Development of a UV Hyperspectral Imaging Prototype for Industrial Applications

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

M.Sc. Mohammad Mahmoud Ahmad Al Ktash

aus Al Ramtha/Jordanien

Tübingen

2023

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	16.05.2024
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter/-in:	Prof. Dr. Marc Brecht
2. Berichterstatter/-in:	Prof. Dr. Alfred Meixner

Kurzfassung

Baumwollfasern sind aufgrund ihrer Weichheit, Haltbarkeit und Saugfähigkeit für die Textilindustrie unverzichtbar. Allerdings wirken sich Verunreinigungen negativ auf die Baumwollqualität aus und es ist notwendig die Stärke der Kontamination zu bestimmen. Eine häufig auftretende Verunreinigung ist Honigtau, welcher zu klebriger Baumwolle führt, und erhebliche Probleme für die Textilindustrie verursacht. Verschiedene Methoden im sichtbaren (Vis) und nahen infraroten (NIR) Spektralbereich werden für Qualitätskontrollen und Sortierverfahren eingesetzt, während der ultraviolette (UV) Bereich bisher kaum genutzt wird.

In den letzten Jahren haben hyperspektrale Bildgebungssysteme aufgrund ihrer Multimodalität, ihre räumliche Auflösung und ihrer Fähigkeit zur quantitativen Analyse im Vergleich zu herkömmlichen Verfahren zunehmend an Aufmerksamkeit gewonnen. Diese Vorteile haben sie für verschiedene Anwendungen, z. B. in der Textilindustrie, sehr attraktiv gemacht.

Das Hauptaugenmerk der vorliegenden Arbeit liegt auf der Erkennung von Honigtaukontaminationen und der Entwicklung eines hyperspektralen Bildgebungssystems im UV-Bereich. Der Aufbau basiert auf einem Spektrographen, der mit einer CCD-Kamera verbunden ist. Die Proben werden auf ein Förderband gelegt, welches die Probe unter der hyperspektralen Kamera bewegt. Dieses Verfahren wird als Pushbroom Imaging bezeichnet. Je nach Anwendung wurden Xenon- oder Deuteriumlampen zur Beleuchtung verwendet, wobei Deuteriumlampen eine höhere Beleuchtungsstärke im UV-C-Bereich im Vergleich zur Xenon-Bogenlampe bieten. Zur Validierung dieser neuartigen Bildgebungseinrichtung wurde eine Reihe von bekannten Substanzen wie aktive pharmazeutische Wirkstoffe (APIs) und Schmerzmittel verwendet. Diese Proben waren Ibuprofen, Acetylsalicylsäure und Paracetamol. Die Ergebnisse wurden mit lokalaufgenommenen Einzelspektren verglichen und mittels multivariater Datenanalyse ausgewertet. Es wurde gezeigt, dass die hyperspektrale Bildgebung im UV-Bereich zuverlässige Ergebnisse erzielt und eine analytische Methode wurde entwickelt, um kommerzielle Schmerzmitteltabletten mit dem neuen Prototyp zu identifizieren. Anschließend wurde eine separate Probenreihe, einschließlich direct bonded copper (DBC) Substrate, für eine sekundäre Bewertung getestet. Der entwickelte Prototyp ist in der Lage wenige Nanometer dicke Oxidschichten, zu erkennen. Dabei können verschiedene Oxidationszustände unterschieden werden, sogar nachdem die Proben vorgesehene Reinigungsverfahren durchlaufen haben. Im nächsten Schritt wurden Baumwollproben aus verschiedenen Ländern verwendet und mittels Vis/NIR hyperspektraler Bildgebung untersucht. Die gewonnenen

Daten wurden mit lokal aufgenommen Einzelspektren verglichen und mittels multivariater Datenanalyse analysiert. Die Ergebnisse zeigen, dass es möglich ist, anhand einiger ausgewählter Wellenlängenbereiche zwischen verschiedenen Baumwollsorten zu unterscheiden. In einem letzten Schritt wurde die Quantifizierung von Honigtaukontaminationen auf Baumwolle durchgeführt. Hierfür wurde ein Verfahren zur Kalibrierung des Prototyps für hyperspektrales Imaging im UV-Bereich entwickelt und mit realen Proben getestet. Mechanisch gereinigte Baumwollproben wurden in eine Lösung getaucht, die Zucker und Eiweiß in bekannten Konzentrationen enthielt, um mit Honigtau verunreinigte Baumwolle zu imitieren. Diese Proben wurden nach 44 Stunden und nach einem Monat untersucht. Anhand dieser Proben wurde der Prototyp erweitert und optimiert. Die gewonnenen Daten wurden chemometrisch analysiert, um ortsaufgelöst die Honigtaumengen in Baumwollproben mit unterschiedlichen Mengen der Substanz erfolgreich vorherzusagen. Zusammenfassend lässt sich sagen, dass die Menge von Honigtau auf Baumwolle lateral aufgelöst quantifiziert wird. Dafür wurde ein Prototyp für hyperspektrale Bildgebung im UV-Bereich entwickelt, der für industrielle Anwendungen geeignet ist.

Die Ergebnisse zeigten, dass die hyperspektrale Bildgebung mehrere Vorteile gegenüber etablierten Bildgebungsverfahren oder der klassischen ortsaufgelösten Spektroskopie bietet, z. B. die laterale Auflösung, die Fähigkeit, Proben zerstörungsfrei zu analysieren, Stoffe in sehr geringen Konzentrationen zu erkennen und Stoffe selbst dann zu identifizieren, wenn sie mit anderen Stoffen vermischt oder durch diese überlagert sind. Außerdem ist sie sehr empfindlich und kann kleinste Veränderungen in der chemischen Zusammensetzung von Materialien in Abhängigkeit der Zeit erkennen.

Abstract

Cotton fiber is essential for the textile industry due to its softness, durability, and absorbency. Therefore, the assessment of the cotton quality is needed, which is determined by the degree of contamination. The predominant contaminants in raw cotton come from insects that excrete sugars called honeydew during feeding. Cotton contaminated by sugar causes significant problems for textile equipment. Honeydew is the most common source of sticky cotton. However, various methods in visible (Vis) and near-infrared (NIR) spectral ranges are regularly used for quality control and sorting procedures, while the ultraviolet (UV) range has not been widely used. In recent years, hyperspectral imaging systems have gained increased attention over traditional techniques due to their multi-modality, spatial resolution, and ability for quantitative analysis. These advantages have made them highly attractive for various applications, such as in the textile industry.

