# 1 Introduction

## 1.1 Problem statement and research questions

For the field of quality assurance, a comprehensive and holistic approach from production to marketing is of high importance. Within the fruit and vegetable sector, this is particularly important, as various factors during the production of fresh produce, such as mineral nutrition, climatic conditions, selection of variety, or maturity at harvest have a significant influence on the subsequent quality along the supply chain (Hewett 2006). Furthermore, the quality of fruit and vegetables changes immediately after harvest due to ongoing metabolic processes (Martínez-Romero et al. 2007). The quality of produce deteriorates at a different pace (Kong and Singh 2016) depending on various influencing factors during the further course along the supply chain (e.g., storage and transport conditions, packaging, mechanical handling, post-ripening processes). Postharvest quality assurance has to monitor the quality of fresh produce and counteract quality-reducing processes in order to keep potential losses along the supply chain as low as possible. Depending on the respective product, specific standards concerning the commercial quality of some fruit parameters must be met to allow the distribution via retail chains (UNECE 2019). In addition to these standards required by law, retail companies oftentimes impose requirements, which exceed statutory provisions.

The verification and assessment of quality of intact fruit can be made based on human senses, such as visual inspection of colors, smelling of volatile organic compounds or touching to control fruit firmness (Walsh 2006). However, these judgements are often subjective, therefore the quality of these evaluations highly depends on personal experience. The utilization of instrumental measurement methods for commercially important quality traits is often preferred over sensory evaluations and help to reduce the variation between individuals, provide high precision and offer a common language between consumers, researchers and industry (Abbott 1999). Furthermore, the determination of internal quality standards is often simply not possible using solely sensory assessments. In order to measure and verify these internal quality parameters, destructive measurements have to be conducted, e.g. the determination of sugar content using pressed fruit juice and a refractometer or measurement of firmness via penetrometer. Other measurements, such as the determination of dry matter, are extremely time-consuming, and in the case of acidity measurement, elaborate sample preparation and handling of chemicals is necessary for titration (OECD 2018).

In the past, various studies investigated the feasibility of optical measurement methods like visible and near-infrared (VIS/NIR) spectroscopy for the determination of important quality parameters of various agricultural and horticultural products (Nicolaï et al. 2007; Cortés et al. 2019). NIR spectroscopy, as a type of vibrational spectroscopy, employs the stimulation of molecular vibration using infrared light in the wavelength range 750 to 2500 nm. NIR electromagnetic waves interact with various molecular bonds (especially C-H, O-H, N-H or S-H bonds) of the respective sample constituents and result in a spectra, which is recorded by an optical sensor (Pasquini 2003). This NIR spectral data can be evaluated using various algorithms and multivariate statistical analysis and allows the development of qualitative and quantitative prediction models (Cozzolino 2009; Pasquini 2018). Whereas qualitative models are designed to enable classification, identification and authenticity verification of products (e.g., origin, variety, grade), quantitative models aim to determine the concentration of known constituents in samples, such as sugar concentration or dry matter content of fruit and vegetables (Figure 1). NIR spectroscopy offers various advantages compared to traditional measurement methods of fruit and vegetables, such as a fast and non-destructive operating principle, simple sample preparation as well as the possibility of using one spectrum for multidimensional evaluation of various parameters in one work step (McClure 1994; OECD 2018).

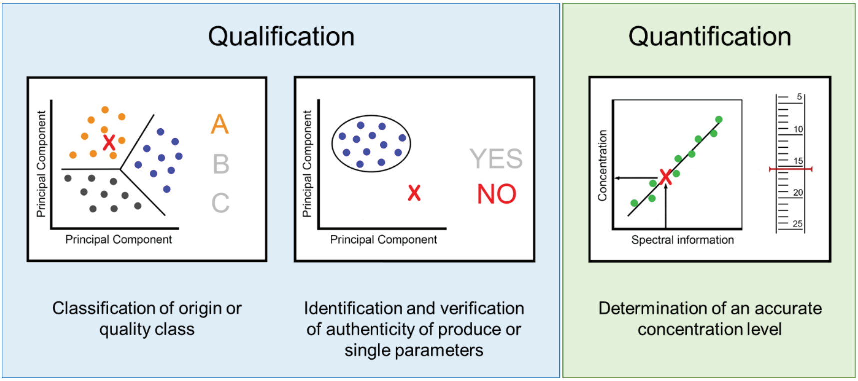


**Figure 1:** Application types of NIR spectroscopy (adapted and redesigned according to Langer (2019))

Next to laboratory applications, which employ benchtop instruments, the technique of NIR spectroscopy found its way into automated sorting and grading machines for fruit and vegetables. These on-line NIR applications are designed for real-time determination of important quality attributes (e.g., sugar content, acidity) and the detection of internal fruit disorders such as internal browning of fruit (Huang et al. 2008). The need for a more cost-effective and compact application of NIR spectroscopy led to the development of portable handheld devices (dos Santos et al. 2013). These sensors allow the transfer of traditional laboratory work to in-field applications, which can be implemented along the whole supply chain of fresh produce. Some companies are already marketing devices specifically designed for horticultural applications, such as the determination of fruit quality parameters in the orchard (Felix Instruments 2020b) as well as measurements along the supply chain to the point of sale (Sunforest 2020).

