldaten k*96	p-Parameter	RMSE MW	MSE MW^2	MRE %	RMSE MW	Q
120	90	98	2,90	8,43	3,50	0,55	0,0278
150	90	98	1,88	3,54	1,79	0,51	0,0280
189	90	98	2,36	5,57	2,73	0,52	0,0441
210	90	98	2,52	6,34	2,65	0,55	0,0199
230	90	98	2,50	6,29	2,56	0,63	0,0220
Prognose mit Variation der Daten aus der Vergangenheit							
140	96	98	2,52	6,35	2,48	0,52	0,0271
175	40	98	2,91	7,57	2,75	0,52	0,0553

Tabelle 6-21: AR-Prognosemodell für Wochentage

Die besten Ergebnisse für das AR-Modell wurden mit einem Datenintervall der Größe 90 bis 100 Tage erzielt. Die Kombination von deterministischen Modellen mit stochastischen Prozessen ermöglicht eine deutliche Reduktion des Prognosefehlers im Vergleich zum reinen deterministischen Modell. Analog der Vorgehensweise folgt eine Prognose für die Wochenenden. Die verwendeten Parameter sind in Tabelle 6-22 dargestellt.

Tage		Anzahl	Prognosegüte			Modellanpassung	
Startpunkt k*96	Modelldaten k*96	p-Parameter	RMSE MW	MSE MW^2	MRE %	RMSE MW	Q
61	60	98	4,02	16,14	5,75	0,65	0,0561
70	60	98	3,48	12,10	4,71	0,58	0,0527
80	60	98	3,49	12,19	4,03	0,49	0,0398
90	60	98	4,08	16,65	5,85	0,44	0,0326
Prognose mit einem kleineren Modell							
61	6	4	1,90	3,62	3,32	0,22	0,0317
70	6	4	2,32	5,39	3,78	0,34	0,1082
80	6	4	3,30	10,87	3,36	0,34	0,0606
90	6	4	4,07	16,60	5,65	0,30	0,0327

Tabelle 6-22: AR-Prognosemodell für Wochenenden

Die Ergebnisse zeigen, dass größere AR-Modelle für die Wochenenden nicht geeignet sind, da sie das recht irreguläre Verhalten zwischen den Wochenendintervallen nur schwer nachbilden können. Der mittlere Prognosefehler der Wochenenden mit großen Modellen liegt durchschnittlich bei einem Wert von 5,08 %, die kleinen Modelle erzielen einen durchschnittlichen Wert von 4,02 %.

Tage		Anzahl		Prognosegüte		Modellanpassung		Anzahl
Start k*96	Modell k*96	p-Parameter	q-Parameter	RMSE MW	MRE %	RMSE MW	Q	G
90	**30**	**98**	**3**	**2,47**	**3,13**	**0,13**	**0,0016**	**15**
90	30	98	3	2,55	3,38	0,14	0,0018	10
90	30	98	3	2,99	3,57	0,49	0,0254	5
100	30	98	3	2,86	3,65	0,11	0,0186	15
100	**30**	**98**	**3**	**2,48**	**3,02**	**0,09**	**0,0028**	**10**
100	30	98	3	2,68	3,06	0,39	0,0027	5
110	30	98	3	Instationär				10
110	30	98	3	Instationär				5
120	30	98	3	3,08	3,69	0,15	0,0064	15
120	30	98	3	3,17	3,78	0,03	3,23e-004	10
120	30	98	3	3,64	4,52	0,17	0,0084	5
140	**30**	**98**	**3**	**2,12**	**2,78**	**0,15**	**0,0067**	**15**
140	30	98	3	2,14	2,81	0,17	0,0085	10
140	30	98	3	2,27	2,95	0,26	0,0173	5
170	30	98	3	1,70	2,07	0,15	0,0059	15
170	30	98	3	1,66	2,02	0,09	0,0020	10
170	**30**	**98**	3	**1,53**	**1,79**	**0,09**	**0,0020**	**5**
200	30	98	3	2,52	2,88	0,15	0.0013	15
200	30	98	3	2,47	2,79	0,20	0.0021	10
200	**30**	**98**	**3**	**2,41**	**2,70**	**0,23**	**0.0030**	**5**

Tabelle 6-23: ARMA-Prognosemodell für Wochentage

In einem weiteren Schritt wird nun das AR-Modell durch ein ARMA- bzw. ein ARIMA-Modell erweitert. Das ARMA/ARIMA-Modell wird mit dem Durbin-Verfahren parametrisiert, was bedeutet, dass die Anzahl der verwendeten Gleichungen eine zusätzlich zu konfigurierende Größe ist. Der AR-Anteil wird auf 98 Parameter gesetzt, da sich dieser Parameter bei den zuvor erstellten AR-Modellen als optimal erwiesen hat. Die Schätzung der Rauschwerte erfolgt mit Hilfe eines AR-Prozesses, der mit 192 Parametern konfiguriert ist. Die Bestimmung der Parameter erfolgt über ein Regressionsmodell, das den Prozess mit zusätzlichen Gleichungen (G) anpasst. Hierbei hat sich gezeigt, dass diese Vorgehensweise stabilere Resultate liefert im Vergleich zur Anwendung der Yule-Walker-Gleichung [Kar05]. Die Anzahl der verwendeten MA-Parameter hat maßgeblichen Einfluss auf die Prognose. Die Untersuchungen zeigen, dass die Verwendung weniger MA-Parameter im Durchschnitt

bessere Prognoseergebnisse liefert. Des Weiteren ist es nicht möglich, eine einheitliche Konfiguration zu bestimmen, da die Parameter für jede Prognose neu eingestellt werden müssen. Deshalb wurde ein Modell mit drei MA-Parametern gewählt, das zwar eine geringfügig schlechtere Prognose liefert, aber deutlich weniger Aufwand bei der Parameterbestimmung verursacht, da die Modelle dann fast immer stationäre Werte annehmen. Die Untersuchungsergebnisse sind in Tabelle 6-23 dargestellt.

Die Prognosen der Wochentage mit dem kleinsten Fehler wurden mit dem ARMA-Modell mit zehn zusätzlichen Gleichungen G=10 erzielt, gefolgt von G=15 Gleichungen. Mit diesem Modell lässt sich für die Wochentage der geringste MSE-Prognosefehler mit 2,79 % beziffern.

Es erfolgt die Prognose der Wochenenden entsprechend den vorausgegangenen Erkenntnissen. Aus der Prognose der Wochenenden mit AR-Prozessen werden dann kleine Modelle betrachtet. Das ARMA-Modell benötigt ein größeres Intervall von Vergangenheitswerten, um in die Lage versetzt zu werden, qualitativ ausreichende Rauschwerte zu schätzen. Die besten Ergebnisse wurden mit einem Intervall von 30 Tagen Länge erreicht. Die Tabelle 6-24 zeigt dieses exemplarisch.

Tage		Anzahl		Prognosegüte		Modellanpassung		Anzahl
Start k*96	Model l k*96	p-Para-meter	q-Para-meter	RMSE MW	MRE %	RMSE MW	Q	G
61	30	4	3	1,86	2,74	0,16	0,0046	10
70	30	4	3	instationär				10
80	30	4	3	instationär				10
90	30	4	3	instationär				10

Tabelle 6-24: ARMA-Prognosemodell für Wochenenden

Jede weitere Anpassung eines ARMA-Modells für Wochenendtage erwies sich als instabil. Ebenfalls wurden Anpassungen mit höheren Ordnungen vorgenommen, die auch zu einem instationären Verhalten des Modells führten.

Abschließend erfolgte der Einsatz eines ARIMA-Modells, das hauptsächlich dann Anwendung findet, wenn die Zeitreihe sich als instationär erweist. Ein Differenzenfilter erster Ordnung wird dem stochastischen Modell vorgeschaltet und differenziert das Residuum des deterministischen Modells. Das Ergebnis wird dann durch einen Summationsfilter wieder transformiert und die deterministischen Komponenten anschließend addiert. Die Parameter wurden wie die des ARMA-Modells gewählt, da sich diese als optimal erwiesen hat.

Die Tabelle 6-25 zeigt die Prognoseergebnisse mit dem ARIMA-Modell. Der durchschnittliche MSE-Prognosefehler liegt bei diesen Modellen bei einem Wert von 3,48 % und ist größer als bei den AR- oder ARMA-Modellen.

Tage		Anzahl		Prognosegüte		Modellanpassung		Anzahl
Start k*96	Modell k*96	p-Para-meter	q-Para-meter	RMSE MW	MRE %	RMSE MW	Q	G
90	30	98	3	2,91	3,73	3,58	18,46	10
110	30	98	3	2,90	3,35	3,28	12,80	10
140	30	98	3	2,14	2,80	2,54	8,59	10
150	30	98	3	2,68	3,05	2,51	10,06	10
160	30	98	3	3,77	3,50	2,79	7,81	10
170	30	98	3	2,98	3,68	2,77	6,73	10
200	30	98	3	4,17	4,36	4,69	12,25	10

Tabelle 6-25: ARIMA-Prognosemodell für Wochentage

Abbildung 6-14 zeigt exemplarisch den Vergleich zwischen ARMA- und ARIMA-Prognosemodell eines Wochentages. Gerade die Abendspitze im Lastgang zwischen 21 und 22 Uhr wird beim ARIMA-Modell weniger stark berücksichtigt. Insgesamt ist die Lastgangkurve des ARMA-Modells näher am Originallastgang, was sich auch in einem kleineren Fehler ausdrückt.

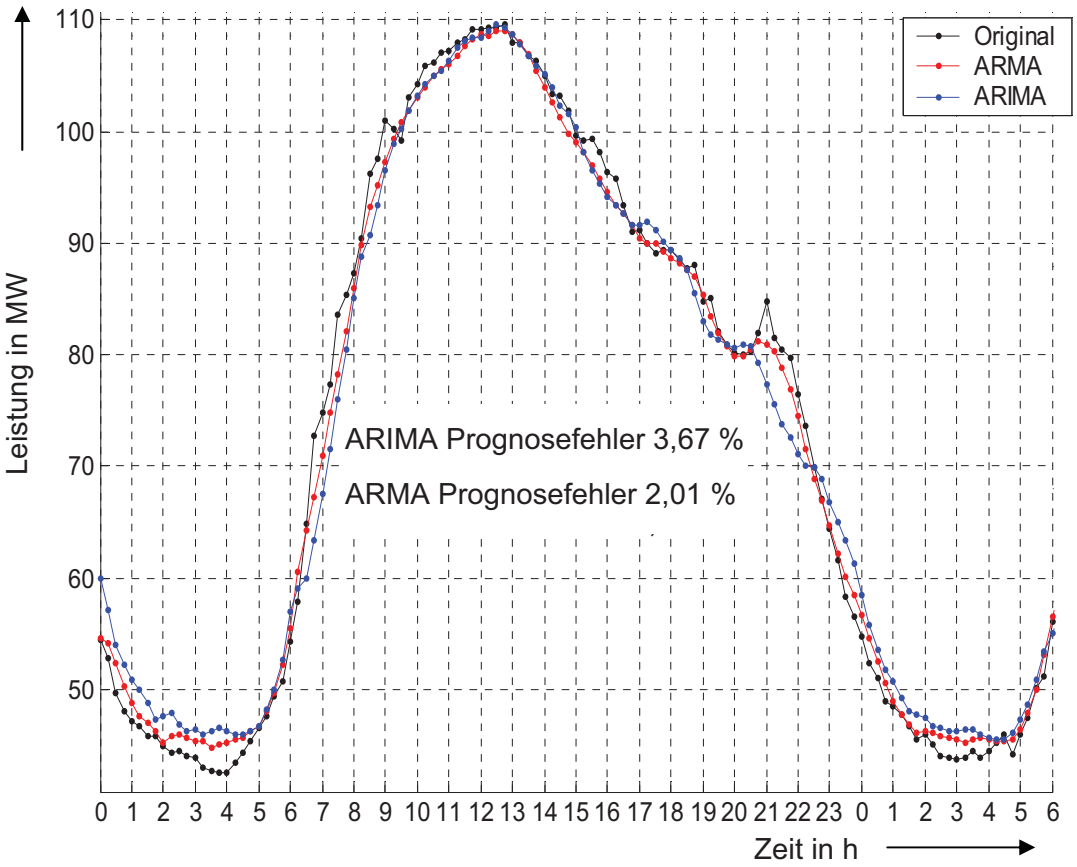


Abbildung 6-14: Vergleich zwischen ARMA- und ARIMA-Prognosemodell

6.4.3 Prognosen mit dem „entkoppelten“ autoregressiven Prozess

Die Untersuchungen der Lastgangdaten haben ergeben, dass ein Lastwert zu einer bestimmten Tageszeit eine sehr große Ähnlichkeit zum Lastwert des Vortags der entsprechenden Zeit aufweist. Dieses verleitet zu der Annahme, dass sich ein Lastwert einer bestimmten Tageszeit im Wesentlichen aus den Lastwerten der entsprechenden Tageszeit der Tage zuvor vorhersagen lässt. Dies bestätigt das Korrelogramm in Abbildung 5-2. Aus dieser Erkenntnis lässt sich ein vereinfachtes Prognoseverfahren ableiten. Die Originalzeitreihe und auch die der exogenen Größen werden in 96 - entsprechend der 1/4h-Werte der Zeitreihe - einzelne Tageszeitreihen zerlegt. Die Werte des Lastganges und der exogenen Größen zu den unterschiedlichen Tageszeiten werden hierdurch entkoppelt. Jede dieser Zeitreihen wird dem Prognoseverfahren „entkoppeltes Modell mit exogenen Größen“ (PARX-Parallel Auto Regressive eXogenous input) unterzogen. Das Trendpolynom beschränkt sich, wie es sich bereits bei den stochastischen Prozessen als zweckmäßig erwiesen hat, auf den Grad Eins. Der Restanteil wird als AR der Ordnung vier aufgefasst. Die Koeffizienten werden nicht durch die Methode der kleinsten Quadrate, sondern durch Lösen der Yule-Walker-Gleichung berechnet.

Die Anwendung des entkoppelten AR-Ansatzes (Kapitel 5.2) wurde jeweils mit (PARX / PAR / AXARMA) und ohne Einbeziehung exogener Größen bei den drei unterschiedlichen Verbrauchskollektiven (Stadt, Land, gemischtes Abnahmekollektiv) angewendet. Aufgrund der Ergebnisse zu den Untersuchungen der Einflüsse von exogenen Wettergrößen wurden als wichtige Einflussgrößen die Bodentemperatur und die Temperatur in 2 m Höhe identifiziert, die im Rahmen der Untersuchungen exemplarisch angewendet werden. Unter Verwendung des relativen mittleren Fehlers zur Bewertung der Qualität sind in Tabelle 6-26 die Ergebnisse der beiden Modellansätze dargestellt.

	städtisches Netzgebiet in %	ländliches Netzgebiet in %	städtisches und ländliches Mischnetzgebiet in %
Trend+zyklisch+ PARX	2,91	3,30	3,20
Trend+zyklisch+ AXARMA	3,72	4,12	3,93
Trend+zyklisch+ PAR	2,23	2,83	2,74
Trend+zyklisch+ ARMA	2,79	3,14	2,97

Tabelle 6-26: MRE - Fehler der entkoppelten Modelle

Damit bestätigt sich die Aussage der Korrelationsanalyse, dass der Einsatz von exogenen Größen bei der Prognose der elektrischen Last nur bedingt sinnvoll ist.

Die Integration der exogenen Größen bei den entkoppelten Modellen wirkt sich positiv auf die Prognosequalität aus, falls ein ausreichend starker Zusammenhang zwischen

den exogenen Größen und der zu prognostizierenden Größe besteht. Für die Prognose der elektrischen Last sind Modelle, die exogene Größen verwenden, nicht zu empfehlen, da die Zusammenhänge bei den untersuchten Lastkollektiven in der Regel gering und daher nicht brauchbar sind. Modelle mit exogenen Größen sind qualitativ besser als rein deterministische Modelle [Kar06], aber im Vergleich zu hierarchischen Modellen unterlegen, die zusätzlich stochastische Prozesse verwenden und sich ausschließlich auf das Lastprofil beschränken.

Im Gegensatz zur elektrischen Last führt die Integration exogener Größen zu einer deutlichen Verbesserung der Prognosequalität des Gasverbrauchs, da hier ein signifikanter Zusammenhang zwischen Last und Temperatur gezeigt werden kann. Die Verwendung exogener Größen kann sich aber auch nachteilig auf das Verhalten des Prognosemodells auswirken. Denn bei der Berechnung der Koeffizienten der AR-Prozesse fließen die exogenen Größen ebenfalls ein. Es kann der Fall auftreten, dass zur Gewährleistung eines stabilen Verhaltens des Prognosemodells oder der Lösbarkeit der eingesetzten Gleichungssysteme ein erhöhter Bedarf an historischen Modelldaten entsteht, wodurch die Flexibilität des Prognosemodells eingeschränkt wird und die Prognosen an Genauigkeit verlieren. Ebenso ist es möglich, dass die bei der Prognose eingehenden Wetterprognosewerte fehlerbehaftet sind und zu einer Verminderung der Prognosequalität führen können.