The main goal of this work is to develop a method to detect honeydew contamination in the UV range. For this purpose, a UV hyperspectral imaging system based on a spectrograph connected to a CCD camera was constructed. The samples were placed on a conveyor belt, which moved them underneath the hyperspectral imaging camera. This technique is called pushbroom imaging. Depending on the application, either Xenon or Deuterium lamps were used for illumination since Deuterium lamps provide a higher illumination strength in the UV-C region compared to the xenon-arc lamp. In order to validate this novel imaging setup, a set of well-known substances, such as active pharmaceutical ingredients (APIs) and painkillers, was used. These sample are ibuprofen, acetylsalicylic acid, and paracetamol. The results were compared with single-point spectroscopy and analyzed using chemometric data analysis. It was shown that the hyperspectral imaging achieved reliable results, and an analytical method was developed to identify commercial painkiller tablets with the new prototype. Subsequently, a separate sample set, including direct bonded copper (DBC) sheets, was tested for a secondary evaluation. The developed prototype is able to detect very thin oxide layers, as thin as a few nanometers. It can also distinguish between various oxidation states via a cleaning procedure for DBC samples. Consequently, cotton samples from different countries were investigated using Vis/NIR hyperspectral imaging. The data obtained were compared to that obtained from single-point spectroscopy and analyzed using multivariate data analysis. The results indicate that it is possible to distinguish between different cotton

types based on specific wavelength ranges. In the last step, the quantification of honeydew contamination on cotton was determined. A calibration procedure was developed using mechanically cleaned cotton samples. These samples were immersed in different concentrations of sugar and protein to mimic cotton contaminated with honeydew. Consequently, they were analyzed after 44 hours and one month. Further improvements were made to the UV hyperspectral imaging setup in the later measurement. The data obtained were analyzed using chemometrics to predict the local quantities of honeydew on cotton samples successfully. In conclusion, the present work aims to quantify the spatial amount of honeydew contaminated on cotton by developing a hyperspectral imaging prototype in the UV region that is advantageous for industrial applications.

The results showed that hyperspectral imaging has several advantages over established analytical techniques, such as lateral resolution, the ability to analyze samples non-destructively, detect materials at very low concentrations, and identify materials even when mixed or obscured by other materials. It is also highly sensitive and can detect subtle changes in the chemical composition of materials over time.

Table of Contents

1 Introduction

1.1 Spectroscopy

Spectroscopy is a method of studying the interaction between light and objects. It helps us understand the properties of different materials based on the absorption or emission over a specific wavelength or energy range [1,2]. Different spectroscopic methods can be used to solve a wide range of analytical problems. These methods differ based on the specific substances or species that need to be analyzed, such as atomic or molecular spectroscopy [3]. It utilizes the ability of molecules and atoms to absorb, emit, and scatter a type of energy called electromagnetic (EM) radiation. EM radiation is a form of energy that covers a wide range of wavelengths and frequencies, ranging from cosmic radiation at 10^{-14} m to infrasonic radiation at 10^{10} m. Spectroscopy can be applied to different frequency ranges, such as Ultraviolet-Visible/Near-infrared (UV-Vis/NIR) and nuclear magnetic resonance (NMR) spectroscopy. UV-Vis/NIR spectroscopy is one of the most commonly used spectroscopic techniques, ranging from 100 nm to 2500 nm [4,5]. The absorption or emission of various forms of EM radiation is associated with different types of transitions. UV-Vis and NIR are associated with electronic transition and molecular vibration [6].

The absorption process is defined as the transfer of EM energy to atoms or molecules. Electrons in the atoms are excited from a lower to a higher energy state, which exists on discrete levels. Absorption occurs when the energy of an exciting photon matches the energy difference between the ground and excited state. These energy differences are unique, providing a means of characterizing a compound or material [7,8]. The following formula gives the energy *E*, which is related to the absorption bands:

$$E = E_{electronic} + E_{vibrational} + E_{rotational} \quad (1.1)$$

Where the molecules' electronic energy is given by $E_{electronic}$, the vibrational and rotational energy is described as $E_{vibrational}$ and $E_{rotational,}$ respectively (see Figure 1.1).

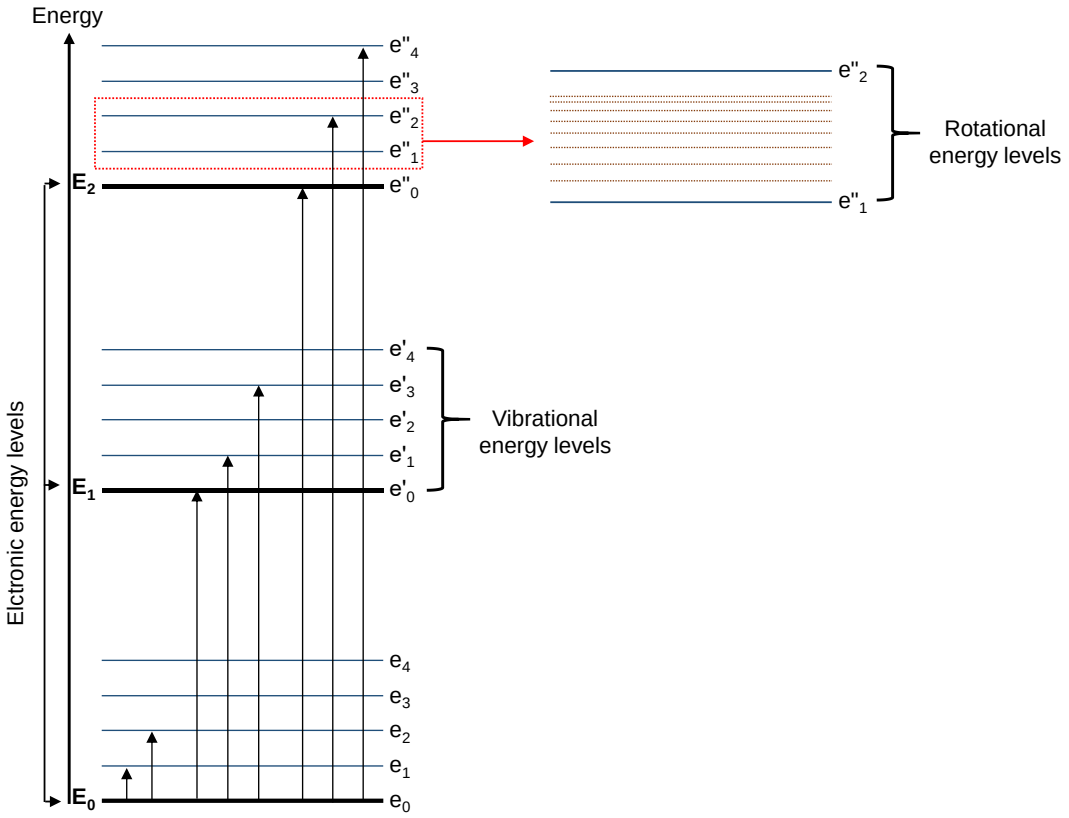


Figure 1.1: Energy level diagram illustrates electronic, vibrational, and rotational energy.

Absorption occurs, for example, when light passes through a solution. Light absorption is defined by Beer-Lambert law and is commonly used in spectroscopy [9]. Beer-lambert's law can be used for pure samples or samples that do not significantly scatter light. It defines the amount of the energy absorbed *A* or transmitted from the solution as proportional to the molar absorptivity co-efficient ε and the concentration of the solute c the optical path length in cm ℓ. The following formula gives the relationship [10]:

$$A = \varepsilon \ell c = -\log_{10}\left(\frac{I}{I_0}\right) \tag{1.2}$$

Although it can also be written in terms of intensities, where I is the light intensity passing through the sample cell, and the initial light intensity is defined by I_0.