Subsequent technological developments led to an increasing miniaturization of NIR sensors. Some of these devices, so called food-scanners, were initially designed for end-consumers and were commercially distributed by various start-up companies (Consumer Physcis 2020; Tellspec Inc. 2020a; Spectral Engines Oy 2020). Oftentimes, these food-scanners work in combination with a mobile app on a smartphone or tablet and make use of cloud-based platforms, which hold vast material libraries and employ advanced algorithms. These applications aim to identify various food-related parameters (e.g. protein, fat, calories, allergens, contaminants, macronutrients) as well as the total energy content and provide important information such as food adulteration, food fraud, and food quality (Rateni et al. 2017). Currently, some of these start-up companies collaborate with industrial partners and target real-time in-field solutions for specific industries, e.g., monitoring of corn, grains and animal feed quality (Consumer Physcis 2020) as well as the analysis of fish, fruit and liquids (Tellspec Inc. 2020a).

Due to the novelty of portable and miniaturized devices, scientific research about the performance and suitability of food-scanners for a reliable determination of the quality of fruit and vegetables is currently still sparse. Initial studies compared various commercially available instruments for their prediction accuracy of specific fruit quality traits such as dry matter (Kaur et al. 2017) or examined the performance of individual devices (Ncama et al. 2018; Li et al. 2018; Choi et al. 2017) . Most of these studies regarding the performance of food-scanners evaluated selected fruit within a laboratory setup without reference to the practical use of these devices in daily processes of quality control along the fresh produce supply chain (FSC). At the beginning of this work, the use of miniaturized sensors for the concrete determination of fruit quality and their application along the FSC had not yet been investigated. In this context, the overall aim of this study was to evaluate the suitability of portable NIR sensors (food-scanners) for the determination of fruit quality along the supply chain of fruit and vegetables. This thesis focused on tomato as model fruit for the development of non-destructive prediction models of a variety of important quality attributes and evaluated the practicability of already existing and commercially available portable and miniaturized food-scanners.

Given this background information, the research questions for this doctoral thesis were formulated as follows:

- Miniaturized NIR instruments (food-scanners) are relatively new to the scientific literature. Is their performance comparable to traditional laboratory benchtop NIR instruments for the prediction of fruit quality traits?
- Currently there are various different food-scanners commercially available. Are there major differences with respect to the predictability of fruit quality parameters?

- With regard to the implementation of these novel devices in processes of quality control along the fresh fruit supply chain, are there differences with respect to the practicalness of commercially available food-scanners?
- Due to their working principle, food-scanners could determine multiple internal fruit quality parameters non-destructively. Is there a demand for these devices from the fresh produce industry?
- Can these food-scanners, which have primarily been used in this doctoral thesis to investigate the performance on predicting tomato fruit quality traits, also be used for a wide-ranging applicability of various types of fruit and vegetables and processes along the FSC?

Based on these research questions, the following hypotheses were phrased for this doctoral thesis:

H1. Actors along the FSC see a great potential for the application of food-scanners;
H2. There are different areas of application for food-scanners along the FSC;
H3. Prediction models derived from food-scanner spectra show comparable results than those computed with benchtop NIR instruments;
H4. Commercially available food-scanners achieve similar prediction model accuracy of selected fruit quality parameters;
H5. Software solutions provided by some food-scanner manufacturers achieve prediction model performance comparable to state-of-the-art multivariate analysis software
H6. Currently available devices differ in practicability and applicability with respect to the requirements of the fruit and vegetable supply chain;
H7. Food-scanners can be used for the prediction of various quality-related parameters on a wide range of fruit and vegetables;

## 1.2 Thesis structure and objectives

This thesis is structured into nine chapters, which comprise the objectives set for this study in order to answer the above stated research questions, and to corroborate the derived hypotheses (Figure 2):