6.4.4 Zusammenfassung der Ergebnisse

Mit Hilfe von Trend- und zyklischen Anteilen konnten bereits Prognosen mit hinreichend guten Ergebnissen erstellt werden. Dabei wurde die Zeitreihe in zwei Abschnitte aufgeteilt. Mit den Daten des vorderen Abschnitts wurden Trend- und zyklischer Anteil ermittelt. Auf deren Basis wurde die Prognose für den Zeitbereich des hinteren Abschnitts erstellt. Die Prognosewerte x_i^{prog} und Messwerte x_i^{real} des hinteren Abschnitts führten zu dem Fehler (MRE). Polynomgrad und Anzahl der größten Frequenzanteile wurden so bestimmt, dass dieser Fehler klein wurde. Mit Anwendung dieses Verfahrens auf die Lastdaten unterschiedlicher Verbrauchskollektive wurde bereits ein mittlerer relativer Fehler unter 5 % erreicht. Ein weiterer Verbesserungsschritt zur Minimierung des mittleren relativen Fehlers ist die Gruppierung der Lastgangdaten mit den Ergebnissen der Datenclusterung mit Hilfe von selbstorganisierenden Karten (SOM). Diese gruppieren die Daten in entsprechende Klassen, z. B. Werk-, Wochenend- und Sondertage. Die Art der Gruppierung erwies sich als abhängig vom jeweiligen Verbrauchskollektiv. Weitere Reduzierungen des Prognosefehlers und damit der Restkomponente r_t sind mit den bisher betrachteten Verfahren nicht möglich.

Um eine zusätzliche Reduzierung des Prognosefehlers zu erreichen, wurde die Restkomponente r_t genauer betrachtet. Sie wurde als Ausprägung eines endlichen linearen

stochastischen Prozesses aufgefasst. Es wurde davon ausgegangen, dass den hier vorkommenden Zeitreihen lineare Prozesse endlicher Ordnung zu Grunde liegen. Es gibt mehrere Typen linearer Prozesse endlicher Ordnung. Die Modellierung der Restkomponenten erfolgte mit drei Ansätzen, dem Autoregressiven Prozess $AR(p)$ der Ordnung p sowie dem Autoregressive-Moving-Average-Prozess $ARMA(p,q)$ der Ordnungen p und q sowie ein $ARIMA$-Prozess**.** Bei dem $ARMA(p,q)$-Prozess wurden dazu die Werte e_t benötigt, die in einer vorgeschalteten AR-Anpassung mit hoher Ordnung p ermittelt wurden. Die Untersuchungen zeigen, dass bessere Prognoseergebnisse mit dem AR-Prozess erzielt werden konnten; wobei allerdings eine sehr hohe Ordnung p in Kauf genommen werden musste. Nachteilig sind aufgrund der hohen Ordnung das zeitaufwändige Lösen großer Gleichungssysteme sowie das mögliche Auftreten von Instabilitäten bei der Extrapolation der Zeitreihe für die Prognose. Um diesem abzuhelfen, wurde ein Ansatz mit einem „entkoppelten“ autoregressiven Prozess entwickelt. Dabei wird die Originalzeitreihe in 96 einzelne Tageszeitreihen zerlegt und somit die unterschiedlichen Tageszeiten entkoppelt. Für jede dieser Zeitreihen wird nun ein Prognoseverfahren erstellt, dabei beschränkt sich das Trendpolynom auf den Grad Eins und der Restanteil wird als $AR(4)$ aufgefasst.

	Wochentage		Wochenenden	
Modell	MRE %	(*s*) Standard-abweichung %	MRE %	(*s*) Standard-abweichung %
Trend+zyklisch	3,99	0,96	3,38	0,79
Trend+zyklisch+ AR	2,62	0,48	4,03	1,10
Trend+zyklisch+ ARMA	2,79	0,59	2,79	0,61
Trend+zyklisch+ ARIMA	3,48	0,51	5,54	2,61
Trend+zyklisch+ PAR	2,23	0,48	3,24	0,81
Trend+zyklisch+ PARX*	2,92	0,58	3,45	0,87
Trend+zyklisch+ AXARMA*	3,72	0,82	-	-

* Temperatur als exogene Größe

Tabelle 6-27: Gesamtübersicht des relativen mittleren Fehlers aller Modelle

Das entkoppelte Modell liefert bessere Ergebnisse als der gekoppelte $ARMA$-Prozess und bietet Geschwindigkeitsvorteile. Nachteilig beim Einsatz des $ARMA$-Gesamtprozesses ist das zeitaufwändige Lösen großer Gleichungssysteme, da diese Modelle relativ viele historische Daten benötigen (mind. 25 Tage), die beim entkoppelten Modell nicht erforderlich sind. Die Integration der exogenen Größen in das entkoppelte Modell führte zu keiner Verbesserung der Prognosequalität. Eine Übersicht über die mit den Zeitreihenansätzen erzielten Ergebnisse liefert die folgende Tabelle 6-27.

Die vorgestellten Modelle verfügen über eine nicht unerhebliche Anzahl von Parametern wie Grad des Trendpolynoms, Anzahl verwendeter Amplituden für die zyklische

Komponente, Anzahl der betrachteten historischen Daten usw. Um unter diesen Bedingungen zu einer geeigneten Konfiguration zu gelangen, die optimale Prognosen ermöglicht, hat sich folgende Heuristik bewährt:

1. Erstelle eine Prognose für die aktuell bekannten Daten.
2. Variiere die Parametrisierung so lange, bis die Prognose im Vergleich zu den bekannten Daten optimal ist.
3. Verwende die gefundene Parametrisierung für die zu erstellende Prognose.

Diese Form von Heuristik setzt jedoch voraus, dass die Daten einander ähnlich sind, plötzliche Veränderungen wie Feiertage können nicht erfasst werden, was durch die Gruppierung der Daten im Modellansatz ausgeschlossen werden kann.

6.5 Lastprognosen mit neuronalen Netzen

Der Einsatz künstlicher neuronaler Netze zur Lastprognose ist im Rahmen sehr verschiedener Strukturen denkbar. Als Eingangsgrößen können exogene Einflussgrößen und im Falle von autoregressiven Zusammenhängen auch Lastdaten vergangener Zeiträume verwendet werden. Bei der Systemarchitektur reicht die Spannbreite vom Einsatz eines einzelnen künstlichen neuronalen Netzes bis zu mehreren miteinander verbundenen bzw. parallel verknüpften Netzen [Kru92]. Prinzipiell lassen sich beliebig komplexe Architekturen erstellen. Auch hinsichtlich der zum Lernen der Netze verwendeten Datenbasis lassen sich Prognosesysteme unterscheiden.

6.5.1 Einschichtiges Netz

Für die erste Prognoseanwendung mit neuronalen Netzen wurden zunächst vorwärtsgerichtete Feedforward-Netze eingesetzt, die mit dem Backpropagationalgorithmus, der standardmäßig in Matlab integriert ist, trainiert wurden. Diese Netze wurden mit nur jeweils einer inneren verborgenen Schicht ausgeprägt. Das „große“ Netz ist mit vielen, im Untersuchungsablauf variablen, Eingabeneuronen ausgestattet und verfügt über nur ein Ausgabeneuron mit dann 96 Ausgabewerten zur Prognose eines Tages. Als Lern- und Testdaten dienten historische Lastwerte zweier Jahre. Um bei der Durchführung vergleichbare Startbedingungen zu erhalten, wurde die Basis der Lern- und Testdaten des neuronalen Netzes konstant gehalten und deren Initialisierung zu Beginn des Lernvorgangs vereinheitlicht. Die Optimierung der einzelnen Netze erfolgte unter Anwendung der in Kapitel 5.3.2 beschriebenen Systematik. Die Untersuchungen haben gezeigt [Dau05], dass bei der Prognoseanwendung mit „großen“ Netzen eine Anzahl von acht inneren Neuronen und eine Anzahl von 500 Lernzyklen bei einer gleichen Aufteilung aller Daten in Lern- und Testdaten gute Ergebnisse erzielen. Die einzelnen Ergebnisse sind in Tabelle 6-28 dargestellt. Innerhalb der Tests wurden dem Netz das Datum, der Wochentag und die Uhrzeit präsentiert, zusätzlich erfolgte die Kombination

unterschiedlichster Netzeingangsgrößen. In der Tabelle 6-28 sind die resultierenden Prognosefehler und Korrelationskoeffizienten dargestellt. Zur Beurteilung der Prognoseergebnisse der neuronalen Netze wurden diese jeweils zehnmal angelernt und der Mittelwert der Fehler und des Korrelationskoeffizienten errechnet und zur Beurteilung der Güte herangezogen. Der hier zugrunde gelegte mittlere absolute prozentuale Prognosefehler (MAPE) bezieht sich jeweils auf die Differenz zwischen tatsächlichem Ausgangswert und dem prognostizierten Ausgangswert des neuronalen Netzes in Relation zum Originallastgang.

Wettereingangsgrößen	RMSE	Korrelations-koeffizient	Ø MAPE %
Temperatur; Windgeschwindigkeit, Vortageswert, Vorwochenwert	2545	0,953	5,09
Temperatur; Windgeschwindigkeit, Vortageswert	2835	0,993	5,07
Vortageswert, Vorwochenwert	3390	0,938	6,58
Windgeschwindigkeit, Vortageswert	3567	0,956	6,79
Temperatur, Vortageswert	4059	0,993	5,84

Tabelle 6-28: Testergebnisse neuronaler Netze

Aus der Übersicht der Ergebnisse geht hervor, dass bei einem „großen“ Netz die Verwendung exogener Größen eine positive Auswirkung auf die Güte der Prognose hat. Eine weitere Verbesserung der Ergebnisse wurde mit der Anwendung des Zerlegungsansatzes (Kapitel 5.1.1) erzielt. Analog der Vorgehensweise der stochastischen Prozesse erfolgt hierbei die Modulierung der Restkomponente mit Hilfe der Methode der neuronalen Netze. Die Bestimmung des Trends erfolgt mit einer Polynomfunktion zweiten Grades, so dass ein relativ schwacher aber dennoch vorhandener Trend zu erkennen ist. Dieser wird abschließend vom Gesamtlastgang abgezogen und mittels der Fast-Fourier-Transformation auf seine Periodizität hin untersucht. Dieses sind die Tages-, Halbtages- und Jahresfrequenz. Mittels dieser Frequenzen wird nun eine Funktion erstellt und ebenfalls vom Lastgang subtrahiert. Die Restkomponente wird dann mit Hilfe des Netzes erzeugt. Die Durchführung einer Prognose zum Teilpunkt t erfolgt durch Addition der Periode $L_{Periode}(t)$, des Trends $L_{Trend}(t)$ und des Netzlastganges $L_{Netz}(t)$.

$$L_{prog}(t) = L_{Periode}(t) + L_{Trend}(t) + L_{Netz}(t) \qquad (6\text{-}3)$$

Analog dem vorherigen Vorgehen wurde ein neuronales Netz mit festen Parametern bestimmt und diesem entsprechend vorbereitete Daten präsentiert. Diese unterschieden sich nur insofern, dass die Anzahl der Eingangswerte durch Hinzu- oder Wegnahme von Wetterdaten voneinander abwich. Bei der Durchführung wurden für jede Kombination der Eingänge mehrere Trainingsläufe durchgeführt und der Mittelwert des MAPE gebildet. Der mittlere prozentuale Fehler bezieht sich in diesem Fall immer auf die Originaldaten, d. h. es wurden zur Bildung der Fehlerdifferenz der

Originallastgang und der prognostizierte Lastverlauf zuzüglich des Trends und der herausgenommenen Schwingungen der FFT verwendet. Nachstehende Kombinationen wurden in Verbindung mit der Zeit, dem Vortageswert und dem Vorwochenwert als Eingangsgrößen dem Netz zugeführt und untersucht:

Wettereingangsgrößen	Ø MAPE in %
Temperatur 2 m	3,06
Windgeschwindigkeit	3,04
Temperatur 2m, Windgeschwindigkeit	3,01
Temperatur 2m, Windgeschwindigkeit, Luftfeuchtigkeit	3,06
Temperatur 2m, Windgeschwindigkeit, Luftdruck	3,03
Temperatur 2m, Windgeschwindigkeit, Bodentemperatur	3,05
Alle verfügbaren Wetterdaten	3,01
ohne Wetterdaten	3,07

Tabelle 6-29: Ergebnis der Wetterdatenanalyse

Die Ergebnisse der Tabelle 6-29 zeigen durchweg bei allen untersuchten Kombinationen einen geringeren Prognosefehler als die Tests ohne eine Trend- und Periodenbereinigung. Dieses lässt sich dadurch erklären, dass periodische, also immer wiederkehrende Muster, die wie in diesem Fall besonders dominant sind, das neuronale Netz so beeinflussen, dass es nur auf diese Muster gut trainiert wird, jedoch nicht auf kleinste Änderungen im Lastverlauf. Diese sind aber von Bedeutung, da sie den Großteil des Prognosefehlers ausmachen.

6.5.2 Untersuchung verschiedener Netztopologien

Die Topologien der einzelnen neuronalen Netze und deren Eigenschaften wurden bereits im Kapitel vier beschrieben. Bei der Initialisierung eines Netzes lassen sich die Anzahl der Ein- und Ausgänge sowie die Aktivierungsfunktion und der Lernalgorithmus, mit dem das Netz trainiert werden soll, variieren. Um auch hierbei exemplarisch einen Vergleich der unterschiedlichen Netztopologien durchführen zu können, wurden die Eingabegrößen vereinheitlicht. Als Trainingsfunktion wurde der Backpropagation-Algorithmus einer Standardfunktion in Matlab verwendet, die Aktivierungsfunktion wurde als Tangens Hyperbolicus festgelegt. Die Anzahl der inneren Neuronen betrug 30. Die feste Datenbasis bestand aus den Daten eines Wochentages, dem Mittwoch eines Jahres, mit den Eingangsgrößen Uhrzeit, Temperatur, Windgeschwindigkeit, Globalstrahlung, Vortageslastwert und Vorwochenlastwert. Als Ausgabegröße wurde der zu prognostizierende Viertelstundenlastgangwert erzeugt. Die Abbildung 6-15 zeigt jeweils vier Trainingsläufe (T1 bis T4) für drei verschiedene Netztopologien und den resultierenden MAPE sowie den gemittelten Wert über alle vier Testreihen. Der MAPE ist in % angegeben.

Das Feedforward-Netz und das Elman-Netz unterscheiden sich in der Lerndauer nur geringfügig voneinander. Beim Radial-Basisfunktionen-Netz (RBF-Netz) trat jedoch eine erheblich höhere Lerndauer auf. Zudem ist bei den RBF-Netzen eine geringere Generalisierungsfähigkeit zu beobachten. Um ausreichend gute Ergebnisse zu erzielen, müssen hier wesentlich mehr innere Neuronen zur Verfügung gestellt werden, was wiederum zu einer weiteren Erhöhung der Trainingszeit führt. Die geringe Generalisierungsfähigkeit liegt bei den Radial-Basisfunktionen darin begründet, dass die Trainingsmuster, die stark vom Durchschnitt abweichen, nicht exakt dargestellt werden können, da durch das Anlernen von RBF-Netzen eine starke Klassifizierung der Trainingsdaten entsteht. Kann ein Datensatz keiner dieser Klassen zugeordnet werden, so werden die Prognoseergebnisse unbeeinflussbar verfälscht. Da bei der Prognose der Lastdaten der Wertebereich der Eingangswerte in Zusammenhang mit der Vielzahl dieser eine relativ hohe Dimension besitzt, tritt der Fall der extrem fehlerhaften Prognose sehr häufig ein. Dieser Netztyp ist damit zum Einsatz für eine Lastprognose ungeeignet und sollte nicht verwendet werden.