Often, samples have varying sizes and shapes and can be either transparent or in-transparent in specific wavelength regions due to absorption characteristics and refractive index in the case of transparent samples or scattering in the case of in-transparent samples. Therefore, many spectral systems acquire the reflected light from the samples. Reflection of light can be classified into two types: classical reflection, such as mirrors, and diffuse reflection from rough surfaces. Diffuse

reflection is described by Lambert's law and is observed in materials such as milk and Spectralon®. Figure 1.2 illustrates both classical and diffuse reflected light [11].

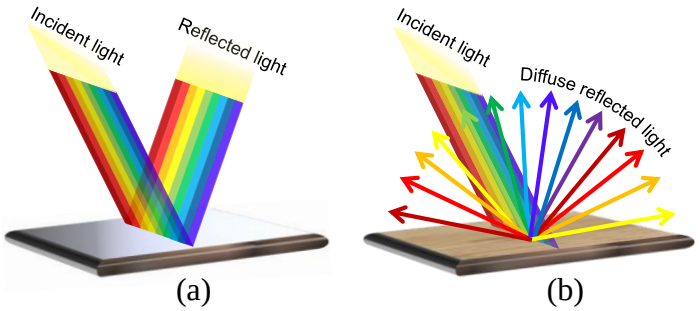


Figure 1.2: (a) Classical and (b) diffuse reflected light.

1.2 Hyperspectral imaging

Optical spectroscopy is a widely used analytical technique in various fields, such as the internal quality analysis of biological samples, due to its rapid, non-destructive, and environmentally friendly nature [12,13]. It often requires minimal sample preparation and handling. These optical methods can be applied as on-line or in-line techniques [14-16]. Over the years, optical spectroscopy has become an essential traditional method in laboratory work and is considered a reference method across a wide range of wavelength regions, from the far ultraviolet to the infrared, including spectral imaging techniques [17,18]. Various instrumental techniques have been developed based on these fundamental principles of optical spectroscopy, such as hyperspectral imaging [19].

Hyperspectral imaging is a technique that involves capturing and analyzing images of an object across a wide wavelength range, typically in the Vis and NIR and, recently, in UV regions [17,19,20]. It allows for identifying materials and compounds within an image by analyzing those materials' unique "spectral fingerprints". The combination of spectroscopic techniques with imaging is, therefore, a fast and cheap option with the ability to cover the whole production, which has not been used so far. This method can operate as on-line or in-line technique [15,21].

Hyperspectral imaging has the potential to meet the requirements for spatial resolution and large-area detection. It enables the rapid spatially resolved spectral analysis of surfaces. This technique is well described in the literature for applications in medicine, food, mineralogy, agriculture, and environmental monitoring [22-27].

The imaging system can be designed in different modes of imaging. The most common are multispectral and hyperspectral imaging. The main difference between these imaging modes is in the number of color channels, and the resolution see Figure 1.3 [28,29]. Multispectral imaging has 4 to 20 color channels and collects non-continued images from specific bands. In contrast, hyperspectral imaging continuously captures tens to hundreds of spectral bands [30].

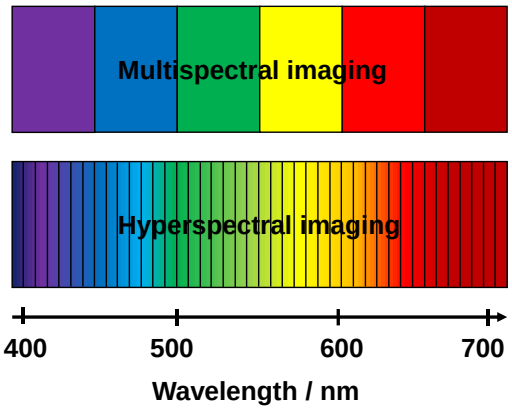


Figure 1.3: Schematic shows the difference between multispectral and hyperspectral.

Hyperspectral imaging is based on one of four different technologies to acquire complete spectral information [30]. Figure 1.4a shows a point-by-point technique known as a whiskbroom or spotlight sensor. It works like a broom sweeping across an area by moving a sensor back and forth to collect data. The sensor collects the data from one pixel in the image at a time, which can obtain a high spectral resolution. Therefore, discrimination between molecules can be acquired. Originally, this technique was used in satellite technology for scanning the earth by the Earth Resources Technology Satellites, and the Airborne Imaging Spectrometer later adopted it. Figure 1.4b shows a new strategy called snapshot. It captures a 3D area (x, y and λ) and spectral information into one single measurement by using an image mapper and prism array. Unlike the other techniques, it uses one exposure to record spatial and spectral information. Therefore, there is no need to scan at all. This technique is suitable for capturing instantaneous images of a scene or object and is often used in applications such as digital photography. Figure 1.4c shows an image-by-image or area-scanning system, also called staring imaging. The staring technique captures the 2D area (x, y) while changing a filter in front of the camera. The light passes through the optics, then it is filtered, which produces a narrowband of the spectrum [18]. It is used in many applications, including surveillance and astronomy, for capturing detailed and steady images of specific targets or regions of interest. Figure 1.4d shows the line-by-line technique called pushbroom. Pushbroom, named along-track scanner based on an on-line scanning system, acquires complete spectral information for each pixel in the line. These data result in a three-dimensional (3D) data matrix with

dimensions x, y, and λ, often referred to as a hypercube. The data of all these technique provides high spectral resolution and is highly stable, making it useful for industrial quality control [31].