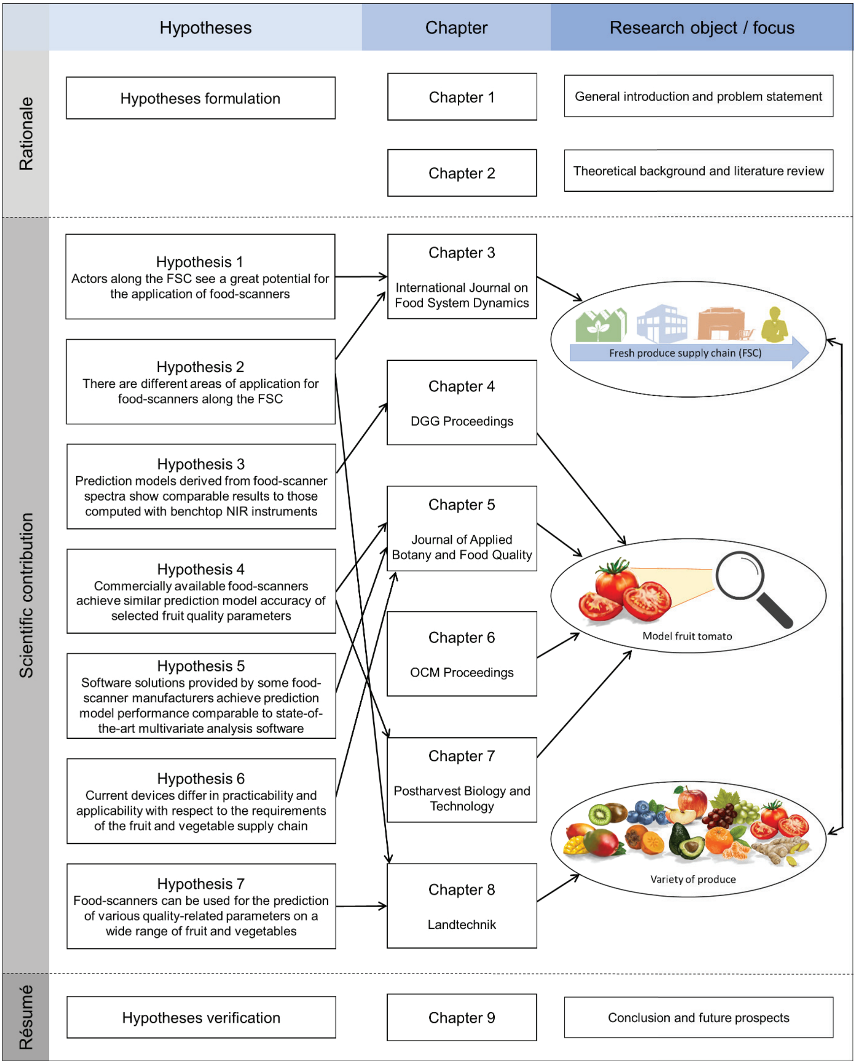


**Figure 2:** Overview of the structure of this doctoral thesis

Chapter 1 serves as introduction and states the status quo and motivation of this thesis. Additionally, the research questions as well as the hypotheses of the doctoral thesis are outlined.

Chapter 2 provides the theoretical background on near infrared spectroscopy and its application within the field of fruit and vegetables. Additionally, the development from laboratory to portable and miniaturized sensors is outlined and an extensive literature review of scientific research utilizing these new devices provided.

Chapter 3 investigates the status quo on how measurement of fruit quality is currently performed on various levels along the fruit and vegetable supply chain. Furhtermore it presents the opinion of experts along the supply chain on the application of food-scanners for non-destructive quality assessment and thereby justifies the basis for this work. This work has been published in 2020 in the International Journal on Food System Dynamics 11 (2), pages 101-106.

Chapter 4 compares the prediction accuracy of a miniaturized food-scanner and a conventional laboratory spectrometer using selected quality parameters of tomato fruit. The results of this work have been presented at the Horticultural Science Conference in Geisenheim 2018 and were issued in the DGG Proceedings 8 (13), pages 1-5.

Chapter 5 evaluates the capability and practical suitability of three commercial food-scanners for non-destructive measurement of important quality parameters using the model fruit tomato. Additionally, prediction models derived from software solutions provided by manufacturer are compared to state-of-the-art software for multivariate analysis. This chapter furthermore illustrates the possibility of utilizing food-scanners for quality control of pre- and postharvest processes. The findings have been published in 2020 in the Journal of Applied Botany and Food Quality 93, pages 204-214.