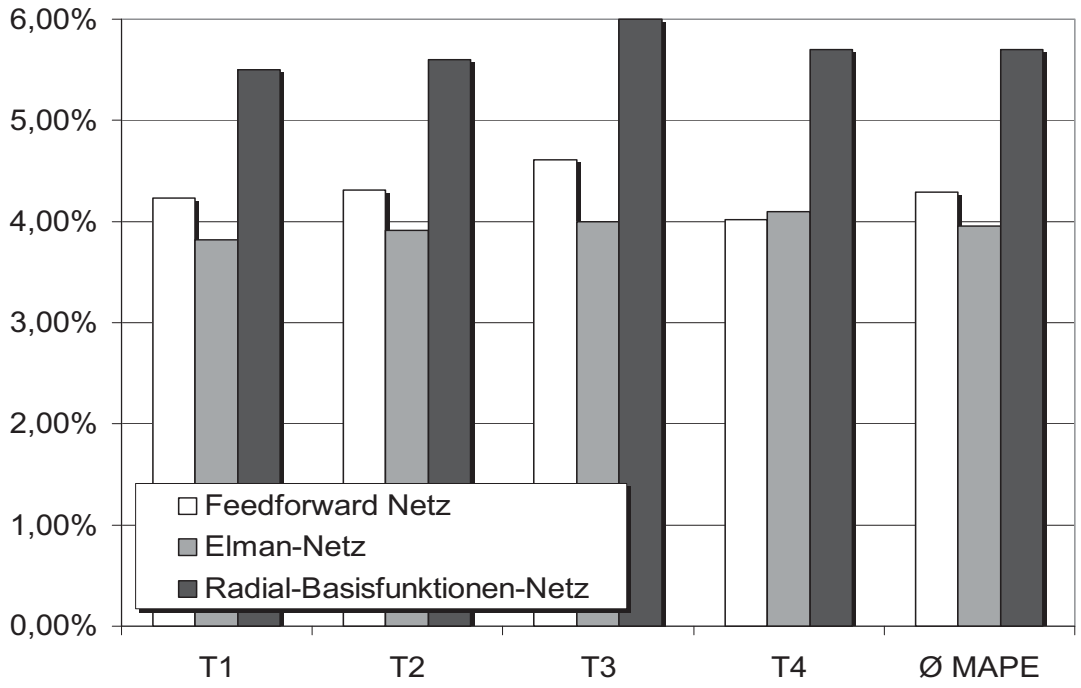


Abbildung 6-15: Testergebnisse unterschiedliche Topologien

Der geringste Prognosefehler konnte durch die Auswahl eines Elman-Netzes erzielt werden. Dieses lässt sich dadurch erklären, dass mittels der Rückkopplung der Ausgangswerte ein zeitlicher Bezug zwischen den auf das Netz gegebenen Datensätzen realisiert wird. Dieses ist durchaus gewollt, da die Datensätze, die jeweils einen Viertelstundenwert beschreiben, so wie sie dem neuronalen Netz präsentiert werden, chronologisch angeordnet sind und tatsächlich eine Abhängigkeit zum Vorgängerdatensatz besteht. Die Untersuchungen zeigen, dass es am sinnvollsten erscheint, die Lastprognose mit Elman-Netzen durchzuführen. Eine weitere Gegenüberstellung von Feedforward- und Elman-Netzen veranschaulicht die Tabelle 6-30 nochmals.

Es sind die jeweils besten Prognoseergebnisse der Netztypen an verschiedenen Wochentagen dargestellt. Verglichen werden der Korrelationskoeffizient r und der mittlere prozentuale Fehler aus Angabe zur Güte der Prognose. Wie aus der Tabelle zu erkennen ist, erzeugen Elman-Netze bei allen Tagtypen bessere Ergebnisse. Des Weiteren ist unabhängig von der Netztopologie zu erkennen, dass grundsätzlich die Wochentage Samstag und Sonntag schlechter zu prognostizieren sind als die Wochentage Montag bis Freitag.

	Feedforward-Netz		Elman-Netz	
Wochentag	r	MAPE in %	r	MAPE in %
Mittwoch	0,74	3,2	0,81	2,84
Freitag	0,75	3,3	0,82	2,97
Samstag	0,74	4,1	0,80	3,65
Sonntag	0,75	4,75	0,825	4,00

Tabelle 6-30: Vergleich Feedforward- und Elman-Netz

6.5.3 Anwendungsstrategie mit kleinen Netzen

Für die weiteren Prognoseanwendungen wurden vorwärtsgerichtete neuronale Netze mit nur jeweils einer inneren verborgenen Schicht verwendet. In den zuvor durchgeführten Untersuchungen wurde ein „großes“ Netz mit 96 Ausgabewerten sowie entsprechend vielen Eingabewerten trainiert. Im Folgenden soll jeder der viertelstündlichen Lastwerte des Zieltags einzeln vorhergesagt werden. Dieses Netz beschränkt sich auf einen einzigen Ausgabewert, dem Lastwert eines bestimmten Viertelstundenintervalls. Für einen Tag erfolgen somit 96 einzelne Auswertungen des entsprechend trainierten Netzes. Als Eingabewerte für die Vorhersage eines einzelnen viertelstündlichen Lastwertes dienten die Lastwerte der beiden Vorgängertage zur jeweiligen Viertelstunde sowie zusätzlich noch die von 0.00 Uhr an gezählte Nummer dieser Viertelstunde (0 bis 95). Die Nummer des Viertelstundenintervalls wurde in die Eingabe des Netzes übernommen, um zwischen verschiedenen Tageszeiten und dabei im Besonderen zwischen Tag und Nacht zu unterscheiden. Zur Verfügung standen hierfür Lastdaten zweier Jahre. Üblich ist es bei vielen Anwendungen neuronaler Netze, als Lerndaten einen Großteil der vorhandenen Daten zu verwenden (ca. 80 %). Eine solche Vorgehensweise konnte bei dieser Anwendung jedoch nicht erfolgen, da über längere Zeiträume das Verhalten der Stromverbraucher - nicht zuletzt jahreszeitlich bedingt - starken Veränderungen unterliegt, die sich deutlich im Lastprofil ausdrücken. Aufgrund dieser Tatsache bot sich an, nicht ein einziges Netz zu trainieren, das für einen langen Zeitraum gültig sein soll, sondern zur Vorhersage des Lastgangs eines jeden einzelnen Tages ein eigenes neuronales Netz zu erstellen, für dessen Lernvorgang ausschließlich auf Lastdaten der acht vorangegangenen Tagtypen zurückgegriffen wird. Ein Vorhersagevorgang für einen Tag besteht somit aus folgenden Schritten:

1. Bilden der Lerndatensätze anhand historischer Lastdaten der acht vorangegangenen Tagtypen.
2. Trainieren des Netzes für diesen Tag einschließlich Vorverarbeitung.
3. Bestimmen der Lastvorhersage für diesen Tag durch Auswerten des trainierten Netzes und Berücksichtigung der Vorverarbeitung

Die Berücksichtigung von genau acht Vorgängertagen hat sich bei Versuchen als geeignete Anzahl erwiesen: Das Verbraucherverhalten in einem Zeitraum von insgesamt neun Wochen weist nur vergleichsweise geringe Änderungen auf, und es steht so noch eine ausreichende Menge an Lerndaten zur Verfügung. Um die Güte dieser gewählten Lern- und Auswertungsstrategie beurteilen zu können, wurden umfangreiche Tests durchgeführt. Auch hierfür wurde wieder auf die historischen Lastdaten zurückgegriffen. Jeder Testvorgang verwendet ein Tageslastprofil eines ausgewählten Tages aus den verfügbaren historischen Lastdaten. Für diesen Tag wurde durch Ausführung des bereits beschriebenen Vorhersageprozesses mithilfe der acht Vorgängertage eine Testlastprognose erstellt und mit dem tatsächlichen Tageslastgang verglichen. Hierbei gelang es, die besten Prognosen mit einem mittleren absoluten Fehler von 2,11 % zu erstellen.

Zunächst wurden durch mehrere Versuche die vorkommenden Dimensionierungen bestimmt. Als optimal haben sich folgende, teilweise schon genannte Werte erwiesen:

- Die Anzahl innerer Neuronen des „kleinen“ Netzes wird auf den Wert 4 festgelegt.
- Zwei Vorgängertage werden für eine Tagesprognose berücksichtigt.
- Beim Lernvorgang werden sechs Vorgängertage berücksichtigt.

Ebenso erfolgten ausgiebige Versuche unter Einbeziehung von Wetterdaten. Die stündlich erhobene Bodentemperatur, Temperatur in 2 m Höhe, die Globalstrahlung und der Bedeckungsgrad wurden hierbei verwendet. Die Versuche haben jedoch gezeigt, dass durch diese Einbeziehung der Wetterdaten keine Verbesserung erzielt werden konnte. Dieses bestätigt die Ergebnisse in [Kar07]. Ebenso wurde erfolglos versucht, eine Ergebnisverbesserung durch vorherige periodische Bereinigung der Lern- und Testdaten zu erzielen. Jedoch konnte auch festgestellt werden, dass die Qualität der Prognoseergebnisse recht unterschiedlich ausfiel [Bie072]. Ein Vergleich der Prognosefehler an unterschiedlichen Testtagen ist mit 17,6 % sehr hoch und kann mit 2,02 % außerordentlich gering ausfallen. Die genauen Analysen zeigten, dass diese Fehlerabweichungen durch die Größenordnung der Lastwerte verursacht werden und die Profile der vorkommenden Lastkurven eine große Ähnlichkeit aufweisen. Das neuronale Netz ist demnach sehr wohl in der Lage, für einen Tag dieser oder ähnlicher Art ein Lastprofil, d. h. einen reinen Lastverlauf ohne Berücksichtigung der Höhe der Kurve mit befriedigender Güte vorherzusagen. Es bietet sich daher an, die beschriebene Prognosestrategie zu verfeinern, indem die Vorhersage des Lastprofils und die Vorhersage der Größenordnung des Lastverbrauchs voneinander abgekoppelt werden. Die Beschränkung auf die Betrachtung der reinen Lastprofile wird hierbei durch

eine Mittelwertbereinigung der verwendeten Lastdaten erreicht. Insgesamt ergibt sich damit folgende erweiterte Auswertungsstrategie:

- Mittelwertbereinigung,
- Prognose des Lastmittelwerts des Prognosetags,
- Zusammenführen der Profil- und Mittelwertprognose.

Zur Mittelwertbereinigung werden von jedem vorhandenen Lastwert x_i die Lastmittelwerte des betreffenden Tags abgezogen. Erforderlich ist dann zusätzlich eine Vorhersage des Lastmittelwertes des Prognosetages. Die Vorhersage des Lastmittelwertes hat sich gerade für die Wochentage als schwierigere Aufgabe dargestellt. Zur automatisierten Vorhersage der Lastmittelwerte wurden mit der vorliegenden Datenmenge mehrere Verfahren getestet, die sich jedoch gar nicht oder als sehr eingeschränkt als erfolgreich erwiesen. Diese Aufgabe konnte im Rahmen dieser Arbeit nicht befriedigend gelöst werden. Der einzige verwertbare Ansatz, der jedoch nicht auf neuronalen Netzen beruht und nur in wenigen Fällen ausreichende Ergebnisse zeigt, beruht auf der Auswahl eines Vorjahrstag der zum Referenztag die größte Ähnlichkeit besitzt.

6.5.4 Zusammenfassung der Prognoseergebnisse mit neuronalen Netzen

Der Einsatz neuronaler Netze liefert mit der hier angewendeten Strategie taugliche Ergebnisse. Die Entwicklungs- und Untersuchungsabfolge sowie die resultierenden Prognosefehler sind zusammenfassend in Tabelle 6-31 dargestellt.

Nr.	Untersuchung- und Entwicklungsschritt	Ø MAPE in %
1	Vorwärtsgerichtete Feedforward-Netze mit resilienten Backpropagationalgoritmus/Datenbasis: alle verfügbaren Daten bei keiner selektive Auswahl der Eingangsgrößen	7…10
2	Spezifikation der Eingangsgrößen wie Wetterdaten und Vortages- und Vorwochenwerte	5…7
3	Einschränken der Datenbasis auf nicht ereignisrelevante Tagetypen und einzelne Wochentage	4…5
4	Trend- und Periodenbereinigung sowie Modellierung der Restkomponente mit Hilfe der Methode der neuronalen Netze	3…4
5	Auswahl einer geeigneten Netztopologie/Elman-Netz	≥ 4

Tabelle 6-31: Prognosefehler neuronaler Netze

Bei den hier durchgeführten Untersuchungen wurden die Ergebnisse des einzelnen Entwicklungsschrittes im jeweiligen Folgeschritt berücksichtigt. Die Tabelle zeigt deutlich die Fortschritte der Prognoseverbesserung in Abhängigkeit der jeweils veränderten Parameter. Es ist vor allem zu erkennen, welchen entscheidenden Einfluss

eine systematische Datenclusterung und die damit verbundene Auswahl von geeigneten Eingangsgrößen auf das Ergebnis nehmen. Innerhalb dieses Schrittes konnte eine Verbesserung des Prognosefehlers von bis zu 5 % erzielt werden.

Des Weiteren wurden maßgebliche Verbesserungen durch Anwendung des Komponentenmodells erzielt. Nach Bereinigung des Trend- und Periodenanteils des Lastgangs wird die verbleibende Restlastkurve mit Hilfe des neuronalen Netzes prognostiziert. Die Auswahl von Netztopologien mit Kontext-Zellen wie dem Elman-Netz führte zur weiteren Verbesserung des Ergebnisses. Die Untersuchung mit „kleinen“ Netzen, d. h. die Prognose des Lastwertes eines bestimmten Viertelstundenintervalls mit Hilfe eines Netzes, führte zu keinen befriedigenden Ergebnissen.

7 Zusammenfassung und Ausblick

Mit dem Fortschreiten der Marktliberalisierung und der damit verbundenen Zuordnung neuer Marktrollen innerhalb der deutschen Energiewirtschaft haben sich in den vergangenen Jahren die Verantwortlichkeiten sowie die Anforderungen an Lastprognosen gänzlich verändert. Entscheidend in dieser veränderten Situation ist neben interner Kostenminimierung der Abschluss geeigneter Erzeugungs- und Liefervereinbarungen, die auf Grundlage qualitativ hochwertiger Prognosen erstellt werden.

Aufgrund des zunehmenden Wettbewerbes im deutschen Energiemarkt haben sich die Anforderungen an Lastprognosen gänzlich verändert. Aktuelle Befragungen zu Prognosemodellen und -methoden, die im Rahmen dieser Arbeit erfolgten, zeigen, dass heute überwiegend monolithische Verfahren eingesetzt werden. Diese Verfahren werden in der Regel in sogenannte Energiedatenmanagementsysteme integriert und bieten dem Anwender in der Regel einige Grundmodelle wie neuronale Netze, Fuzzy-Systeme, AR-, $ARMA$-Prozesse oder allgemeine Regressionsverfahren zur Erstellung von Lastprognosen an. Da dem Anwender diese Modelle häufig nur grob bekannt sind, bleibt hierbei oft die Qual der Wahl des eingesetzten Verfahrens. Nachteilig sind vor allem auch die aufwendige Konfiguration dieser Werkzeuge und damit auch gleichzeitig die Konfiguration der Modelle. Dabei muss entschieden werden, welche zur Verfügung stehenden historischen Daten und welche Auswahl exogener Größen zur Konfiguration herangezogen werden. Modell und Konfiguration können zudem noch stark vom jeweiligen Verbraucherkollektiv, sei es innerstädtisch, ländlich oder gemischt städtisch-ländlich, abhängig sein. Die Qualität der Vorhersage hängt unter anderem wesentlich von der Auswahl und dem Umfang historischer Datensätze ab, die dem Prognoseverfahren als Eingabe und somit als Basis für die Prognoseerstellung dienen. Im Rahmen der vorliegenden Arbeit wurden anhand von empirischen Untersuchungen unterschiedlicher Grundmodelle wie neuronale Netze, klassische Zeitreihenanalyse und die Methode der stochastischen Prozesse im Anwendungsbereich der kurzfristigen Lastprognose gegenübergestellt.

Zunächst wurden die Einflüsse exogener Größen und Wettereinflüsse auf unterschiedliche Kollektive detailliert mit klassischen statistischen Methoden analysiert. Sollte das Wetter tatsächlich einen nachweisbaren Einfluss auf die Zunahme elektrischer Last haben, so könnte durch ein durchdachtes Einbringen des Wetters in die Lastprognose der Prognosefehler gesenkt werden. Hierbei konnte jedoch festgestellt werden, dass der Einfluss von Feiertagen und einigen besonderen Ereignissen auf den zu prognostizierenden Tageslastverlauf evident ist, der Einfluss von Wettergrößen mit statistischen Mitteln jedoch nur sehr schwer nachzuweisen ist. Am deutlichsten wird ein derartiger Einfluss bei den Lastdaten eines eher ländlichen Versorgers, wie umfangreiche Korrelationsuntersuchungen gezeigt haben. Dabei wurden unterschiedliche Untersuchungszeiträume (Jahr, Monat oder Tag) und unterschiedliche Tagtypen zu Grunde gelegt. Wie sich dabei zeigte, sind diese Größen in den meisten

Fällen nur schwach mit dem elektrischen Lastgang korreliert und die Integration dieser in ein Prognosemodell nicht zu empfehlen. Da die Größen teilweise ein unterschiedliches Verhalten mit dem Jahresverlauf aufweisen, führt in einigen Fällen die Bereinigung um die Jahresperiode zu stärkeren Korrelationen. Eine hinreichend starke Korrelation, und damit ein hinreichend starker linearer Zusammenhang zwischen Wettergrößen und der elektrischen Last, liegen somit nur unter bestimmten Gegebenheiten vor. Analysen des Zusammenhangs zwischen Gasabsatz und dem Einfluss exogener Größen zeigen, dass eine Integration dieser in ein Prognosemodell sich positiv auf die Güte auswirkt.