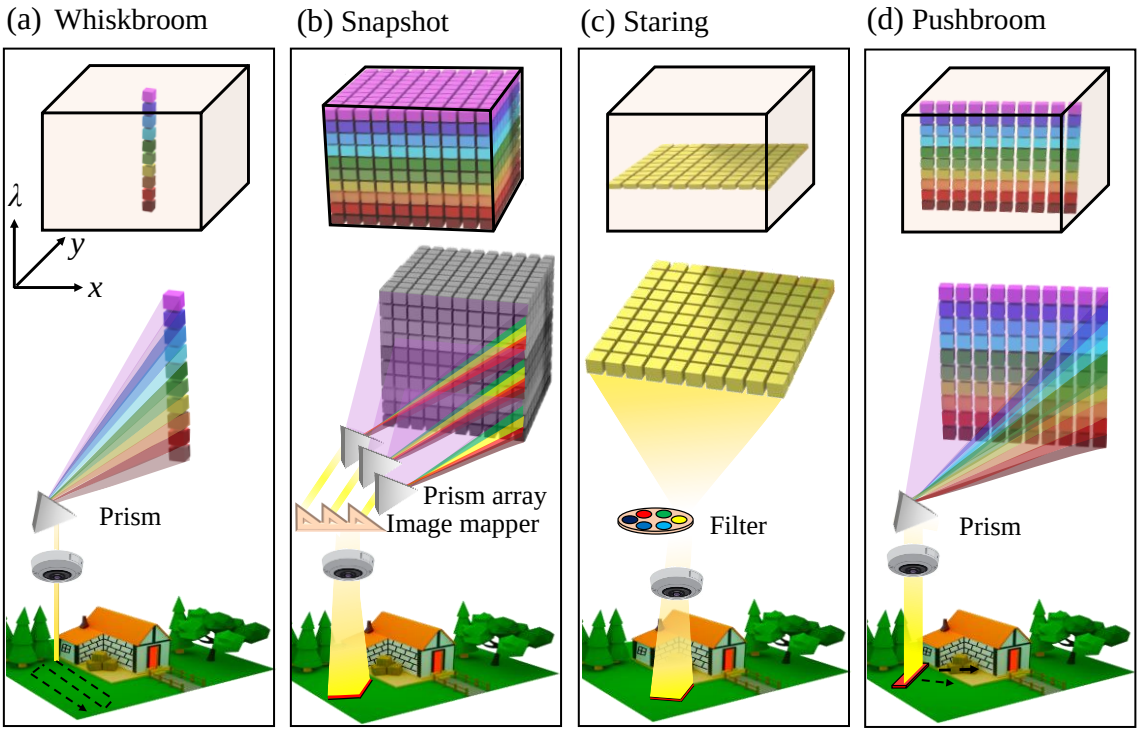


Figure 1.4: Visualization of the different imaging technologies: (a) Whiskbroom imaging (single point scanning) (b) Snapshot imaging (c) Staring imaging (2D scanning) (d) Pushbroom imaging (line scanning).

Such inspection systems require a minimum of sample preparation and can scan several samples swiftly with high spectral resolution [23]. Hyperspectral imaging systems based on a pushbroom scanner and a conveyor belt are commonly used in industrial applications [27,32-34]. The data is collected with the camera placed perpendicular to the conveyor belt. As the conveyor belt moves, the pushbroom scanner scans each line of the scene, and these lines are captured and combined to form a complete image. Figure 1.5 illustrates the scheme and general setup of a pushbroom scanner with a conveyor belt.

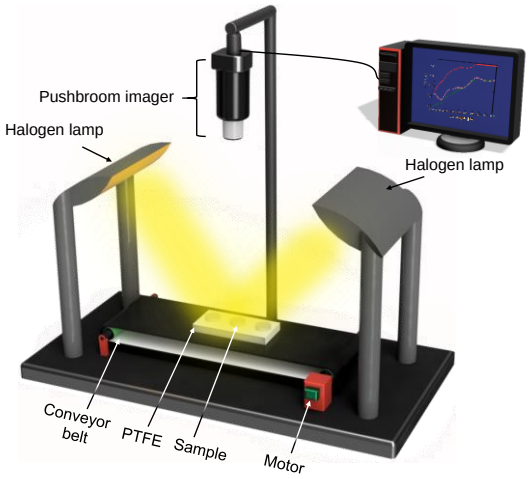


Figure 1.5: Scheme of Vis/NIR hyperspectral imaging (pushbroom).

Such systems have been primarily developed in the Vis/NIR region [23,35,36], but there is further information on quality control in the UV region. Therefore, in this dissertation, we will address the development of a new hyperspectral imaging prototype in the UV region and some applications combined with chemometrics.

A hyperspectral imager was developed in the UV region to quantify and classify different cotton types with different amounts of honeydew contamination. Figure 1.6 shows a scheme for UV hyperspectral imaging. Figure 1.6a illustrates the principle of continuous line-by-line spectral data collection. The data results in a lateral resolved (x, y) 2D image, as shown in Figure 1.6b, c, whereas each location contains a further spectroscopic dimension (λ), as shown in Figure 1.6d. Thus, a 3D data matrix (hypercube) was recorded [19-21,37-39].

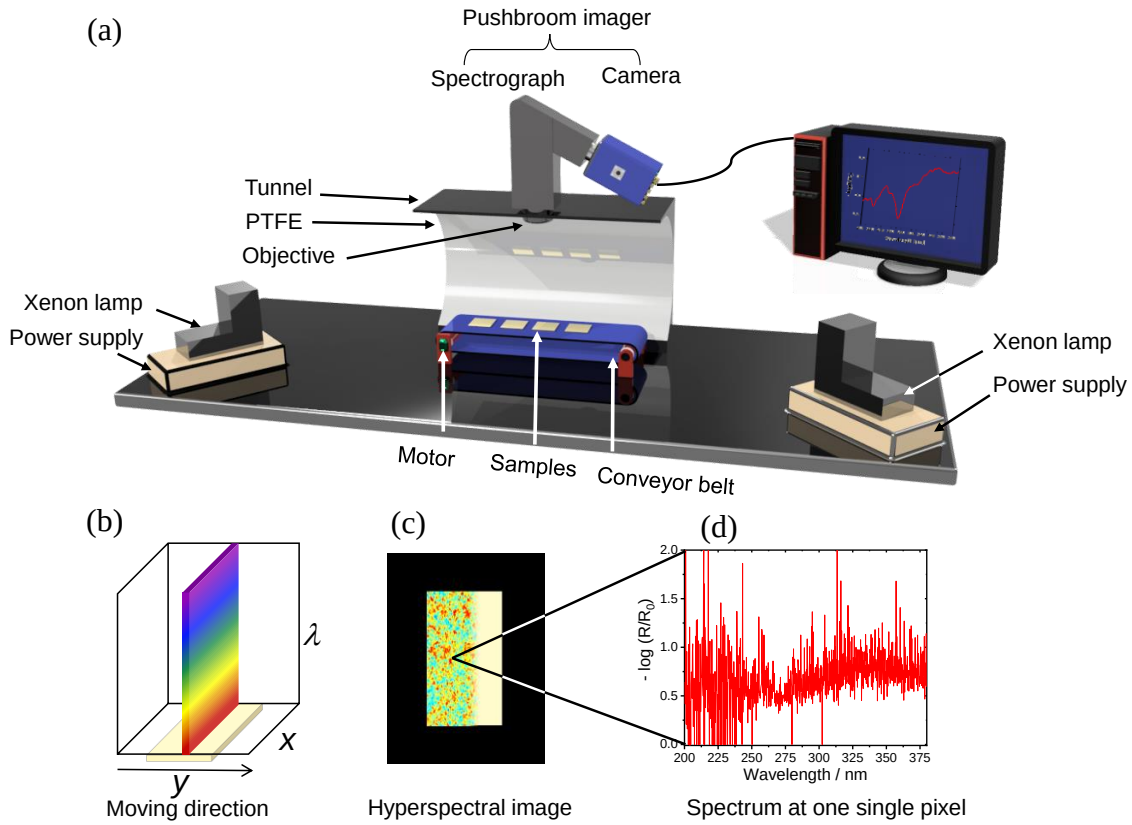


Figure 1.6: (a) Schematic shows the concept of hyperspectral imaging based on the pushbroom (the tunnel in the scheme was cut to show the inside). (b) Pushbroom Imager scanning principle. (c) Hyperspectral image produced immediately during sample scanning. (d) UV spectrum for one single pixel extracted from the image in (c).