Chapter 6 examines the possibility of shelf-life assessment of tomatoes using a miniaturized NIR device. For this purpose, storage trials were carried out in combination with NIR measurements on different tomato varieties. Mathematical models were computed to estimate the tomato firmness over the storage period. The findings of this research have been presented at the 4$^{th}$ International Conference on Optical Characterization of Materials in Karlsruhe in 2019 and were subsequently published in the OCM Proceedings, pages 1-12.

Chapter 7 investigates the measurement of secondary plant metabolites with the help of food-scanners using lycopene content in tomato as example. Three commercially available food-scanners are tested and compared with respect to their accuracy and performance to each other as well as to alternative measurements of lycopin by using color values derived from a colorimeter. The results have been issued in 2020 in Postharvest Biology and Technology 167 (111232).

Chapter 8 evaluates the performance of three portable and miniaturized food-scanners with regard to their predictive accuracy of important quality parameters on a wide range of produce from the fruit and vegetable assortment. Additionally, results of a first in-field practical application of a food-scanner for quality control during incoming goods control are presented and further areas of application are discussed. The findings of this work have been published in 2021 in the issue (76) 1 of the journal Landtechnik, pages 52-67.

Chapter 9 provides a comprehensive discussion and links the hypothesis to the findings obtained in this thesis. Future prospects of food-scanner applications along the supply chain and remaining challenges for their implementation are exposed.

# 2 Theoretical background and literature review

## 2.1 Operating principle of near infrared spectroscopy

The near infrared (NIR) radiation is located between the visible and mid infrared part of the electromagnetic spectrum (Figure 3), which relates to the wavelength range of 780 to 2500 nm (Sandorfy et al. 2007). NIR radiation, due to its nature, acts like a wave and can be described with the two properties frequency and wavelength. NIR directed towards an object (e.g., liquid, solid) is absorbed by molecules, which as a result leads to the vibration of chemical bonds between atoms within these molecules (Osborne 2006). In a simplified approximation, these vibrations react as a harmonic motion, which can be described using a diatomic oscillator (Osborne 2006). Based on this simple vibrating system, additional concepts have been developed to more complex polyatomic molecules. The fundamentals and physical principles of these oscillating motions of di- and polyatomic molecules have been described in the literature in great detail (Bokobza 2002; Sandorfy et al. 2007).

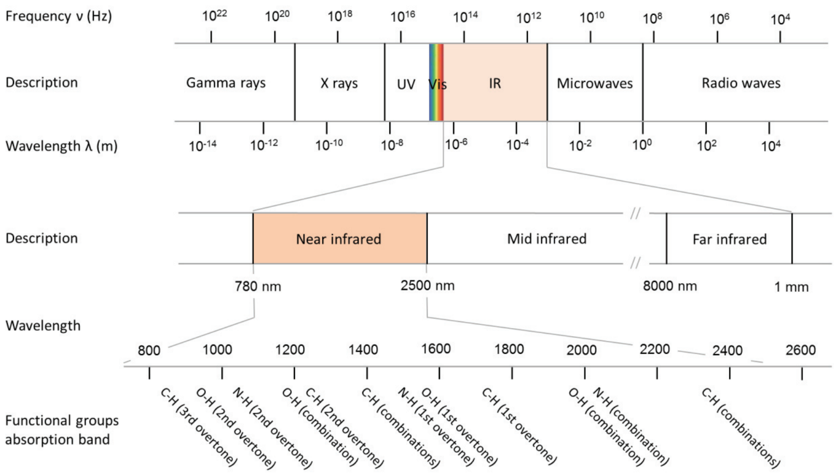


**Figure 3:** Location of near infrared radiation on the electromagnetic spectrum and approximate locations of absorption bands of important functional groups (own illustration based on Osborne (2006), Xiaobo et al. (2010))

Samples containing organic matter display distinct spectral fingerprints in the NIR region due to relatively prominent absorption of overtones and combination modes related to several functional groups (Figure 3) typically present in organic compounds, e.g., aliphatic and aromatic (C-H), carboxyl (C-O), hydroxyl (O-H) and amine and amide (N-H) groups (Xiaobo et al. 2010). Therefore, NIR is sometimes called the “overtone region” (Sandorfy et al. 2007). Due to the chemical composition of organic molecules and their particular absorption patterns in the NIR region, NIR radiation is suitable for the analysis of organic material and samples. Provided that the frequency of the radiation matches the frequency of the vibrating molecule, a transfer of energy from radiation to the molecule occurs, which subsequently can be measured and displayed as a plot of energy vs. wavelength, called a spectrum (Osborne 2006).

In some cases, in addition to NIR radiation, light visible to the human eye (VIS) is used for spectroscopic applications. In that case it is referred to as VIS/NIR spectroscopy. For reasons of simplification, only the term NIR spectroscopy will be used in the following sections.