In einem weiteren Arbeitsschritt erfolgte die Identifizierung von Sondertagen und die Einteilung in Lastmuster typischer Gruppen. Hierzu wurde eine Lastmusteranalyse mit Hilfe von SOM (selbstorganisierende Karten, self organizing maps) entwickelt und Untersuchungen mit Datensätzen verschiedener Kollektive durchgeführt. Zur Lastdatenanalyse wurde ein Datensatz als Vektor der Dimension 96 modelliert, wobei der Merkmalsraum der viertelstündlich gemessenen Lastgangdaten eines Tages abgebildet wird.

Das erste Ziel der Datenanalyse bestand darin, eine vorgegebene Anzahl von Elementen (d. h. der Lastgänge) anhand ihrer Merkmale (d. h. der Lastwerte) in möglichst homogene Gruppen einzuteilen, die ihrerseits möglichst heterogen sind. Da die Einflussgrößen auf die elektrische Last in Zusammenhang mit weiteren kalendarischen, lokalen wie auch globalen Ereignissen stehen, ist eine Datenclusterung zur Bestimmung des Tagtyps notwendig. Sowohl die Identifizierung der Sondertage als auch die Einteilung der Lastmuster in typische Gruppen wurde mit einer Lastmusteranalyse mit Hilfe von SOM im Rahmen von Untersuchungen mit Datensätzen verschiedener Versorgungsunternehmen erfolgreich durchgeführt.

Im Rahmen der Weiterentwickelung des Verfahrens wurde die Kohonen-Schicht um eine weitere aktorische Output-Schicht ergänzt. Diese dient in besonderer Weise der Interpretation der von der Kohonen-Schicht unabhängig erlernten Tagtyporganisation. Die „lernende“ Kohonen-Schicht wurde als zweidimensionales Raster ausgeprägt, wodurch zusätzlich zwischen den einzelnen Neuronen eine Nachbarschaftsbeziehung hinsichtlich der Lage berücksichtigt wird. Jedes Neuron der Kohonen-Schicht ist mit den sieben Neuronen der Outputschicht verbunden, die auch als motorische Karten fungieren und den Eingangsdatensatz einem Wochentagtyp zugeordnet werden. Nach Abschluss des Lernvorgangs konnten erfolgreich Datenanalysen mit diesem Verfahren durchgeführt werden. Hierzu wurden die bereits während des Lernvorgangs verwendeten Lastgangdaten und die zugehörigen, über das Datum referenzierten Merkmale des jeweiligen Tages benötigt. Um die Karten für Analysezwecke zu nutzen, wurde eine Filterfunktion eingesetzt, die einzelne Tage, deren Merkmale und den jeweiligen Lastgang des Tages der Karte berücksichtigt. Die mit Hilfe des Filters ermittelten Lastgangtage wurden der sensorischen Karte zugeführt und, ähnlich wie

beim üblichen Lernvorgang einer SOM, wurde über eine Metrik ein Gewinnerneuron der sensorischen wie auch motorischen Karte ermittelt. Mit Hilfe einer grafischen Ausgabe der sensorischen wie auch der motorischen Karte ließen sich nun Datencluster entsprechend der Filtereinstellung erfolgreich detektieren.

In einem weiteren Schritt wurden Prognosen mit Hilfe der klassischen Zeitreihenanalysen durchgeführt. Das hierbei entwickelte Modell beruhte wesentlich auf Methoden der Zeitreihenanalyse und im Speziellen auf der Theorie stochastischer Prozesse. Eine Zeitreihe ordnet diskreten Zeitwerten jeweils einen reellen Wert zu, wobei im vorliegenden Fall den Zeitwerten die Viertelstundenwerte des jeweils aktuellen gültigen Lastverbrauchswertes zugeordnet werden. Die klassische Zeitreihenanalyse betrachtet eine Zeitreihe als Überlagerung dreier Komponenten. Diese drei Komponenten wurden getrennt modelliert. Für die Trendkomponente wurde ein polynomiales Modell verwendet, wobei die Bestimmung der Koeffizienten mit der „Methode der kleinsten Quadrate“ nach Gauß erfolgte. Die Modellierung des zyklischen Anteils erfolgte analog einer Fourierreihe. Mit Anwendung dieses Verfahrens auf die Lastdaten unterschiedlicher Verbrauchskollektive wurde bereits ein mittlerer relativer Fehler unter 5 % erreicht. Ein weiterer Verbesserungsschritt zur Minimierung des mittleren relativen Fehlers ist die Gruppierung der Daten in entsprechende Klassen, wie Werk-, Wochenend- und Sondertage. Die Art der Gruppierung erwies sich als abhängig vom jeweiligen Verbrauchskollektiv. Die Verwendung des Verfahrens auf die zuvor gruppierten Lastdaten erreichte bei der getrennten Prognose von Werktagen und Wochenenden einen Fehler < 4 %. Weitere Reduzierungen des Prognosefehlers und damit der Restkomponente sind mit den bisher betrachteten Verfahren nicht möglich.

Um eine zusätzliche Reduzierung des Prognosefehlers zu erreichen, wurde die Restkomponente als Ausprägung eines endlichen linearen stochastischen Prozesses aufgefasst. Lineare stochastische Prozesse liefern eine Näherung der untersuchten Zeitreihe durch lineare Verknüpfung der Vergangenheitswerte, wobei die Gewichte den jeweiligen linearen Zusammenhang zwischen zwei Zeitpunkten repräsentieren. Es wird dabei davon ausgegangen, dass den hier vorkommenden Zeitreihen lineare Prozesse endlicher Ordnung zu Grunde liegen. Es gibt mehrere Typen linearer Prozesse endlicher Ordnung. Zur Modellierung der Restkomponenten erfolgten zwei Ansätze mit zwei unterschiedlichen Typen, dem Autoregressiven Prozess $\mathrm{AR(p)}$ sowie dem Autoregressive-Moving-Average-Prozess $\mathrm{ARMA(p,q)}$. Um mit diesen Verfahren hinreichend gute Prognoseergebnisse zu erzielen, müssen hohe Ordnungen der Prozesse in Kauf genommen werden, die zeitaufwändig zu lösen sind.

Um diesen Aufwand zu reduzieren, wurde ein Ansatz mit einem „entkoppelten“ autoregressiven Prozess (PARX-Parallel Auto Regressive eXogenous input) angewendet. Untersuchungen haben ergeben, dass ein Lastwert zu einer bestimmten Tageszeit eine sehr große Ähnlichkeit zum Lastwert des Vortags der entsprechenden Zeit aufweist.

Dieses verleitet zu der Annahme, dass sich ein Lastwert einer bestimmten Tageszeit im Wesentlichen aus den Lastwerten der entsprechenden Tageszeit der Tage zuvor vorhersagen lässt. Aus dieser Erkenntnis lässt sich ein vereinfachtes Prognoseverfahren ableiten. Die Originalzeitreihe wird in 96 einzelne Tageszeitreihen zerlegt. Die Lastwerte zu den unterschiedlichen Tageszeiten werden hierdurch entkoppelt. Jede dieser Zeitreihen wurde einem Prognoseverfahren unterzogen. Es zeigte sich sogar, dass mit diesem Ansatz des entkoppelten AR-Prozesses nicht nur der Rechenaufwand erheblich reduziert, sondern auch das Prognoseergebnis zum Vergleich zum üblichen AR-Prozess weiter verbessert werden konnte.

Für die Prognose mit der Methode der neuronalen Netze wird mittels Zuordnungsberechnung zwischen Ein- und Ausgabewerten ein gewünschter funktionaler Zusammenhang realisiert. Dieses wird insbesondere bei einem solchen funktionalen Zusammenhang angewendet, der nicht auf einfachem algorithmischem oder formelbasiertem Weg hergestellt werden kann. Ein erfolgreich trainiertes neuronales Netz sollte nun nicht nur zu den Beispielwerten, sondern zu beliebigen an den Eingängen anliegenden Werten die korrekten Ausgangswerte liefern, das heißt Ausgangswerte, die mit nur geringer Abweichung den zugehörigen Funktionswerten entsprechen. Im Rahmen der Arbeit wurde ein Ansatz verwendet, der im Wesentlichen drei Teilschritte enthält. Diese umfassen die Datenaufbereitung sowie die Lern- und Testphase. In einem ersten Untersuchungsschritt wurden mit vorwärtsgerichteten Feedforward-Netzen Prognosen erstellt. Da diese kontinuierliche Einflussgrößen verwenden und auf Diskretisierung und symbolische Zuordnungen verzichten, bieten sich die Feedforward-Netztopologien an. Hierbei erfolgt zunächst keine selektive Auswahl der Eingangsgrößen. Diese wurden im Rahmen der Versuchsreihen variiert. Hierbei konnten bereits Ergebnisse mit einem Fehler, zwischen 7-10 % erzielt werden. Durch Spezifikation der Eingangsgrößen sowie Auswahl geeigneter Lern- und Testdaten wurden weitere Verbesserungen erreicht.

Innerhalb eines weiteren Schrittes erfolgte analog der Vorgehensweise der klassischen Zeitreihenanalysen die Anwendung des Komponentenmodells, bei dem ausschließlich die Modulierung der Restkomponente mit Hilfe der Methode der neuronalen Netze erfolgte. Trend- und Periodenkomponente wurden dabei mit den klassischen Verfahren ermittelt. Bei der Modellierung der Restkomponente wurde ein neuronales Netz mit festen Parametern bestimmt, um eine weitere Verbesserung des zuvor erreichten Ergebnisses von ein bis zwei Prozent zu erreichen. Die Auswahl von Netztopologien mit Kontext-Zellen, wie beim Elman-Netz, führte zur weiteren Verbesserung des Ergebnisses.

Die ständig steigende Menge von elektrischer Energie aus dezentraler Erzeugung stellt neue Anforderungen an die Betreiber von Stromnetzen. Laut Dena Netzstudie wird bei einem politisch gewollten Ausbau der regenerativen Energien auf 20 % des Strombedarfs bis 2020, der Bedarf an Regelenergie bereits bis 2015 um 50 % bis 250 % (je

nach Szenario) zunehmen [Den05]. Dies liegt zum einen daran, dass insbesondere Wind- und Sonnenenergie sich nicht nach dem Erzeugungsbedarf des Verteilnetzes richten, sondern von sich kurzfristig ändernden Wetterbedingungen abhängen. Zusätzlich richtet sich das Erzeugungsmanagement dezentraler BHKW normalerweise nach dem Wärmebedarf und nicht nach dem Bedarf elektrischer Energie. Dieses erfordert eine Erweiterung und Anpassung der ursprünglich als unidirektional konzipierten Verteil- und Übertragungsnetze an die neuen Aufgaben. Es werden im Rahmen mehrerer Forschungsprojekte Anstrengungen unternommen, die Vernetzung von vielen dezentralen Erzeugungseinheiten zu virtuellen Kraftwerken über schnelle Kommunikationstechnologien vorzunehmen, mit dem Ziel einer Erzeugungsbündelung und dadurch bedingt verbessertem Management von Regelenergie. Andere Ansätze fassen unter dem Begriff „Virtuelles Power System“ [BMW09] [Wba08] nicht nur die Bündelung des Erzeugungsmanagements zusammen, sondern weiten diesen Ansatz zusätzlich mit dem Lastmanagement durch unterschiedliche Interventionsansätze auch auf die Kundenseite aus. Unstrittig ist bei all diesen Ansätzen, dass der Funktionsbaustein einer Erzeugungs- und Absatzprognose künftig weiterhin an Bedeutung gewinnen wird.

Die Entwicklungen werden zunehmend den Fokus auf eine dynamische Prognoseerstellung legen, d. h. abweichend von den in Kapitel 2.5 eher zeitpunktbezogenen Prognosenerstellung definierten Anforderungen, hin zu einer zeitraumbezogenen, sich ständig an die aktuelle Situation angepasste Prognose.

Ziel dieser Entwicklungen sollte dabei die Optimierung der Betriebsstrategien der eingesetzten Systeme sein unter Berücksichtung verschiedener meteorologischer, technischer und wirtschaftlicher Randbedingungen sowie unter Berücksichtigung der Intervention durch den Kunden selbst, der künftig innerhalb solcher Systeme eine eher aktivere Rolle einnehmen wird.

8 Literaturverzeichnis

[Ade86] Adermann, Hans-Jürgen; Wagner, Herbert: Verfahren zur Kraftwerkseinsatzoptimierung, VDE Verlag, Berlin Offenbach, etz Report 22, ISBN 3-8-007-1473-6, 1986

[Als95] Al-Sakaf, Omar Hassan Saleh: Tageslastprognose mit Hilfe künstlich neuronaler Netze im Vergleich mit praxisbewährten Verfahren, Dissertation, 1.Auflage RTWH Aachen, ISBN 3-86073-327-3, Dezember 1995

[Als08] Al-Shareef, A. J.; Mohamed, E. A.; Al-Judaibi, E.: One Hour Ahead Load Forecasting Using Artificial Neural Network for the Western Area of Saudi Arabia, World Academy of Science, Engineering and Technology 37, S. 219-224, 2008

[Amj01] Amjady, N.: Short-Term Hourly Load Forecasting Using Time-Series Modeling with Peak Load Estimation Capability, IEEE Transactions on Power Systems, Vol. 16, Nr. 3, 2001

[And97] Anders, Ulrich: Statische neuronale Netze, Dissertation, Fakultät für Wirtschaftswissenschaften der Universität Karlsruhe, Franz Vahlen Verlag GmbH, ISBN 3-8006-2184-3, 1997

[Asb75] Asbury, C. E.: Weather local model for electric demand and energy forecasting, IEEE Transactions on Power Apparatus and Systems, Vol. PAS –94 No. 4. 1975, Seite 1111-1116

[Bar82] Bartsch, Hans-Jochen: Taschenbuch Mathematische Formelsammlung, Verlag Harri Deutsch, Thun, Frankfurt; 1982

[Bär92] Bärmann, F.; Biegler-König, F.: On a class of Efficient Learning Algorithms for Neural Networks. Neural Networks, Pergamon Press, Vol. 5. pp 139-144F.,1992

[Ber07] Berger, Wolfgang: Die Last des Fußballs, ew, Das Magazin für die Energiewirtschaft; Jg. 106, Heft 21 Seite 18-20, 2007

[Bie071] Biermann, Jürgen; Fiedeldey, Michael; Gervens Theodor; Karrasch, Christian: Lastprognose mit neuronalen Netzen, VDE Verlag, Berlin Offenbach, etz Heft 9, Seite 52-59, 2007

[Bie072] Biermann, Jürgen; Fiedeldey, Michael; Gervens, Theodor; Karrasch, Christian: Lastprognose mit neuronalen Netzen, VDE Verlag, Berlin Offenbach, etz Heft 10, Seite 48-53, 2007

[Bil03] Bilsing, Markus: Erstellung eines Programms zur Beurteilung der Qualität von Prognosebibliotheken, Praxissemesterarbeit, FH-Osnabrück, 2003

[Bis97] Bishop, C. M: Neural networks for pattern recognition, Oxford University Press, 1997

[Bof06] Bofinger, S.; Heilscher, G.: Solar Electricty Forecast – Approaches And First Results, 20th European PV Conference, 4.-8. September 2006, Dresden, http://www.envisolar.com /factsheets/ meteo-control _bofinger _dresden.pdf, 2006

[Bof07] Bofinger, S.; Beyer, H.G.; Lorenz, E.; Wikramarathne, H.: Validierung und Optimierung eines Solarstromprognosesystems unter Berücksichtigung von Ensemble-Effekten, meteocontrol GmbH, http://www.energie-meteorologie.de/ publications/solar /conference/ 2007/staffelstein /validier -ung_und_optimierung_ eines_ solarstromprognosesystems _unter_ beruecksichtigung_ von_ensemble_ effekten.pdf, 2007

[Box70] Box, G, E. P.; Jenkins G. M.: Time Series Analysis, Forecasting and Control, Holden Day, San Francisco, 1970

[Bra95] Brause, Rüdiger: Neuronale Netze, Teubner Stuttgart, ISBN 3-519-12247-2, 1995