Selecting an appropriate light source for UV irradiation is a significant challenge in various fields. Several sources are available in the UV region, but only a limited number are suitable for our purpose. Among others, there are synchrotron radiation [40], light-emitting diodes (LED), laser-induced plasma, deuterium, xenon-arc (XBO), and mercury-arc (HBO) lamps.

Synchrotron radiation is a form of electromagnetic radiation emitted by a particle accelerator when the velocity of the electrons is approaching the speed of light. These electrons are forced to travel in curved paths by a magnetic field, which causes them to emit radiation when they are moved within the synchrotron. This intense radiation covers a wide range of wavelengths, including X-rays, UV-Vis, and IR light. These properties make synchrotron radiation a valuable tool for scientists studying the nature and structure of molecules and materials. Having a synchrotron radiance covering the entire UV range is challenging, but a spectrum within the 225 – 325 nm range was obtained [40]. However, synchrotrons perform off-line measurements due to their transportability limitations, making them unsuitable for industrial quality. Moreover, it is considered one

of the most expensive radiation techniques, with approximately 70 synchrotrons worldwide. Therefore, researchers need a cheaper and more available light source [41-43].

LEDs are diodes with several advantages over traditional incandescent and fluorescent lamps, including their small size and ease of integration into various designs. LEDs have high quantum efficiency, which refers to the ratio of the number of photons emitted by the LED to the number of electrons that pass through it. They have a long lifetime of over 100,000 hours of operation. However, one limitation of UV LED lamps is that they typically emit radiation within a narrow wavelength band, making it challenging to generate a continuous spectrum of radiation with UV LEDs [44,45].

Laser-induced plasma lamps, also known as plasma globes or spheres, are a light source containing a partially ionized gas in a glass sphere. It utilizes laser-driven technology, which is considered a leading choice for the UV-Vis/NIR region. These lamps have high spectral radiance intensity across the aforementioned wavelength range and a longer lifetime than traditional lamps. Due to their energy consumption and heat generation, plasma lamps require water cooling to prevent overheating. Additionally, plasma lamps are relatively expensive compared to other lighting options [46].

Deuterium lamps are considered one of the most stable lamps in the UV region. This is due to the ceramic electrode structure within the lamp, which results in low fluctuation levels in peak intensity. One of the key advantages of deuterium lamps is their ability to produce a continuous spectrum of UV light from approximately 115 – 400 nm. This makes these lamps suitable for experiments that require UV radiation. However, deuterium lamps have a few limitations, such as restricted spectral range and low spectral radiance intensity in the UV region. These lamps are commonly used in various scientific and analytical applications, such as the pharmaceutical industry and for atomic absorption spectroscopy [47,48].

High-pressure mercury vapor lamps, also known as mercury-arc lamps (HBO), are a type of light source that can produce intense, bright illumination. These sources are stable and have a high flux density, making them widely used in fluorescence microscopy. However, they have a relatively low intensity in the UV region compared to xenon-arc (XBO) lamps. XBO lamps have a continuous spectrum and produce high-intensity spectra within 240 nm to 400 nm. Despite the low stability, the broad spectrum range and the long-life feature are added advantages of XBO lamps. On the other hand, XBO lamps produce ozone and heat. In this dissertation, a xenon lamp was used because it is inexpensive, easy to replace, and the aperture of the illumination is adjustable

[49-52]. One of the critical factors to consider when evaluating these sources is the distinct emission spectra they produce, see Figure 1.7. It is important to ensure that the chosen light source emits the desired wavelengths to meet the application's specific requirements. This often involves comparing the emission spectra of the different sources and selecting the one that best matches the desired wavelength range. These lamps are available for use in our experiments, and the prototype tested five different light sources: LED, plasma EQ-77, deuterium, XBO, and HBO. Figure 1.7 shows the different light source spectra used pushbroom hyperspectral imaging setup. The pushbroom imager contains a back-illuminated CCD camera (Apogee Alta F47: Compact, inno-spec GmbH, Nürnberg, Germany) and is connected to a spectrograph (RS 50-1938, inno-spec GmbH, Nürnberg, Germany) [19,23,53-55].

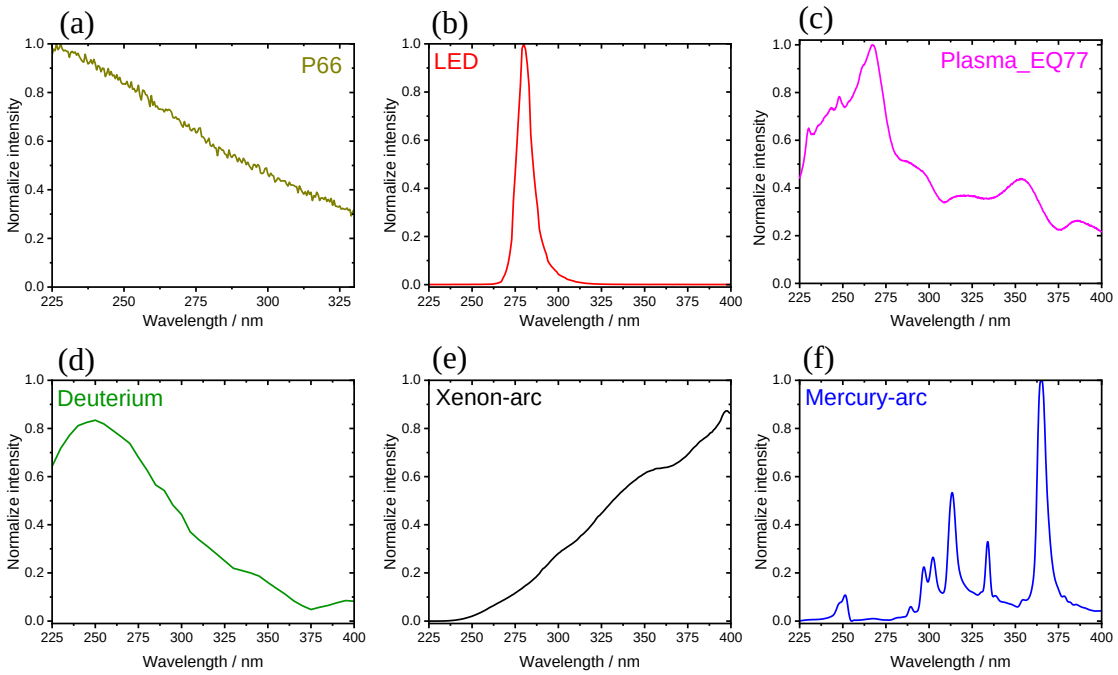


Figure 1.7: Spectral radiance for a different light source in the UV region. (a) Synchrotron radiation of P66 beamline (b) LED radiance (Roithner LaserTechnik GmbH, Wien, Germany) (c) Plasma radiance (EQ-77, Energetiq Technology LDLS™, Wilmington, MA, USA) (d) Deuterium radiance (SL 3, StellarNet Inc, 24 V, 65.04 W, Tampa, Florida, USA) (e) Xenon-arc radiance (XBO, 14 V, 75 W, Osram, München, Germany) (f) Mercury-arc radiance (HBO, 14 V, 75 W, OSRAM, München, Germany). All light sources were tested by UV hyperspectral imaging except Synchrotron radiation of P66 taken from reference [40].