## 2.2 Measurement modes in NIR spectroscopy

With regard to the sample composition and the trait of interest for analysis there are four different modes of measurement available in NIR spectroscopy: transmittance, reflectance, interactance and transflectance (Figure 4). The main operation principles and differences between these modes are as follows (Tsuchikawa 2007; Pasquini 2003): In transmission mode, the sample (S) is illuminated perpendicular from one side by a light source (L) and the transmitted light is subsequently detected on the opposite side of the sample with a detector (D). In a similar manner, the sample is illuminated perpendicular to the sample surface in reflection mode. However, in contrast to transmission, the reflection mode employs reflected light from the sample surface, in which the incident light, after illumination, underwent various effects such as scattering, absorption and transmission. Reflection measurements are oftentimes conducted in a specific distance (d) between illumination source and sample, whereas the reflected light is in general collected by the detector in a fixed angle (e.g., 45°) and distance to the sample.

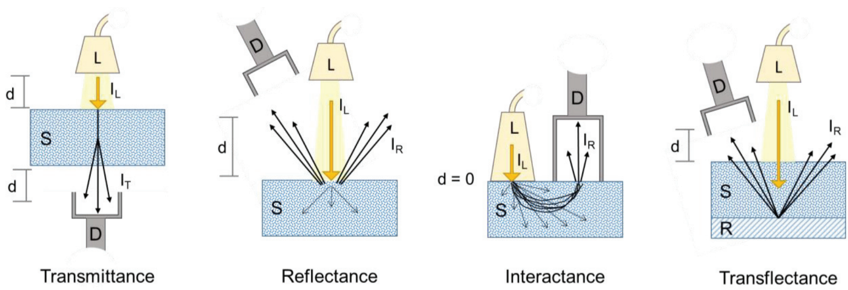


**Figure 4:** Different modes of measurement used in NIR spectroscopy, where $I_L$ = incident light, $I_T$ = transmitted light, $I_R$ = reflected light, D = detector, R = reflector, S = sample, d = distance (own illustration based on Tsuchikawa (2007), Pasquini (2003), and Saranwong and Kawano (2007))

Interactance mode usually makes use of a specifically designed interaction probe, which generally consists of an illuminator and detector both in direct contact with the sample surface but arranged in certain distance to each other. Usually the light source and detector are positioned in a parallel setup so that the light, due to specular reflection, cannot directly be detected (Nicolaï et al. 2007). As a result, only the light directly transmitted through the sample is detected, and more information about the actual sample composition and sample constituents can be obtained (Pasquini 2003). The transflection mode constitutes a combination of transmission and reflection, where the illuminated light, after passing through the sample, is scattered back from a reflector (R) and passed through the sample a second time (Tsuchikawa 2007; Pasquini 2003). The reflected light is then collected by a detector closely positioned to the sample. A detailed explanation of the use of the different NIR measurement methods for agricultural and horticultural products analysis can be found in the following chapter 2.5.

## 2.3 Chemometrics

Due to the complexity of organic material, the crucial information present in NIR spectra about the chemical composition is scattered over the whole spectrum and generally not directly available for analytical purpose. The evaluation of NIR spectra requires a special type of data analysis, which simultaneously recognizes changes in the whole or part of the spectrum as a function of the changes in the sample properties or/and analytes contents (Pasquini 2018). This technique is called multivariate analysis or chemometrics (Brown 2016; Cozzolino 2009), which since its development in the middle of the 1970 became “one of the sustaining pillars of modern NIRS” (Pasquini 2018). The application of chemometric techniques helps to obtain quantitative as well as qualitative relations between NIR spectra and the trait of interest (Figure 1).

Partial least squares regression (PLSR), a method which assumes a linear relationship between the spectral vibration and the modeled parameter or concentration (Pasquini 2018), is commonly used for quantitative evaluations, e.g., the estimation of soil element contents (Yu et al. 2016), the determination of the concentration of quality parameters in alcohol (Valderrama et al. 2007) and the analysis of sugar, acid and phenol content of fruit (Guo et al. 2016). In recent years, several additional techniques based on PLSR algorithm have been developed, e.g. nonlinear iterative partial least squares (NIPALS), sparse partial least squares (SPLS) and mixed-norm partial least squares (MNPLS) (You et al. 2016).

Qualitative evaluations are used in order to classify and authenticate product origin and quality as well as identify product adulterations and falsifications. The most prominent methods include linear discriminant analysis (LDA), cluster analysis (CA), soft independent modeling of class analogy (SIMCA), k-nearest neighbor (KNN), support vector machine discriminant