[Bra02] Braun, Markus: Untersuchungen zum Einsatz von Lastprognosen mit Hilfe künstlich Neuronaler Netze bei der Stadtwerke Osnabrück AG, Diplom-arbeit, Fachhochschule Osnabrück, Juni 2002

[Bun85] Bunn, D. W.; Framer, E. D.: Comparative Models for Electrical Load Forecasting, John Wiley and Sons, New York, 1985

[BMW09] Bundesministerium für Wirtschaft und Technologie: E-Energy - IKT-basiertes Energiesystem der Zukunft; Geförderte Projekte, Projekt-beschreibungen, http://www.e-energie.info/de/gefoerderte %25 projekte .php, 2009

[CIG92] Present Practices On Load Forecasting and Load Management: A Survey. By members of CIGRE Working Group WG 3903. Electra, No. 145, S. 69-89, Dezember 1992

[Cha82] Chatfield, Christopher: Analyse von Zeitreihen, Carl Hauser Verlag, ISBN 4-446-13597-9, München, 1982

[Cic93] Cichcki, A.; Unbehauen, R.: Neural Networks for Optmization and Signal Processing, B.G Teubner Stuttgart, ISBN3 519 06444, 1993

[Cor99] Cormen, Th. H.; Leiserson, Ch. E.; Rivest, R. L.: Introduction to Algorithms. Cambridge, Massachusetts/USA: MIT Press, 1999 (ISBN 0-262-03141-8)

[Cze96] Czernichow, T.; Piras, A.; Imhof, K.; Caire, P.; Jaccard, Y.; Dorizzi, B; Germond, A.: Short term electrical load forecasting with artificial neural networks. In: Engineering Intelligent Systems, Vol. 4, No. 2, 1996;

[Dau05] Dauwe, Denis: Datenbank gestützte Lastprognose mit Hilfe neuronaler Netze, Diplomarbeit, Fachhochschule Osnabrück, Januar 2005

[Den05] Studie im Auftrag der Deutschen Energie-Agentur GmbH (dena): Energiewirtschaftliche Planung für die Netzintegration von Windenergie in Deutschland an Land und Offshore bis zum Jahr 2020, Konsortium: DEWI / E.ON Netz / EWI / RWE Transportnetz Strom / VE Transmission, 2005

[Dav59] Davies, M.: The relationship between weather und electricity demand, IEEE Proceeding-C, Vol. 106 No.9, S 27-37, 1959

[DuM07] Richtlinie Datenaustausch und Mengenbilanzierung (DuM) Geschäftsprozesse: Unterbrechung und Wiederherstellung der Anschlussnutzung im Auftrag des Lieferanten (Gemeinsames Teilkapitel von VDN, VDEW und der EDNA-Initiative), 28. September 2007

[Ede98] Edemann, G. M.: Building a Picture of the brain, Daeddalus 127, Seite 37-69, 1991

[Edw81] Edwin, K.; Röllinger, H.: Untersuchungen zur kurzfristigen Prognose der elektrischen Leistung, Forschungsauftrag des Landes Nordrhein-Westfalen, Seite 14, 1981

[Edv90] Arbeitskreis EDV-Optimierung Kraftwerkeinsatz: DV-Optimierung des Kraftwerkseinsatzes - Definition, Anforderungen, Verfahren, Elektrizitätswirtschaft, Jg. 89, Heft 15, Seite 848-855, 1990

[Eick06] Eickhof, Norbert; Holzer, Verena Leila: Die Energierechtsreform von 2005, Ziele, Maßnahmen und Auswirkungen, Universität Potsdam, Potsdam, 2006

[Eld90] Lars, Elden, Linde Wittmeyer-Koch: Numerical Analysis: An Introduction, Boston, Massachusetts/USA Academic Pr, ISBN-13: 978-0122364303, Juni 1990

[EnWG98] Bundesgesetzblatt Jahrgang 1998: Gesetz zur Neuregelung des Energiewirtschaftsrechts (EnWG), Berlin, 24. April 1998

[EnWG07] Gesetz über die Elektrizitäts- und Gasversorgung (Energiewirtschaftsgesetz - EnWG), Berlin, 26.03.2007

[Fie051] Fiedeldey, Michael: Anforderungen und Einsatz von kurzfristiger Lastprognose im deutschen liberalisierten Energiemarkt, Das Magazin für die Energiewirtschaft, Jg. 104, Heft 11 Seite 20-23, 2005

[Fie052] Fiedeldey, Michael; Biermann, Jürgen; Gervens, Theodor: Mustererkennung in Lastgangdaten mit Hilfe selbstorganisierender Karten, VDE Verlag, Berlin Offenbach, etz Heft 7, Seite 30-34, 2005

[Fie07] Fiedeldey, Michael, Karrasch, Christian; Biermann, Jürgen; Gervens, Theodor: Lastprognose: Kombination von Zeitreihen und exogenen Einflussgrößen, VDE Verlag, Berlin Offenbach, etz Heft 6, Seite 56-60, 2007

[Fre05] Freie, Thorsten: Multivariate Zeitreihenanalyse, Diplomarbeit, FH-Osnabrück, 2005

[Gna04] Gnabs, Timo: SPSS: Zusammenhangshypothesen Tutorium III. 2004. – http://timo.gnambs.at/pages/spss/tut3/

[Gei01] Geiger, B.; Hellwig, M: Entwicklung von Lastprofilen für die Gaswirtschaft im Auftrag des Bundesverbandes der deutschen Gas- und Wasserwirtschaft e. V. und des Bundesverbandes kommunaler Unternehmen e. V., TU München Energiewirtschaft und Anwendungstechnik, September 2001

[Ger04] Gervens, Theodor: Wissensbasierte Systeme. 2004. - Vorlesungsunterlagen (Sommersemester), FH-Osnabrück, 2004

[Gil77] Gilchrist, W.: Statistical Forecasting, Wiley, 1977

[Gra01] Graf, F.-R.: Regelzonenübergreifender Einsatz von Regelenergie: Erfahrung aus der Abwicklung, DVG/VDN-Fachtagung Regelenergiemarkt in Deutschland, Dortmund, November 2001

[Gro87] Gross, G.; Galiana, F.: Short-Term Load Forecasting, Proceedings of the IEEE, Vol. 75, No. 12, S.1558-1572, Dezember 1987

[Gro93] DeGroot, Class: Nonlinear Time Series Analysis with Connectionist Networks, Dissertation Eidgenössische Technische Hochschule Zürich, März 1993, IPS Research Report No. 93-03 Dissertation ETH No. 10038

[GPKE06] Geschäftsprozesse und Datenformate für die Belieferung von Kunden mit Elektrizität, Bundesnetzagentur 11.07.2006 (Az. BK6-06-009)

[Hab93] Haubrich, J.: Elektrische Energieversorgungssysteme, Skriptum zur Vorlesung „Elektrische Anlagen“; RWTH Aachen ABEV, Band 13, Verlag Augustinus Buchhandlung, Aachen 1993

[Hah09] Hahn, H.; Meyer-Nieberg, S.; Pickl, S.: Electric load forecasting methods: Tools for decision making, European Journal of Operational Research 199 (2009), Seite 902-907, 2009

[Hai04] Hail, Henrick: ew, Das Magazin für die Energiewirtschaft; Risiko-absicherung bei Übernahme und eigenverantwortlicher Führung des Differenzbilanzkreises durch den Netzbetreiber; Jg. 103, Heft 15 Seite 20-21, 2004

[Ham93] Hamilton, Patrick: Künstlich Neuronale Netze, VDE Verlag GmbH, Berlin, Offenbach; ISBN3-8007-1824-3, 1993

[Han83] Hansmann, Karl-Werner: Kurzlehrbuch Prognoseverfahren, Verlag Gabler, ISBN 3-4091-3444-1, 1983

[Hay07] Hayati, M.: Short Term Load Forecasting Using Neural Networks for the West of Iran, Journal of Applied Sciences 7 (12), Seite 1582-1588, ISSN 1812-5654, 2007

[Heb49] Hebb, D. O.: The Organisation of Behaviour, Wiles, New York, Introduction and Chapter 4: The first stage of perception: growth of an assembly Seite 60-78, 1949

[Hec90] Hechtr-Nielsen, R: Neurocomputering, Addison-Wesley,1990

[Hob85] Hopfield, J.;Tank, D., W.: „neural“ computation of decisions in optimization problems, Biological Cybernetics, 52, Seite 141-152, 1985

[Hof92] Hoffmann, N.: Simulation Neuronaler Netze, ISBN3-528-15140-4, Vieweg Verlag Braunschweig, 1992

[Hof04] Hoffmann, Michaela: Künstlich neuronale Netze im empirischen Vergleich zu regressionsanalytischen Verfahren in der Werbungsforschung, Dissertation, Universität Osnabrück, Fachbereich Wirtschafts-wissenschaften, Shaker Verlag, ISBN 3-8322-2501-3, Aachen 2004

[Huf99] Hufendiek, Kai; Kaltschmitt, Martin: Einsatz künstlich Neuronaler Netze bei der kurzfristigen Lastprognose, Universität Stuttgart, Institut für Rationelle Energieanwendung (IER), 1999

[Huf00] Hufendiek, Kai: Systematische Entwicklung von Lastprognosesystemen auf der Basis neuronaler Netze, VDI Verlag, ISBN 3-18-345506-4, Dissertation, Universität Stuttgart 2000

[Huf97] Hufendiek, Kai; Albiger, J.: Künstliche neuronale Netze bei der kurzfristigen Lastprognose für die Kraftwerkseinsatzoptimierung. VDI-GET-Tagung „Optimierung in der Energiewirtschaft II“, 10.9.97,

Leverkusen, VDI-Berichte Nr. 1352, S. 325–341, VDI-Verlag, Düsseldorf, 1997;

[Hüt86] Hüttner, Manfred: Prognoseverfahren und ihre Anwendung, Verlag Walter de Gruyter, ISBN 3-1101-0826-7, 1986

[Jan96] Jan O.; Kehrberg, C.: Die Entwicklung des Elektrizitätsrechts in Deutschland - Der Weg zum Energiewirtschaftsgesetz von 1935 (= Rechtshistorische Reihe Nr. 157). Lang, Frankfurt u. a. 1996, ISBN 3-631-30797-7 (zugleich Dissertation, Universität Kiel 1996)

[Jes00] Jessen, Holger: Test- und Bewertungsverfahren zur regelbasierten Modellierung und Anwendung in der Lastprognose, VDI Verlag, ISBN 3-18-383608-4, Dissertation, Universität Dortmund 2000

[Joe96] Joeressen, A. : Systematisierung der Datenvorbereitung für die Prognose, Sitzung der Arbeitsgruppe „Prognoseverfahren“ und „Fuzzy Sets und Neuronale Netze“ in der DGOR, Stralsund, September 1996

[Kaa96] Kaastra, I., Boyd, M.: Designing a Neural Network for Forecasting Financial and Economic Time Series, Neurocomputering 10, Page 215-236, 1996

[Kar05] Karrasch, Christian: Lastprognose mit Hilfe von Zeitreihenuntersuchungen, Diplomarbeit, Diplomarbeit Fakultät Ingenieurwissenschaften und Informatik Fachhochschule Osnabrück, 2005

[Kar06] Karrasch, Christian; Fieldeldey, Michael; Biermann Jürgen; Gervens, Theodor: Lastprognose mit Hilfe von Zeitreihenuntersuchungen, VDE Verlag, Berlin Offenbach, etz Heft 2, Seite 22-26, 2006

[Kar07] Karrasch, Christian; Fieldeldey, Michael; Biermann Jürgen; Gervens, Theodor: Lastprognose: Kombination von Zeitreihen und exogenen Einflüssen, VDE Verlag, Berlin Offenbach, etz Heft 8, Seite 40-42, 2007

[Kas98] Kaski, S.; Kohonen, T.: Tips for Processing and Color-Coding of self-organizing maps. In Deboeck, G., Kohonen, T.: Visual Explorations in Finance with Self-Organizing-Maps, London S. 195-201, 1998

[Kin92] Kinnebrock, Werner: Neuronale Netze, Oldenbourg Verlag GmbH München 1992

[Kin94] Kinnebrock, Werner: Neuronale Netze, Grundlagen, Anwendung und Beispiele, Oldenbourg Verlag, ISBN 3-486-22947-8, 1994

[Koh72] Kohonen, Teuvo: Correlation Matrix Memories, IEEE-Transactions of Computers C-21: Seite 353-359, 1972

[Koh91] Kohonen, Teuvo: Self-organizing maps: optimization approaches. In Kohonen, T.; Simual, O.; Kangags, J.: Artificial Neural Networks. Proceeding of ICANN 91, International Conference of Artificial Neural Networks. Amsterdam, Vol. II, S. 981-990

[Kop08] Kopf, Jürgen: Skript Zeitreihenanalyse, Uni Würzburg www.wifak.uni-wuerzburg.de/ewf/doku/info-zra.htm

[Kra95] Kraaijveld, M. A.; Jain, A.: A nonlinear projection method based on Kohonen´s topology preserving maps. IEEE Transactions on Neural Networks Vol. 6, No. 3. (1995), Seite 548-559.

[Kru92] Kruse, Mangold, Mechler, Penger: Programmierung neuronaler Netze, Addison-Wesley ISBN 3-89319-293-X, 1992

[Kre06] Kreiß, Jan-Peter.; Neuhaus, Georg: Zeitreihenanalyse, Springer-Verlag, ISBN 3-5540-2562-8, Berlin, Heidelberg, New York, 2006

[Kuh04] Kuhn, Carsten: Konzeption und Realisierung eines Programms zur Analyse von Lastdaten mit Neuronalen Netzen, Diplomarbeit, Fakultät Ingenieurwissenschaften und Informatik Fachhochschule Osnabrück, 2004

[Law92] Lawrence, Jeanette: Neuronale Netze Computersimulation biologischer Intelligenz, Systema Verlag GmbH, ISBN 3-89390-271-6, 1992

[Lei86] Leiner, Bernd: Einführung in die Zeitreihenanalyse, Oldenbourg Verlag, ISBN 3-486-20299-5, München, Wien, 1986

[Lew74] Lewandowski, R.: Prognosesysteme und ihre Anwendung, Band 1, Walter Gruyter, 1974

[Lin96] Linke, Klaus: Kraftwerksführungssysteme der Verbundebene/ Energiewirtschaftliche Tagesfragen/46. Jg. (1996), Heft 10, Seite 618-624, 1996

[Lor07] Lorenz, E.; Hurka, J.; Heinemann, D.; Beyer H.G.; Schneider, M.: Solarleistungsvorhersage zur Netzintegration von Solarstrom, Universität Oldenburg, Institut für Physik, Hochschule Magdeburg-Stendal, Institut für Elektrotechnik, Meteocontrol GmbH, http://www.energiemeteorologie.de/ publications/solar/ conference/ 2009/ staffelstein/ solarleistungsvorhersage_staffelstein09.pdf, 2009

[Mak98] Makridakis, G; Wheelwright, S; Hyndman, R: Forecasting: Methods and Applications, Wiley & Sons, ISBN 0-471-53233-9, 1998

[Mal73] von der Malsberg, Christoph: Self-organization of orientation sensitve cells in the striata cortex, Kybernetik 14: Seite 85-100, 1973

[Man05] Mandal, P.; Senjyu, T.; Uezato, K.; Funabashi, T.: Several-Hours-Ahead Electricity Price and Load Forecasting Using Neural Networks, Power Engineering Society General Meeting, IEEE, ISBN: 0-7803-9157-8, Seite 2146 - 2153 Vol. 3, 2005

[Mat08] Matecki, Ute: Klassifikation mit Neuronalen Netzen in der Bildverarbeitung; Fachbereich Mathematik/Informatik, Universität Osnabrück, 2008; http://www.informatik.uni-osnabrueck.de/um/96/96.5/matecki/ matecki.html

[Mar85] Mareke, A.; Tillwicks, T.: Tageslastprognose: Bedeutung, Lösung offener Fragen. Energiewirtschaftliche Tagesfragen/35 Jg. (1985), Heft 11, Seite 772-782, 1985

[McP43] McCulloch, W. S.; Pitts, W. H.: A Logical Calculation of the ideas imminent in nervons activity, Bulletin of Mathematical Biophysics (1943) Vol. 5; Seite 115-133

[Mell03] Meller, Eberhard: ew, Das Magazin für die Energiewirtschaft, Sind die Stromkunden reif für den Wettbewerb?, Jg. 102, Heft 3 Seite 18-23, 2003

[Mie03] Miethe, Robert: Entwicklung und Anwendung einer Methode zum systematischen Aufbau von Lastprognosesystemen und Fallstudien zur Prognose von regenerativ erzeugter Energie, Praxissemesterarbeit, FH-Osnabrück, 2003