As a result, the synchrotron was excluded from this study because it did not serve our purpose; it was deemed unsuitable for on-line measurements due to lack of portability and expense. LED and mercury lamps were ruled out due to their narrow wavelength bands, see Figure 1.7b and f. Using plasma was limited due to cost considerations and the requirement for cooling. In contrast, Deuterium and XBO-arc lamps were selected, offering a continuous spectrum (Figures 1.7d and e) and low cost.

1.3 Chemometrics

Chemometrics can be derived from “chemo” related to chemistry, and “metric” means measurement. It is a sum of statistical methods that have introduced new algorithms capable of handling the massive amount of chemical data using multivariate data analysis (MVA) [56,57]. MVA can be a powerful tool for data reduction to find a small number of variables capable of explaining all the variations from the data. Typical MVA techniques are principal component analysis (PCA), Discriminant analysis (DA), and partial least squares regression (PLS-R) [58].

1.3.1 Principal component analysis (PCA)

This section is partially based on the work by Rebner K. [59]and has been further modified.

Principal Component Analysis (PCA) is a statistical method used to identify similarities and differences among data sets. It reduces the dimensionality of the variables to their most essential features without losing important information [60-62]. Data reduction and features are extracted from a data matrix with p objects (rows) and t variables (columns). The first summary index, so-called latent variables, the principal components (PCs), are calculated from the original variables via a principal axis transformation. Mathematically, these factors are linear combinations of the original variables. For example, an object in a two-dimensional space described by variables x_1 and x_2 can be transformed into a new vector space represented in PCs. The transformation is given as an orthogonal matrix formed from the eigenvectors of the covariance matrix. The coordinate system is rotated so that it points in the direction of maximum variance in the data, thus describing the information content of the data (see Figure 1.8).

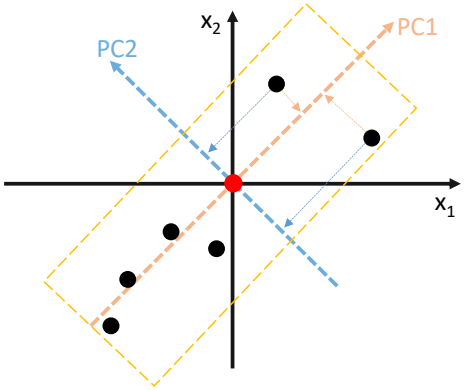


Figure 1.8: Graphical representation of the principal components. The original data in the original data space (x1, x2) are transformed into new principal axes (PC1, PC2).

In data analysis, the first principal component explains the most variance. The second principal component explains the second highest variance in what remained after the effect of the first component was removed [63]. Each new principal component describes the maximum variance that was not captured by the previous components. The number of variables – 1 (maximum number of PCs) is mathematically determined until the data explain a certain percentage of variance. The PCs are orthogonal to each other and thus are independent, which means the data is de-correlated. This also means that the principal axis transformation can look different depending on the problem, and a separate transformation matrix must be calculated for each data set.

The principal components are determined by decomposing the data matrix $\boldsymbol{X}$ into a weight matrix $\boldsymbol{T}$, a transpose factor matrix $\boldsymbol{P}^{\mathrm{T}}$, and an additional residual matrix $\boldsymbol{E}$.

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{E} \tag{1.3}$$

$\boldsymbol{T}\boldsymbol{P}^{\mathrm{T}}$ describes the new data structure, and the residual matrix $\boldsymbol{E}$ consists of noise and unexplained data. Figure 1.9 shows how a matrix with PCA is decomposed. Where the data matrix X is decomposed to PC1 (scores t_1 and loading p_1) and residual $\mathbf{E_1}$, PC2 will be calculated from residual matrix $\mathbf{E_1}$, which contains information not explained by the first PC.

$$X = t_1p_1 + t_2p_2 + \ldots + t_ap_{a+} E \tag{1.4}$$

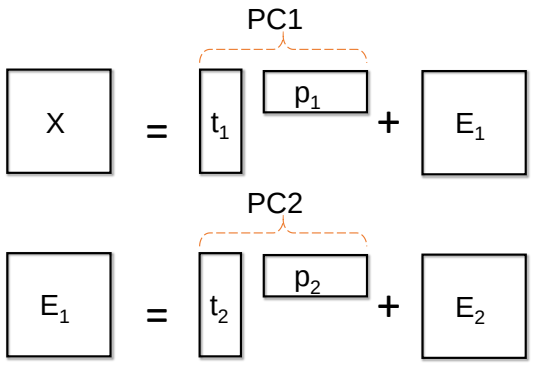


Figure 1.9: Schematic description of a decomposition of a matrix X with PCA using two PCs.

In context of data analysis, after performing PCA to reduce the dimensionality of the data, another statistical method called discriminant analysis (DA) can be used to validate the PCA model. DA is a supervised classification method that assigns an unknown pattern to a group of similar objects. Also, it is a separation technique that optimally divides a training set of objects into two or more groups using a border, which can take the form of linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), or mahalanobis discriminant analysis (MDA) distance-based separators see Figure 1.10. DA is a powerful classification method that simultaneously minimizes the

variance within each group and maximizes the distance between the groups. This gives better separation of the groups compared to PCA, which only looks at the variance of the data. The optimized border may allow for minor group overlaps, and the choice of border shape depends on the nature of the data. For example, quadratic borders help separate groups with differently-oriented main variances, while Mahalanobis distance is suitable for measuring the distance between objects and class centers using ellipses as distance calculators [64-66].

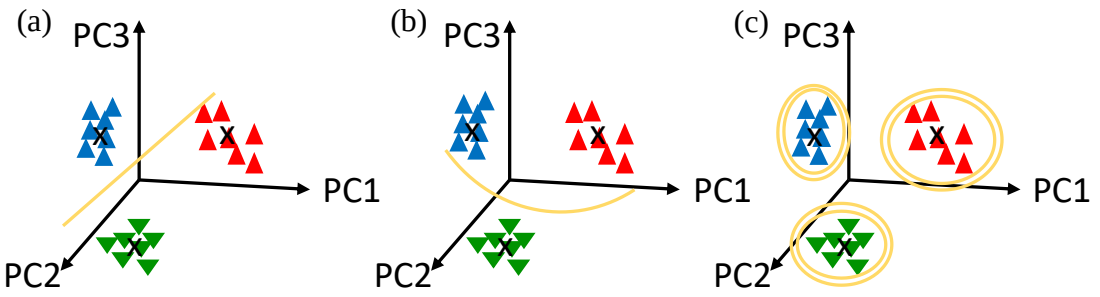


Figure 1.10: Schematic shows different discriminant analysis (DA). DA function creates a border of variable shape that optimally separates a training data set into multiple groups. PCA can be used to reduce the dimension of the training set before creating the discriminant function. The three most common types of DA separators are linear, quadratic, and mahalanobis distance-based separators. Linear separators are represented by straight orange lines (a), quadratic separators by orange curves (b), and mahalanobis distance-based separators by ellipses (c). This figure is taken and modified from reference [66].