[Min88] Minsky M.; Papert S.: Perceptrons, 2nd Ed., MIT Press, Cambridge, 1988

[Moz93] Mozer, M.C.: Neural net architectures for temporal sequence processing, in Wiegend, N. Gersenfeld: Predicting the future an understanding the past, Addison Wesly, 1993 Preprint and Neuroprose-Archiv, 1993

[Pet86] Petritsch, Gerd: Orakel: Bedarfsprognosen elektrischer Last mit Regression und Karhunen Loe´ve Entwicklung, Dissertation Technische Universität Wien, Wien 1986

[PIR96] Piras, A.; Germond, A.; Buchenel, B.; Imhof, K.; Jaccard, Y.: Heterogeneous artificial neural network for short term electrical load forecasting. Proc. of IEEE Power Industry Computer Applications Conference, 7-12/5/95, Salt-Lake-City (UT), U.S.A, pp. 319-324, 1995

[PWC07] PricewaterhouseCoopers (Hrsg.): Entflechtung und Regulierung in der deutschen Energiewirtschaft. Praxishandbuch zum Energiewirtschafts-gesetz. Haufe, ISBN 978-3-448-08025-4, 2007

[Pri05] Prins, Jack: NIST/SEMATECH e-Handbook of Statistical Methods, Chapter 6 - Process or Product Monitoring and Control. http://www.itl.nist.gov/div898/handbook/ (Abruf 20.04.2005)

[Pod05] Poddig, Thorsten; Sidorovitch, Irina: Künstliche Neuronale Netze: Überblick, Einsatzmöglichkeiten und Anwendungsprobleme, http://www.ku-eichstaett.de/Fakultaeten/WWF/(Abruf 19.04.2005)

[Reh95] Rehkugel, H.; Kerling, M: Einsatz Neuronaler Netze für Analyse- und Prognosezwecke; Betriebswirtschaftliche Forschung und Praxis, Heft 3, Seite 306-324; 1995

[Rey08] Rey, Günter Daniel: Neuronale Netze; Bern: Huber, 2008, ISBN: 978-3-456-84513-5

[Rl 03] Rl 2003/54/EG des Europäischen Parlaments und des Rates vom 26.06.2003 über gemeinsame Vorschriften für den Elektrizitätsbinnenmarkt und zur Aufhebung der RL 96/92/EG, Abl. Nr. L176 S.37

[Rip97] Ripley, B.B: Pattern Recognition and Neural Networks, Oxford University Press, 1997

[Rit90] Ritter, H.; Martinetz, T.; Schulten, K.: Neuronale Netze - Eine Einführung in die Neuroinformatik selbstorganisierender Netzwerke, Addison-Wesley, ISBN 3-89319-172-0, 1990

[Roj93] Rojas, Raul: Theorie der neuronalen Netze. Eine systematische Einführung; Springer Verlag Berlin Heidelberg New York, 1993

[Roj96] Rojas, Raul: Neural Networks - A Systematic Introduction. Springer-Verlag, Berlin, New-York, (502 p., 350 illustrations), 1996

[Ros58] Rosenblatt, Frank: The perceptron: a probabilistic model for information storage and organization in the brain, psychological Review 65, Seite 368-408, 1958

[Rum86] Rumelhart D. E.; Hilton, G. E.; Williams, R. J.: Learning representations by back-propagation errors; in Rummelhard D. E., MC Clelland J. L.: Parallel Distributed Processing: Exploration in the Microstructure of Cognition; Vol 1 MIT Press, Cambridge; MA Seite 318-362, 1986

[Sch89] Schlittgen, Rainer; Streitberg, Bernd: Zeitreihenanalyse, Oldenbourg Verlag, ISBN 3-4862-5725-0, 2001

[Sch91] Schmitter, E. D.: Neuronale Netze, Hofackerverlag, ISBN 3-88963-312-9, 1991

[Sch97] Schlittgen, Rainer: Einführung in die Statistik, Oldenbourg Verlag, ISBN 3-4862-4198-2, 7. Auflage 1997

[Sch04] Schwarz, H. R.; Köckler, N.: Numerische Mathematik. Stuttgart: Teubner, 2004 (ISBN 3-219-42960-8)

[See08] Seehafer, Norbert; Spahn, Frank: Das Klassische Komponentenmodell einer Zeitreihe, http://www.agnld.uni-potsdam.de/~reik/dynModelle/Kapitel03.pdf, Universität Potsdam, 2008

[Sen02] Senjyu, T.; Takara, H.; Uezato, K.; Funabashi, T.: One-hour-Ahead Load Forecasting Using Neural Network, IEEE Transactions on Power Systems, Vol. 17, No. 1, Februar 2002

[Sid06] Sideratos, G.; Hatziargyriou; N.D.: Application of Radial Basis Function Networks for Wind Power Forecasting, Artificial Neural Networks – ICANN 2006, Springer Berlin Heidelberg, ISBN 978-3-540-38871-5, 2006

[Spo08] Sponsel; Rudolf; Rathsmann-Sponsel, Irmgard: Fast-Kollinearität in Korrelationsmatrizen mit Eigenwert-Analysen erkennen. Ergänzungsband - Band II zu "Numerisch instabile Matrizen und Kollinearität“; http://www.sgipt.org/wisms/statm/kor/kurkor.htm, iec-verlag; ISSN-0944-5072; ISBN 3-923389-13-2; 2008

[Ste04] Steen, Annette: Entwicklung und Anwendung einer Methode zum systematischen Aufbau von Lastprognosesystemen und Fallstudien zur Prognose von regenerativ erzeugter Windenergie, Diplomarbeit, Universität Flensburg, 2004

[Sti01] Stier, Winfried: Methoden der Zeitreihenanalyse, Springer-Verlag, Berlin, Heidelberg, New York, ISBN 3-540-41700-1, Heidelberg 2001

[Sto00] Stolzke, Ulf: Neuronale Netze zur Prognose von Warenterminpreisen, Peter Lang, Europäischer Verlag der Wissenschaft/Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Wien, Lang, 2000

[StromNZV] Verordnung über den Zugang zu Elektrizitätsversorgungsnetzen (Stromnetzzugangsverordnung - StromNZV) Bundesgesetzblatt Jahrgang 2005 Teil I Nr. 46, ausgegeben zu Bonn am 28. Juli 2005

[StromNEV] Verordnung über die Entgelte für den Zugang zu Elektrizitätsversorgungsnetzen (Stromnetzentgeltverordnung - StromNEV) Bundesgesetzblatt Jahrgang 2005 Teil I Nr. 46, ausgegeben zu Bonn am 28. Juli 2005

[Str02] Stecker, Stefan; Göbelt, Mathias: Liberalisierte Energiemärkte, Strategie, Prognose, Handel, VDI Verlag, ISBN 3-18-314716-5, Beiträge der Forschungsgruppe Energiemärkte, Leipzig 2002

[Tag01] Tagscherer, Michael: Künstliche Neuronale Netze, Kontinuierliches Lernen, Adaption, Prozessautomatisierung, Dissertation, Technische Universität Chemnitz 2001

[Thi98] Thiesing, Frank M.: Analyse und Prognose von Zeitreihen mit neuronalen Netzen, Shaker Verlag, ISBN 3-8265-4242-8, Dissertation, Universität Osnabrück, Mai 1998

[Ult93] Ultsch, A.: Knowledge extraction from self-organizing neural networks. In: Opitz, O.; Lausen, B.; Klar, R. (Eds.): Information and Classification. Berlin, S. 301-306, 1993

[VDN02] Verband der Netzbetreiber-VDN-e.V beim VDEW: Kommentarband, Umsetzung der Verbändevereinbarung über Kriterien zur Bestimmung von Netznutzungsentgelten für elektrische Energie und über Prinzipien der Netznutzung vom 13. Dezember 2001, November 2002

[Wba08] M. Delfanti, M. Merlo – Fondazione Politecnico di Milano M. Mura, S. Lukovic, U. Bondi – AlaRI-USI M. Stöhr, L. Karg – B.A.U.M. Consult Virtual Power Systems White Book, Dezember 2008

[Web90] Weber, Karl: Wirtschaftsprognostik, Verlag Vahlen, ISBN 3-8006-1385-9, 1996

[Wel05] Welling, Matthias: Entwicklung und Implementierung von Verfahren zur Clusterung von Lastgangdaten mit der Methode der Künstlich Neuronalen Netze, Diplomarbeit, Fakultät Ingenieurwissenschaften und Informatik, Fachhochschule Osnabrück, 2005

[Zel94] Zell, Andreas: Simulation neuronaler Netze, Oldenbourg Verlag München Wien, ISBN 3-486-24350-0, Juli 1994

[Zha03] Zhang, G. Peter. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing Band 50, 2003, S. 159–175.

[Zöf01] Zöfel, Peter: Statistik verstehen, Addison-Wesly scientific computing, ISBN 3-8273-1690-1, 1. Auflage 2001

9 Anhang

9.1 Lineare und Polynominale Trends

Die einfachste Trendfunktion ist eine Gerade $m_t = a_1 + a_2 \cdot t$ ausgehend von $x_t = m_t + e_t$ mit $t = 1...N$ erhält man $e_t = x_t - m_2 = x_t - a_1 - a_2 \cdot t$ aus der „Methode der kleinsten Quadrate“ folgt:

$$Q = \sum_{t=1}^{N} \left(x_t - a_1 - a_2 \cdot t\right)^2 \rightarrow \min_{a_1, a_2} \tag{A-1}$$

Für Polynome der Ordnung k gilt: $m_t = a_1 + a_2 \cdot t + a_3 \cdot t^2 + a_4 \cdot t^3 + ... + a_k \cdot t^{k-1}$

$$Q = \sum_{t=1}^{N} \left(x_t - a_1 - a_2 \cdot t - a_3 \cdot t^2 - a_4 \cdot t^3 - ..... - a_k \cdot t^{k-1}\right)^2 \rightarrow \min_{a - a_k} \tag{A-2}$$

Beispiel für $k = 4$ mit polynomialem Trend: $m_t = a_1 + a_2 \cdot t + a_3 \cdot t^2 + a_4 \cdot t^3$

$$Q = \sum_{t=1}^{N} \left(x_t - a_1 - a_2 \cdot t - a_3 \cdot t^2 - a_4 \cdot t^3\right)^2 \rightarrow \min_{a_1, a_2, a_3, a_4} \tag{A-3}$$

$$\begin{aligned}
\frac{\partial Q}{\partial a_1} &= 2 \cdot \sum_{t=1}^{N} \left(x_t - a_1 - a_2 \cdot t - a_3 \cdot t^2 - a_4 \cdot t^3\right) \cdot 1 = 0 \\
\frac{\partial Q}{\partial a_2} &= 2 \cdot \sum_{t=1}^{N} \left(x_t - a_1 - a_2 \cdot t - a_3 \cdot t^2 - a_4 \cdot t^3\right) \cdot t = 0 \\
\frac{\partial Q}{\partial a_3} &= 2 \cdot \sum_{t=1}^{N} \left(x_t - a_1 - a_2 \cdot t - a_3 \cdot t^2 - a_4 \cdot t^3\right) \cdot t^2 = 0 \\
\frac{\partial Q}{\partial a_4} &= 2 \cdot \sum_{t=1}^{N} \left(x_t - a_1 - a_2 \cdot t - a_3 \cdot t^2 - a_4 \cdot t^3\right) \cdot t^3 = 0
\end{aligned} \tag{A-4}$$

$$\begin{aligned}
&\sum_{t=1}^{N} \left(x_t - a_1 - a_2 \cdot t - a_3 \cdot t^2 - a_4 \cdot t^3\right) = 0 \\
&\sum_{t=1}^{N} \left(x_t \cdot t - a_1 \cdot t - a_2 \cdot t^2 - a_3 \cdot t^3 - a_4 \cdot t^4\right) = 0 \\
&\sum_{t=1}^{N} \left(x_t \cdot t^2 - a_1 \cdot t^2 - a_2 \cdot t^3 - a_3 \cdot t^4 - a_4 \cdot t^5\right) = 0 \\
&\sum_{t=1}^{N} \left(x_t \cdot t^3 - a_1 \cdot t^3 - a_2 \cdot t^4 - a_3 \cdot t^5 - a_4 \cdot t^6\right) = 0
\end{aligned} \tag{A-5}$$

Umgestellt nach x_t:

$$
\begin{aligned}
a_1\sum_{t=1}^{N}1 + a_2\sum_{t=1}^{N}t + a_3\sum_{t=1}^{N}t^2 + a_4\sum_{t=1}^{N}t^3 &= \sum_{t=1}^{N}x_t \\
a_1\sum_{t=1}^{N}t + a_2\sum_{t=1}^{N}t^2 + a_3\sum_{t=1}^{N}t^3 + a_4\sum_{t=1}^{N}t^4 &= \sum_{t=1}^{N}x_t\cdot t \\
a_1\sum_{t=1}^{N}t^2 + a_2\sum_{t=1}^{N}t^3 + a_3\sum_{t=1}^{N}t^4 + a_4\sum_{t=1}^{N}t^5 &= \sum_{t=1}^{N}x_t\cdot t^2 \\
a_1\sum_{t=1}^{N}t^3 + a_2\sum_{t=1}^{N}t^4 + a_3\sum_{t=1}^{N}t^5 + a_4\sum_{t=1}^{N}t^6 &= \sum_{t=1}^{N}x_t\cdot t^3
\end{aligned}
\qquad \text{(A-6)}
$$

Man erhält das Gleichungssystem, welches über die Methode der kleinsten Quadrate zu lösen ist.

$$
\begin{pmatrix} c_{00} & c_{01} & c_{02} & c_{03} \\ c_{10} & c_{11} & c_{12} & c_{13} \\ c_{20} & c_{21} & c_{22} & c_{23} \\ c_{30} & c_{31} & c_{32} & c_{33} \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^{N} x_t \\ \sum_{t=1}^{N} x_t\cdot t \\ \sum_{t=1}^{N} x_t\cdot t^2 \\ \sum_{t=1}^{N} x_t\cdot t^3 \end{pmatrix} \Rightarrow \underline{A}\cdot\vec{x} = \vec{b} \Rightarrow \vec{x} = \left(\underline{A}^T\cdot A\right)^{-1}\cdot\underline{A}^T\cdot\vec{b}
$$

$$
\text{mit } c_{ij} = \sum_{t=1}^{N} t^{i+j} \qquad \text{(A-7)}
$$

Man erhält die Koeffizienten a_j mit $j = 1.....k$ des verwendeten Polynoms.

9.2 Periodogramm und Fourieranalyse

Dazu wird die um den Mittelwert bereinigte Zeitreihe $(x_t - \overline{x})$ angewendet. Aus dem Ansatz der „Methode der kleinsten Quadrate“ ist bekannt:

$$
Q = \sum_{t=1}^{N}\left[\left(x_t - \overline{x}\right) - \beta_1\cdot\cos(2\pi\cdot f\cdot t) - \beta_2\cdot\sin(2\pi\cdot f\cdot t)\right]^2 \rightarrow \min_{\beta_1,\beta_2} \qquad \text{(A-8)}
$$

erhält man das aus der Trendbestimmung bekannte Gleichungssystem:

$$
\begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^{N}\left(x_t - \overline{x}\right)\cdot\cos(2\pi\cdot f\cdot t) \\ \sum_{t=1}^{N}\left(x_t - \overline{x}\right)\cdot\sin(2\pi\cdot f\cdot t) \end{pmatrix} \qquad \text{(A-9)}
$$

$$
c_{ij} = c_{ji} = \sum_{t=1}^{N}\cos(2\pi\cdot f\cdot t)\cdot\sin(2\pi\cdot f\cdot t) \text{ für } i \neq j \text{ sonst} \qquad \text{(A-10)}
$$

$$c_{00} = \sum_{t=1}^{N} \cos^2(2\pi \cdot f \cdot t) \;;\; c_{11} = \sum_{t=1}^{N} \sin^2(2\pi \cdot f \cdot t)$$

Aufgrund der Orthogonalitätsrelation für Fourierfrequenzen $f = \frac{k}{N}$; $k \in$ IN gilt:

$$\left.\begin{aligned} &\sum_{t=1}^{N} \cos\left(2\pi \cdot \frac{k}{N} \cdot t\right) \cdot \sin\left(2\pi \cdot \frac{k}{N} \cdot t\right) = 0 \\ &\sum_{t=1}^{N} \cos^2\left(2\pi \cdot \frac{k}{N} \cdot t\right) = \sum_{t=1}^{N} \sin^2\left(2\pi \cdot \frac{k}{N} \cdot t\right) = \frac{N}{2} \end{aligned}\right\} \Rightarrow \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix} \Rightarrow \begin{pmatrix} N/2 & 0 \\ 0 & N/2 \end{pmatrix} \tag{A-11}$$

Die Kovarianzen der Zeitreihe mit Sinus und Cosinus ergeben [Sch87]:

$$\left.\begin{aligned} C(f) &= \frac{1}{N} \sum_{t=1}^{N} (x_t - \bar{x}) \cdot \cos(2\pi \cdot f \cdot t) \\ S(f) &= \frac{1}{N} \sum_{t=1}^{N} (x_t - \bar{x}) \cdot \sin(2\pi \cdot f \cdot t) \end{aligned}\right\} \Rightarrow \begin{pmatrix} N/2 & 0 \\ 0 & N/2 \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} N \cdot C(f) \\ N \cdot S(f) \end{pmatrix} \tag{A-12}$$

eingesetzt erhält man:

$$\begin{aligned} &\beta_1 \cdot \frac{N}{2} = N \cdot C(f); \qquad \beta_2 \cdot \frac{N}{2} = N \cdot S(f) \\ &\beta_1 = 2 \cdot C(f); \qquad \beta_2 = 2 \cdot S(f) \end{aligned} \tag{A-13}$$

Die Intensität der Frequenz f innerhalb der Zeitreihe wird durch die Summe der quadrierten Residuen gemessen. Für die angepasste harmonische Schwingung gilt [Sch87]:

$$Q = \sum_{t=1}^{N} \left[(x_t - \bar{x}) - 2 \cdot \left[\underbrace{C(f)}_{\beta_1} \cdot \cos(2\pi \cdot f \cdot t) + \underbrace{S(f)}_{\beta_2} \cdot \sin(2\pi \cdot f \cdot t) \right] \right]^2 \tag{A-14}$$

Das Periodogramm ist folgende Funktion der Frequenz:

$$\begin{aligned} I(f) &= N\left[C(f)^2 + S(f)^2\right] \\ &= N\left[\frac{1}{N} \sum_{t=1}^{N} (x_t - \bar{x}) \cdot \cos(2\pi \cdot f \cdot t)\right]^2 + N\left[\frac{1}{N} \sum_{t=1}^{N} (x_t - \bar{x}) \cdot \sin(2\pi \cdot f \cdot t)\right]^2 \end{aligned} \tag{A-15}$$

weiter gilt:

$$e^{j \cdot 2\pi \cdot f \cdot t} = \cos(2\pi \cdot f \cdot t) + j \cdot \sin(2\pi \cdot f \cdot t)$$

$$C(f) = \frac{1}{N} \sum_{t=1}^{N} (x_t - \bar{x}) \cdot \cos(2\pi \cdot f \cdot t)$$

$$S(f) = \frac{1}{N}\sum_{t=1}^{N}(x_t - \bar{x}) \cdot \sin(2\pi \cdot f \cdot t)$$

$$D(f) = C(f) + j \cdot S(f) \qquad \text{(A-16)}$$

$$D(f) = \frac{1}{N}\sum_{t=1}^{N}(x_t - \bar{x}) \cdot \cos(2\pi \cdot f \cdot t) + j \cdot \frac{1}{N}\sum_{t=1}^{N}(x_t - \bar{x}) \cdot \sin(2\pi \cdot f \cdot t)$$

$$D(f) = \frac{1}{N}\sum_{t=1}^{N}(x_t - \bar{x}) \cdot e^{j \cdot 2\pi \cdot f \cdot t}$$

$$I(f) = N \cdot |D(f)|^2$$

Der Ausdruck $\frac{1}{N}\sum_{t=1}^{N}(x_t - \bar{x}) \cdot e^{j \cdot 2\pi \cdot f \cdot t}$ kann als Fouriertransformation

$$C_k = \frac{1}{N}\sum_{t=1}^{N} z_t \cdot e^{-j2\pi \cdot \frac{k}{N} \cdot t} \quad \text{mit} \quad z_t = (x_t - \bar{x}) \Rightarrow \text{Hintransformation}$$

(A-17)

$$z_t = \sum_{t=1}^{N} C_k \cdot e^{j2\pi \cdot \frac{k}{N} \cdot t} \qquad \Rightarrow \text{Rücktransformation}$$

der Zeitreihe aufgefasst werden. Man erhält ein zweiseitiges Spektrum:

$$I(f) = N \cdot |C_k|^2 \text{ mit } f = \frac{k}{N} \; ; \; k = 0,....,N-1 \qquad \text{(A-18)}$$

Hinweis: Hier soll die mathematische Standarddefinition benutzt werden. Die hier verwendete Fouriertransformation benutzt im Gegensatz zur Definition des Periodogramms [Sch87] einen negativen Exponenten. Das Periodogramm repräsentiert das quadrierte und normierte Spektrum der Fouriertransformation einer Zeitreihe. Anzumerken ist, dass zu jeder Frequenz ein konjugiert komplexes Gegenstück mit identischem Absolutwert existiert. Je nachdem, ob das einseitige oder das zweiseitige Spektrum betrachtet wird, muss das Periodogramm angepasst werden [Bar82].

9.3 2-Stichproben T-Test

Ein statistischer Test untersucht unter Berücksichtigung einer Irrtumswahrscheinlichkeit eine aufgestellte Hypothese, die entweder abgelehnt oder angenommen wird. Der 2-Stichproben T-Test untersucht, ob die Mittelwerte zweier Stichproben gleich sind. Als Nullhypothese wird angenommen, dass die Mittelwerte mit einer Irrtumswahrscheinlichkeit von $\alpha = 0{,}05$ gleich sind. Die zugehörigen Quantile

entstammen den Tabellen aus [Bar82]. Die Stichproben können hierbei verschiedene Längen und Varianz haben. Es gilt:

$$H_0 : x_t = \overline{x}$$
$$H_1 : x_t \neq \overline{x} \Rightarrow x_t - \overline{x} = d_0 \qquad \text{(A-19)}$$

Stichprobe: X, Y mit σ_x^2, σ_y^2 und der Länge N_x, N_y:

$$T = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{\sigma_x^2}{N_x} - \frac{\sigma_y^2}{N_y}}} \qquad \text{(A-20)}$$

für die Varianz gilt dann:

$$T = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{\sigma_x^2}{N_x} - \frac{\sigma_y^2}{N_y}}} \Rightarrow \sigma_p^2 = \frac{(N_x - 1)\sigma_x^2 + (N_y - 1)\sigma_y^2}{(N_x + N_y - 2)} \qquad \text{(A-21)}$$

Ablehnung der Nullhypothese, wenn:

$$T < t_{(a/2,v)} \Leftrightarrow T > t_{(a/2,v)} \qquad \text{(A-22)}$$

Für den Freiheitsgrad bei ungleicher Varianz gilt:

$$\nu = \frac{\left(\frac{\sigma_x^2}{N_x} + \frac{\sigma_y^2}{N_y}\right)^2}{\frac{\left(\frac{\sigma_x^2}{N_x}\right)^2}{(N_x - 1)} + \frac{\left(\frac{\sigma_y^2}{N_y}\right)^2}{(N_y - 1)}} \qquad \text{(A-23)}$$

Bei gleicher Varianz gilt:

$$\nu = N_x + N_y - 2 \qquad \text{(A-24)}$$

9.4 Chi-Quadrat-Test

Der Chi-Quadrat-Test überprüft, ob eine Standardabweichung s einer Stichprobe einem spezifischen Wert entspricht. Es existiert ebenfalls eine zweiseitige Version dieses Testes; wobei geprüft wird, ob die Standardabweichung der Stichprobe größer oder kleiner als der spezifizierte Wert ist. Es gilt:

$$H_0 : s_t^2 = s^2$$
$$H_a : s_t^2 < s^2 \qquad \text{(A-25)}$$
$$H_a : s_t^2 > s^2$$

$$T = (N-1)\left(\frac{a}{s}\right) \qquad \text{(A-26)}$$

mit : N Stichprobengröße
α Standardabweichung

Ablehnung der Hypothese:

$$T > \chi^2(\alpha, N-1) \Rightarrow \quad \text{unterer Bereich} \qquad \text{(A-27)}$$
$$T < \chi^2(1-\alpha, N-1) \Rightarrow \quad \text{oberer Bereich}$$

mit α Quantil der Chi-Verteilung mit N-1 Freiheitsgraden

9.5 Yule-Walker-Gleichung

Ist ein allgemeingültiger autoregressiver Prozess $AR(p)$ wie folgt definiert:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \alpha_2 X_{t-3} + \ldots + \alpha_p X_{t-p} + e_t \qquad \text{(A-28)}$$

dann ergibt eine Multiplikation mit $X_{t-\tau}$ und die Bildung des Erwartungswerts $E(X_t X_{t-\tau})$. Dann gilt:

$$\gamma(\tau) = \alpha_1 E(X_{t-1} X_{t-\tau}) + \alpha_2 E(X_{t-2} X_{t-\tau}) + \ldots + \alpha_p E(X_{t-\tau} X_{t-\tau-p}) + E(e_t X_{t-\tau})$$

$$\gamma(\tau) = \alpha_1 \gamma(\tau - 1) + \alpha_2 \gamma(\tau - 2) + \ldots + \alpha_p \gamma(\tau - p) + \underbrace{E(e_t X_{t-\tau})}_{=0} \qquad \text{(A-29)}$$

Die Division mit $\gamma(0)$ ergibt:

$$\rho(\tau) = \alpha_1 \rho(\tau - 1) + \alpha_2 \rho(\tau - 2) + \ldots + \alpha_p \rho(\tau - p)$$

$$\begin{bmatrix} \rho(1) \\ \rho(2) \\ \rho(3) \\ \vdots \\ \vdots \\ \rho(p) \end{bmatrix} = \begin{bmatrix} 1 & \rho(1) & \rho(2) & \cdots & \cdots & \rho(p-1) \\ \rho(1) & 1 & \rho(2) & \cdots & \cdots & \rho(p-2) \\ \rho(2) & \rho(1) & 1 & \cdots & \cdots & \rho(p-3) \\ \vdots & \vdots & \cdots & 1 & \cdots & \vdots \\ \vdots & \vdots & \cdots & \cdots & 1 & \vdots \\ \rho(p-1) & \rho(p-2) & \rho(p-3) & \cdots & \cdots & 1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \vdots \\ \alpha_p \end{bmatrix} \qquad \text{(A-30)}$$

$\vec{a} = \underline{P}^{-1}\vec{p} =$ Parameter des AR-Prozesses $\rho(\tau) = \rho(-\tau)$/Quelle: [Sch87]

9.6 Durbin-Methode

Das Verfahren baut auf dem Ansatz der „Methode der kleinsten Quadrate“ auf. Das Modell wird dabei in folgender Form aufgefasst:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \alpha_2 X_{t-3} + ... + \alpha_p X_{t-p} + e_t - \beta_1 e_{t-1} - \beta_2 e_{t-2} - ... - \beta_q e_{t-q} \qquad \text{(A-31)}$$

Die unbekannten Residuen werden durch die Anpassung eines AR-Prozesses großer Ordnung an die Zeitreihe approximiert.

$$e_t = X_t - \alpha_1 X_t - .... - \alpha_k X_{t-k} \qquad \text{(A-32)}$$

Die Parameter des AR-Prozesses können über die Yule-Walker-Gleichung bestimmt und alternativ kann die Rekursion nach Levinson benutzt werden[1].

$$\sum_{t=1}^{p+q} \left[X_t - \left(\alpha_1 X_{t-1} + \alpha_1 X_{t-2} + ... + \alpha_p X_{t-p} - \beta_1 e_{t-1} - \beta_2 e_{t-2} - ... \beta_q e_{t-q} \right) \right] \rightarrow \min \qquad \text{(A-33)}$$

Hinweis: Der Fehlerterm e_t des AR-Anteils des Modells wird nicht berücksichtigt. Der Grund für die Vernachlässigung des Fehlerterms e_t wird in der Literatur nicht genannt (Definition Schlittgen Streitberg, Zeitreihenanalyse). Der Parameter $\beta_0 = 1$ ist gesetzt und das Gleichungssystem kann auch in der Form

$$\underbrace{\begin{bmatrix} X_{t-1} & \cdots & X_{t-p} & e_{t-1} & \cdots & e_{t-q} \\ X_{t-2} & \cdots & X_{t-p-1} & e_{t-2} & \cdots & e_{t-q-1} \\ X_{t-3} & \cdots & X_{t-p-2} & e_{t-3} & \cdots & e_{t-q-1} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ X_{t-p-q-1} & \cdots & X_{t-2p-q-1} & e_{t-p-q-1} & \cdots & e_{t-p-2q-1} \end{bmatrix}}_{\underline{A}} \cdot \underbrace{\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_p \\ \beta_1 \\ \vdots \\ \beta_q \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} X_t \\ X_{t-1} \\ X_{t-2} \\ \vdots \\ \vdots \\ X_{t-p-q} \end{bmatrix}}_{\vec{b}} \qquad \text{(A-34)}$$

$$\left(\underline{A}^T \underline{A} \right)^{-1} \underline{A}^T \vec{b} = \vec{x}$$

aufgestellt werden. Es wird vorgeschlagen, mehrere Kombinationen der Ordnungen zu berechnen und aus der Anpassung an den Originalprozess auf die passende Modellordnung zu schließen[2]. Die Schätzwerte der Parameter sind konsistent, aber nicht effizient[3]. Dennoch liefert die Anpassung der Parameter durch dieses Verfahren gute Prognosen.

Um das Gleichungssystem für die „Methode der kleinsten Quadrate“ eindeutig zu bestimmen, sind der Anzahl der unbekannten Parametern entsprechend viele

[1] Quelle Schlittgen Streitberg Zeitreihenanalyse, 2. Auflage 1987
[2] Quelle Schlittgen Streitberg Zeitreihenanalyse, 2. Auflage 1987
[3] Quelle Schlittgen Streitberg Zeitreihenanalyse, 2. Auflage 1987

Gleichungen nötig. Testprognosen haben gezeigt, dass die Überbestimmung des Gleichungssystems durch mehrere Datensätze eine bessere Schätzung der Parameter liefert. Mit der Trail & Error-Methode kann schnell ein Optimum gefunden werden wobei gute Ergebnisse bereits mit fünf bis 20 zusätzlichen Datensätzen erreicht werden.

9.7 Ergebnisse des Mittelwertstationaritäts - sowie des Varianz-stationaritätstests

Länge der Lastzeitreihe: 70.176
Anzahl der Tests: 20
Länge der Stichprobe: 96
α: 0.05
Freiheitsgrade (t-Test): 95
Freiheitsgrade (χ2-Test): 95

Nr.	t-Test			χ2-Test			
	Stichproben-beginn	p-Wert	x_t=x	Stichproben-beginn	p_1-Wert	p_2-Wert	s_t^2=s^2
1	6.239	0,0010	0	10.033	0,9357	0,0643	0
2	34.708	0,0000	0	31.490	0,0433	0,9567	0
3	43.131	0,0000	0	66.295	0,2818	0,7182	0
4	36.453	0,0000	0	646	0,6056	0,3944	0
5	32.976	0,0000	0	62.125	1,0000	0,0000	-1
6	67.165	0,0000	0	42.674	1,0000	0,0000	-1
7	61.008	0,0008	0	26.336	0,1084	0,8916	0
8	18.831	0,0910	1	62.335	0,9046	0,0954	0
9	65.352	0,0000	0	59.485	1,0000	0,0000	-1
10	21.082	0,0000	0	50.409	0,7897	0,2103	0
11	31.860	0,2389	1	53.238	0,8569	0,1431	0
12	66.101	0,6253	1	15.622	0,9918	0,0082	-1
13	10.963	0,9772	1	55.152	0,9843	0,0157	-1
14	34.141	0,0000	0	25.923	1,0000	0,0000	-1
15	42.696	0,0000	0	28.547	1,0000	0,0000	-1
16	38.631	0,0000	0	37.807	0,8785	0,1250	0
17	70.056	0,0000	0	58.064	1,0000	0,0000	-1
18	2.306	0,0000	0	743	0,7693	0,2307	0
19	53.007	0,1099	1	56.166	1,0000	0,0000	-1
20	10.269	0,0000	0	16.203	0,7208	0,2792	0

Tabelle 9-1: Ergebnisse des Mittelwertstationaritäts- und des Varianzstationaritätstests

9.8 Analyse ausgewählter Wetterelemente

Zu einigen exemplarisch ausgewählten Wetterelementen wurden die für die Zeitreihen-analyse wichtigsten Informationen ermittelt. Dazu gehört eine Grafik mit dem Verlauf der Messungen aus zwei Jahren. Die Grafik zeigt des Weiteren das Jahresmittel, den linearen Trend und die Schwingung mit der Jahresperiode. Eine 3-D-Grafik zeigt das

jeweilige Wetterelement in Abhängigkeit von der Tageszeit und vom Jahresverlauf sowie das Periodogramm. Darunter folgen die Auswertungen der empirischen Momente, der fünf Schwingungen mit den größten Amplituden und der Koeffizienten des Trendpolynoms. Für diese Berechnungen wurden die Daten aus den Jahren 2003 und 2004 verwendet.