To create a discriminant function, the object groups must exhibit significant differences in their variables, and the variables with the maximum variance are identified to achieve the highest possible separation between the groups. The performance of the discriminant functions can be evaluated using a classification matrix or confusion matrix. The confusion matrix, also known as the error matrix, provides a table used to evaluate the performance of a model by comparing the predicted and actual values. It consists of four values: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) see Table 1.1. This table is used to evaluate the performance of a binary classification model by comparing the predicted and actual values. It consists of four values: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). Various performance measurements can be derived from the confusion matrix, such as precision, specificity, and accuracy, which provide insights into the model's effectiveness and potential limitations (see Chapter 3.4, Chapter 4.4) [67].

Table 1.1: Confusion Matrix.

		Predicted value	
		Positive (1)	Negative (0)
Actual value	Positive (1)	TP	FP
	Negative (0)	FN	TN

1.3.2 Partial least squares regression (PLS-R)

This section is partially based on the work by Rebner K. [59] and has been further modified.

PLS is considered one of the most important data analysis tools for regression and classification. This analysis tool decomposes the variance between independent X variables (predictor values) and dependent Y variables (response values) to calculate the variance and correlation between X and Y to estimate PLS-R components [68,69]. The straight-line equation gives the relationship for linear regression:

$$y = b_0 + b_1 x \tag{1.5}$$

Where b_0 is the y-axis intercept and b_1 is the slope of the straight line. This function is fitted to the data x and y to calculate the variance and maximize the correlation between the data representing the first PLS. When using MVA, whose structure may be complex or even error-prone, using a multivariate regression method is preferable to increase predictive accuracy. The PLS method is related to PCA and already uses the structure of the Y data to find PCs. This has the advantage that only a few PCs are often sufficient for the complete data description and can be interpreted easily.

For each object, a target value y_i is measured, which forms the vector $\boldsymbol{y}$ and, in the case of multiple objects, the target value matrix $\boldsymbol{Y}$. The target values can be chemical or physical parameters. As in PCA, the data matrix is decomposed into a $\boldsymbol{T}$ and $\boldsymbol{P}$ matrix. As an intermediate step of the PLS, an additional matrix is needed that calculates weighted loadings $\boldsymbol{W}$ and provides a link to the $\boldsymbol{Y}$ variables. In this case, not the $\boldsymbol{P}$ loadings but the $\boldsymbol{W}$ loadings are orthogonal to the $\boldsymbol{T}$ scores. Therefore, PLS components are used instead of PCs, as in PCA. The target matrix $\boldsymbol{Y}$ is decomposed into a factor matrix $\boldsymbol{Q}$ and a score matrix $\boldsymbol{U}$, as well as an additional residual matrix $\boldsymbol{F}$.

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{E} \tag{1.6}$$

$$\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{Q}^{\mathrm{T}} + \boldsymbol{F} \tag{1.7}$$

Two data sets ***X*** and ***Y*** are now contrasted with their respective score vectors ***T*** and ***U***, respectively. Figure 1.11 shows a schematic description for a PLS-R model. The direction of each PLS latent variable of the ***X*** matrix is now changed such that the covariance between it and the vectors of the ***Y*** matrix is maximized. The exact intermediate mathematical steps are described in detail in the literature [61].

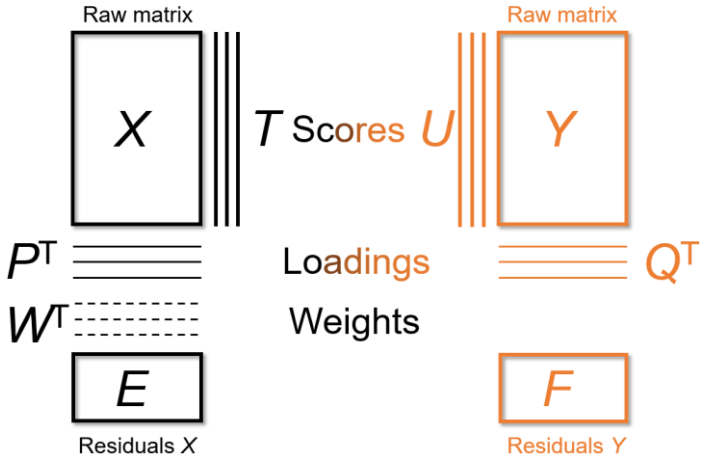


Figure 1.11: Partial least square regression method. The Y variables influence the X variables.

However, unlike PCA, the PLS components do not contain the largest differences in variance between spectra but the most relevant differences concerning the reference data. In the regression approach between the x- and y-variables, the coefficients are calculated as follows:

$$\boldsymbol{B} = \boldsymbol{W}(\boldsymbol{P}^{\mathrm{T}})^{-1}\boldsymbol{Q}^{\mathrm{T}} \tag{1.8}$$

and

$$b_0 = \bar{y} - \bar{\mathbf{x}}^{\mathrm{T}}\boldsymbol{B} \tag{1.9}$$

Where ***B*** is the regression matrix ***W*** represents the loadings matrix for the predictor (***X***) variables, ***Q*** represents the loadings matrix for the response (***Y***) variables, and ***P*** represents the score matrix for the predictor variables. b_0 is the y-axis intercept, and $\bar{y}$ is the mean of the response variable (***Y***). Finally, the target quantity y_i can be given by

$$y_i = b_0 + \mathbf{x}_i^{\mathrm{T}}\boldsymbol{B} \tag{1.10}$$

by using the measured values $\mathbf{x}^{\mathrm{T}}$ for each object. The equation represents the predicted value of the response variable y_i for a given set of predictor variables x_i in a linear regression model. $\mathbf{x}_i^{\mathrm{T}}$ represents the transpose of the predictor variable vector x_i. If the PLS regression is performed with one Y variable, it is called PLS1. In contrast, PLS2 calculates a model for several Y variables

simultaneously. The advantage is that all X and Y data are included in a common model; thus, all correlations are considered. Mathematically, instead of the vectors, the matrices ***Y***, ***F***, and ***Q*** are used and supplemented by the score matrix ***U*** of the ***Y*** values.

In addition to calibration, regression includes steps of validation and prediction. PLS-R is utilized to unite so-called "factors" that describe the relationship between X and Y variables. The optimal number of factors used to describe the model is essential. Too few factors (under fitting) or too many factors (over fitting) lead to additional prediction errors. For an optimal model, this error is minimal, but a difference must be made between a calibration error due to under fitting and an estimation error due to overfitting. The calibration error and the residual variance will decrease with the first PLS components as the relevant information increases to a certain degree. The optimum is exceeded if random changes in the form of noise are modeled with the addition of further components. In this case, the estimation error increases, and the prediction error of unknown data becomes larger than the calculated calibration error once a suitable model can be applied to unknown data [61,62,68].