9.8.1 Temperatur

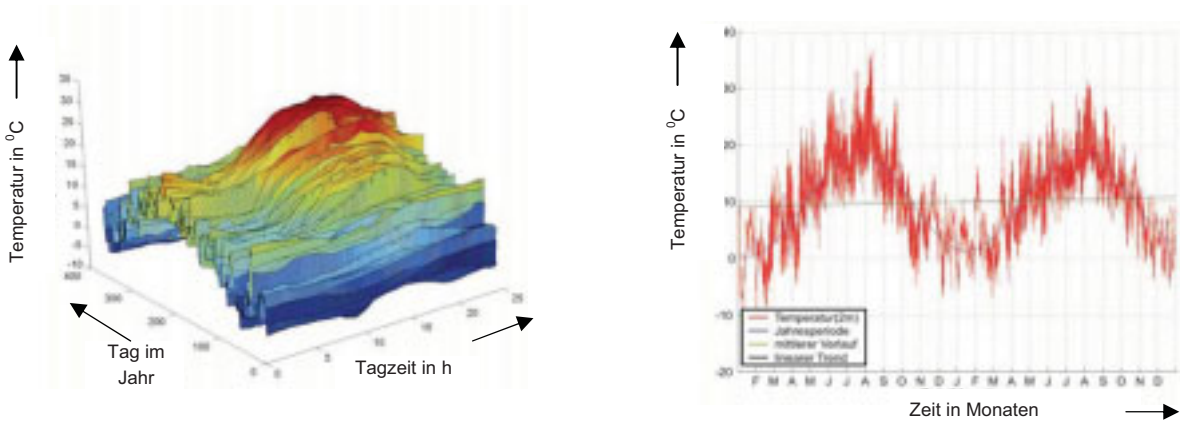


Abbildung 9-1: Temperaturverlauf

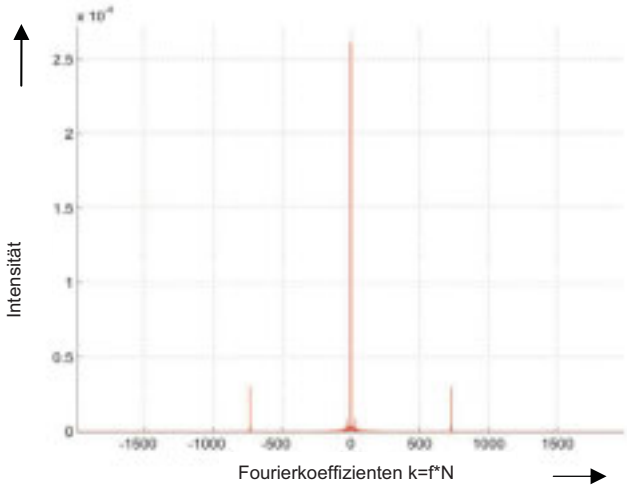


Abbildung 9-2: Periodogramm Temperaturverlauf

Mittelwert $\bar{x}$ in ^0C	Standardabweichung s in ^0C	Varianz s^2 in $^0C^2$
9,9580	7,6463	58,4665

Trendpolynome	
1,00	x
9,012189	0,000026955

	Periode (Tag)	Amplitude absolut in ^0C	Amplitude $s_n/s1$	Phase
s1	365,5000	8,0678	1,0000	-1,8396
s2	1,000	2,9030	0,3598	-2,4503
s3	22,1515	1,4900	0,1847	2,5392
s4	45,6875	1,0997	0,1363	-2,3989
s5	0,9973	1,0558	0,1309	0,8849

Tabelle 9-2: Ergebnisse Temperatur

9.8.2 Globalstrahlung

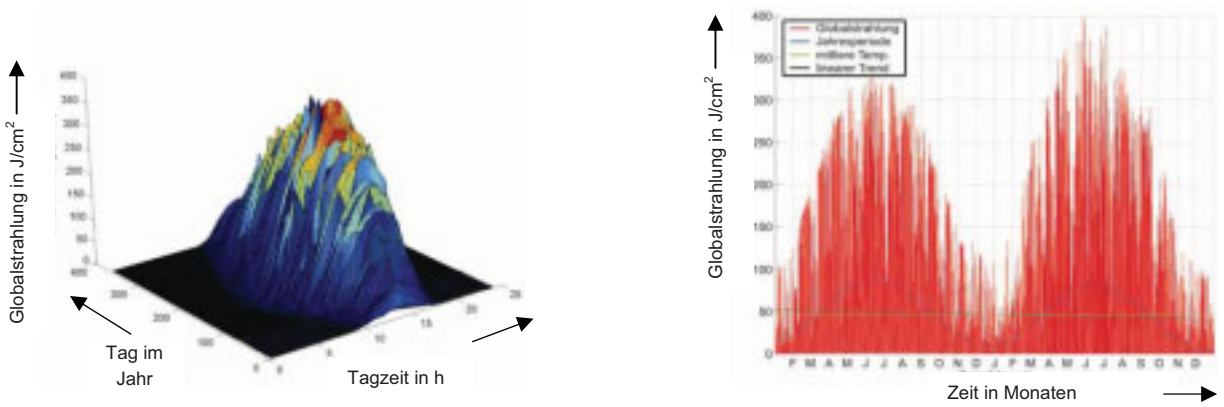


Abbildung 9-3: Globalstrahlungsverlauf

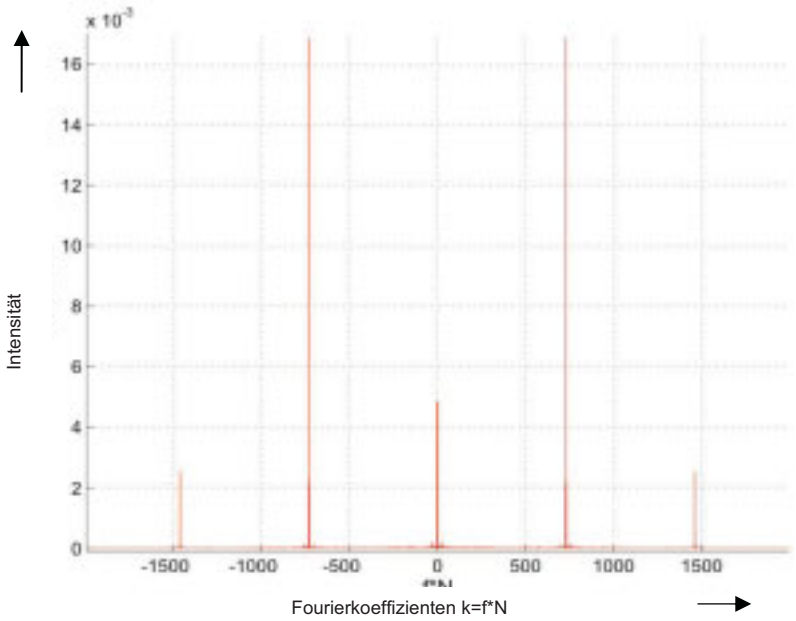


Abbildung 9-4: Periodogramm Globalstrahlung

Mittelwert $\bar{x}$ in J/cm²	Standardabweichung s in J/cm²	Varianz s^2 in $(J/cm^2)^2$
46,33	73,93	5466,37

Trendpolynome	
1,00	x
52,4246	-0,00017355

	Periode (Tag)	Amplitude absolut in J/cm²	$s_n/s1$	Phase
s1	1,0000	68,8671	1,0000	-1,8685
s2	365,5	36,9539	0,5366	-1,4105
s3	0,500	26,6154	0,3865	0,9902
s4	0,9973	25,4273	0,3692	1,4921

Tabelle 9-3: Ergebnisse Globalstrahlung

9.8.3 Windgeschwindigkeit

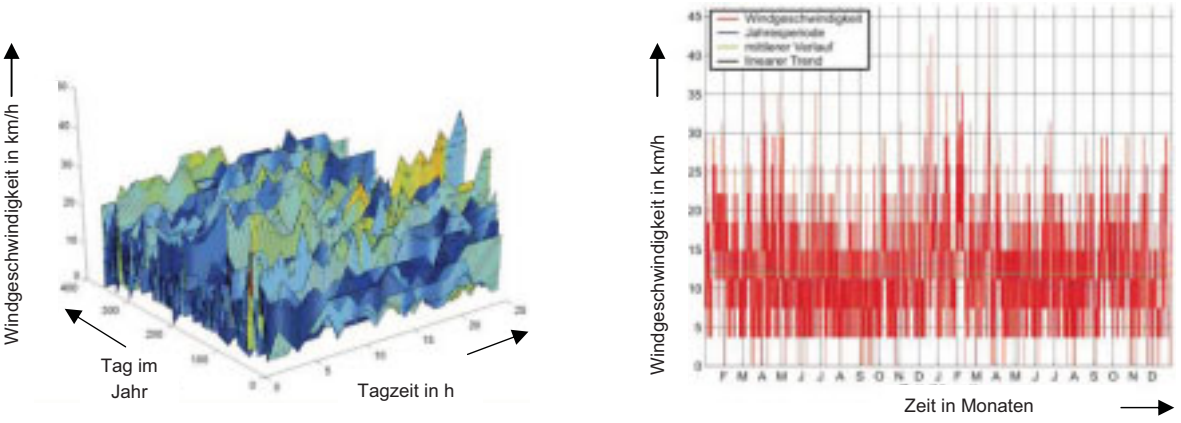


Abbildung 9-5: Windgeschwindigkeitsverlauf

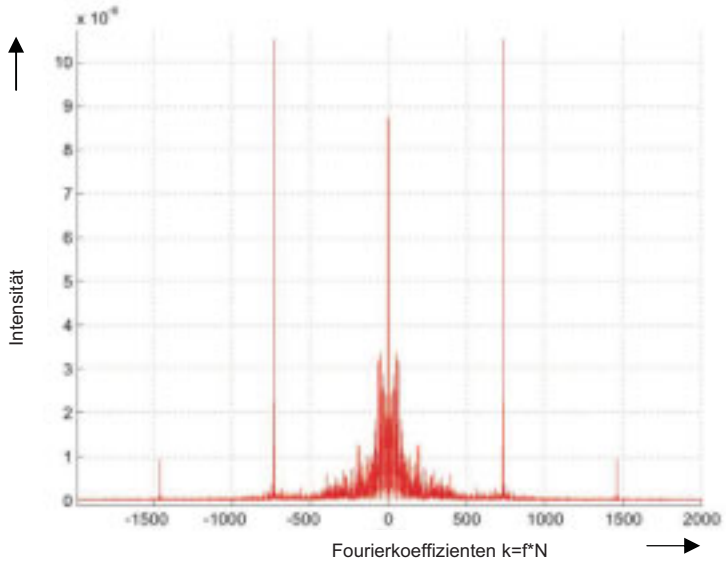


Abbildung 9-6: Periodogramm Windgeschwindigkeit

Mittelwert $\bar{x}$ in km/h	Standardabweichung s in km/h	Varianz s^2 in $(km/h)^2$
11,9180	5,8812	34,5889

Trendpolynome	
1,00	x
12,1954	0,000079077

	Periode (Tag)	Amplitude absolut in km/h	$s_n/s1$	Phase
s1	1,0000	1,7184	1,0000	-2,3591
s2	365,5	1,5679	0,9124	1,2521
s3	14,6200	0,9754	0,5676	-0,0906
s4	243,666	0,9715	0,5653	-1,8332

Tabelle 9-4: Ergebnisse Windgeschwindigkeit

9.8.4 Luftdruck

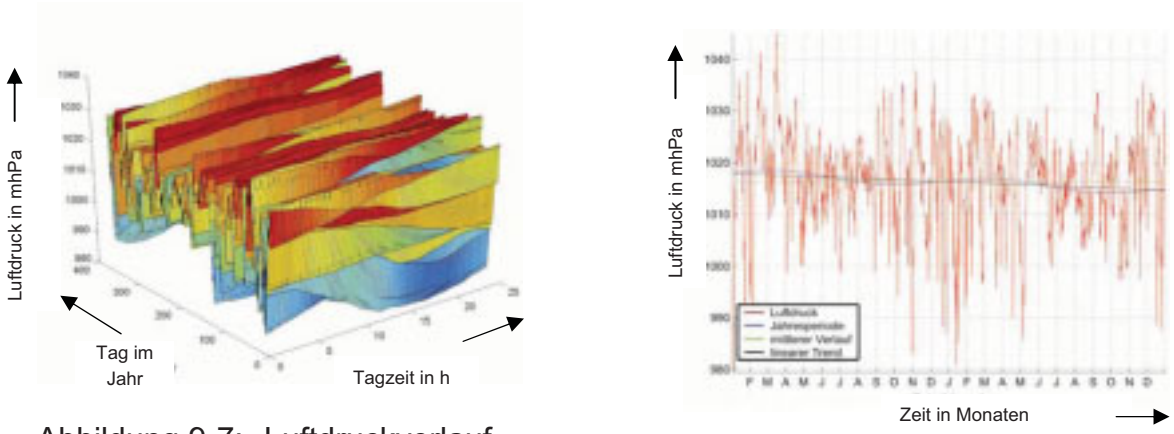


Abbildung 9-7: Luftdruckverlauf

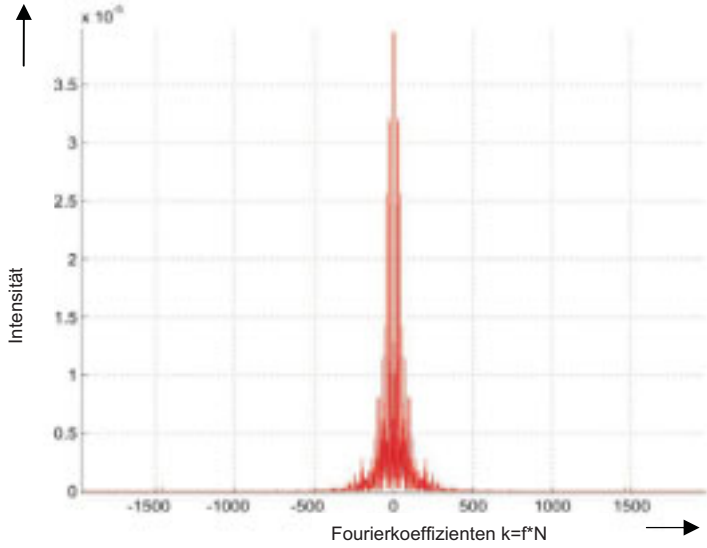


Abbildung 9-8: Periodogramm Luftdruck

Mittelwert $\bar{x}$ in mhPa	Standardabweichung s in mhPa	Varianz s^2 in $mhPa^2$
1016,268	9,6164	92,4751

Trendpolynome	
1,00	x
1017,8740	-0,00004574

	Periode (Tag)	Amplitude absolut in mhPa	$s_n/s1$	Phase
s1	91,3750	3,3281	1,0000	-2,7800
s2	23,5806	3,0069	0,9035	-0,2514
s3	30,4583	2,9849	0,8969	-0,7595
s4	16,2444	2,6838	0,8064	-2,8764

Tabelle 9-5: Ergebnisse Luftdruck

Lebenslauf

Persönliche Daten

Michael Fiedeldey
geboren am 01.05.1966 in Osnabrück
Familienstand verheiratet

Schulausbildung

1972 - 1977	Grundschule
1977 - 1979	Orientierungsstufe
1979 - 1983	Realschule
1983 - 1986	Fachgymnasium Technik in Osnabrück

Wehrdienst

07/1986 – 07/1987	Wehrdienst in München und Osnabrück

Berufsausbildung

08/1987 – 06/1990	KM - Kabelmetal AG Ausbildung zum Energieelektroniker
07/1990 – 09/1990	KM - Kabelmetal AG Elektroniker

Studium

10/1990 – 02/1997	Universität Hannover Elektrotechnik mit der Fachrichtung ‘Elektrische Energieversorgung’

Berufstätigkeit

04/1997 – 08/1997	Ingenieurbüro für Elektrotechnik H. D. Wallenhorst, Osnabrück
09/1997 – 02/1999	RWE Energie AG Traineeprogramm
03/1999 – 06/2001	RWE Energie AG Hauptabteilung Gas, Wasser, Fernwärme

07/2001 – 03/2007	Stadtwerke Osnabrück AG Leitung Netzführung, Netzbetrieb, Asset-Management ,Regulierungsmanagement und Netzwirtschaft
seit 04/2007	Allgäuer Überlandwerk GmbH Prokurist / Bereichsleiter Technik
seit 04/2007	AllgäuNetz GmbH & Co. KG Geschäftsführer
seit 01/2009	Erdgas Kempten-Oberallgäu Netz GmbH & Co. KG Geschäftsführer