In chemometrics, a distinction is made between external and internal validation. Internal validation is often used for smaller data sets because the same data set can be used for calibration and validation. Especially for method developments, cross-validation offers a very efficient method when a large number of samples is not yet available. For this purpose, some objects are omitted from the calibration data during the calculation, and a calibration model is created. Afterward, the calibration model is used to predict the omitted objects and to determine the residuals [62]. This process is repeated several times until all objects have been omitted once and predicted with the model. The quality of calibration or validation can be indicated with different values. The most important ones are briefly explained here.

1.3.2.1 Coefficient of determination

The correlation between the reference value y and the predicted value $\hat{y}$ is often given for the regression coefficients, and the coefficient of determination R^2 is calculated. This expresses how much of the variance of the dependent variable y can be explained by the independent variable x and is calculated by the following formula [70]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{1.11}$$

1.3.2.2 Root mean square error (RMSE)

Root mean squared error (RSME) is called the root mean square deviation (RSMD). It is considered a common function and evaluation matrix used in a regression model. RMSE is mainly used to calculate the difference between values predicted by a model and observed values. It is given with the suffix C for "Calibration" in the case of calibration RMSEC, the suffix CV for "Cross-Validation" in case of internal validation RMSECV or the suffix P for "Prediction" in the case of external validation RMSEP. The mean error is calculated by [71,72]:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{1.12}$$

Where n is the number of samples.

1.3.2.3 Standard deviation

The standard deviation is a measure that shows how much data is scattered around the true value. For example, how different the answers of your respondents are. It is summarized into two forms: standard error of calibration (SEC) and standard error of prediction (SEP). The standard error (SE) is the residuals' standard deviation [73]. Mathematically, the SE systematic error between prediction and reference value (BIAS) must be determined in advance by the following formulas [62,74].

$$\mathrm{BIAS} = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n} \tag{1.13}$$

$$\mathrm{SE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i - \mathrm{BIAS})^2}{n-1}} \tag{1.14}$$

1.3.2.4 Prediction

PLS regression is also used for predicting dependent Y values from independent X values. Decompose the following equation gives X value used for building up and predicting Y values [62]:

$$y = b_0 + b_1 x \tag{1.15}$$

Where b_0 is the y-axis intercept and b_1 is the slope of the straight line.

1.4 Model systems

Quality control is an essential feature of products in industrial environments. Several studies verified UV hyperspectral imaging to achieve the quality aspect. This dissertation used various applications, such as pharmaceuticals, metals, thin layers, and natural products.

1.4.1 Pharmaceutical tablets

Chemical drugs have been increasingly used globally for treating and preventing infections and pain [75]. Therefore, countless medicines, tablets, liquids, gels, and powders are made daily. Painkillers were used as a model system; painkillers may also consist of different liquids or powder mixtures [76]. In this case, qualitative and quantitative analyses are needed to identify active compounds, determine the content of active compounds, and measure the significant impurities of the medicine [77]. Thus, these parameters are required rapidly and non-destructively during the manufacturing process.

In the pharmaceutical industry, separation techniques are used for qualitative and quantitative analysis. These analyses are mainly achieved by high-performance liquid chromatography (HPLC) [63]. Snakar et al. [78] described the application of HPLC in pharmaceutical analysis. This technique has many disadvantages, such as being expensive, time-consuming, needing sample preparation, and destructive. Using UV spectroscopy, Saeed et al. [79] estimated the quantity of active ingredients in different tablets of some commercial dosages, such as ibuprofen, aspirin, and paracetamol.

In the last 20 years, the food and drug administration (FDA) in the United States started the process of analytical technology (PAT) to control medicine production processes [19,20]. For this purpose, hyperspectral imaging is an excellent PAT tool for ensuring product quality. One of the aims of this dissertation is to develop a laboratory prototype for hyperspectral imaging in the UV region based on the pushbroom technique in combination with multivariate data analysis.

Active pharmaceutical ingredients (API) 100% and commercial painkiller tablets were used for general testing of UV hyperspectral imaging [19]. The APIs were ibuprofen (IBU), acetylsalicylic acid (ASA) and paracetamol (PAR). Ibuprofen tablets from two companies (ratio pharm IBU_{ratio}, and beta pharm IBU_{beTa}), aspirin (Bayer GmbH ASP_{BAYER}), paracetamol (ratio pharm PAR_{ratio}), and thomapyrin were used (see Figure 1.12). The overview of the APIs and painkiller tablets is given in Table 3.1.

4 g of ASA_{pure} and IBU_{pure} were pressed at 10 tons for 2 min by a hydraulic press into the depicted disc shape. Then, 4 g PAR_{pure} powder were dried in a vacuum oven for 1 h at 120 °C, and pressed at 10 tons for 20 min (see Table 3.1). A mixture of 2 g ASA_{pure} and 2 g PAR_{pure} was prepared by using a speed mixer and pressed at 10 tons for 2 min [19].

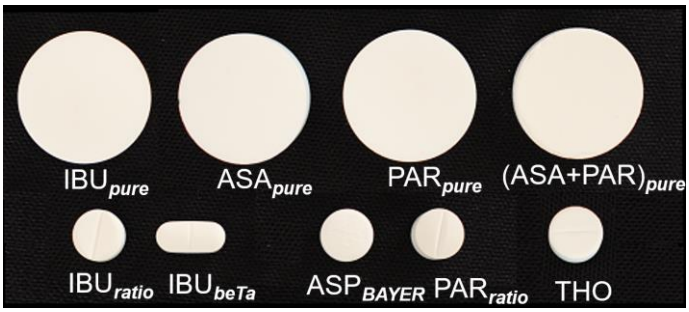


Figure 1.12: Drug samples. Reference API samples and painkiller tablets. This figure is taken from reference [19].

1.4.2 Direct bonded copper

Direct Bonded Copper (DBC) refers to a process in which copper and ceramic material are directly bonded. DBC substrates have been proven to be an excellent solution for electrical insulation and thermal management of high-power semiconductor modules [80]. Therefore, they are considered the most significant conductors compared to other materials, such as aluminum. The advantages of DBC substrates are high electrical and thermal conductivity due to thick copper metallization and thermal expansion close to the silicon copper surface due to the high adhesion strength of copper to ceramic [81]. However, the copper surface interacts with oxygen to produce copper (I) oxide (Cu_2O) and copper (II) oxide (CuO). Therefore, the efficiency of the conductivity becomes poor [81,82]. Optical techniques such as Auger electron and X-ray photoelectron spectroscopy were applied to solve this problem [83,84]. Such systems are time-consuming, destructive, and expensive.

In the past, progress in research led to the development of sensor technology [32,35,77,83,85,86]. Stiedl et al. [26,87] investigated the copper oxide layer thicknesses and copper state on a metallic DBC using UV-Vis spectroscopy and visible hyperspectral imaging. In this thesis, we verified the capability of the new UV hyperspectral imaging prototype to study the changes in the copper state and thickness of the copper oxide on DBC (see Figure 1.13).

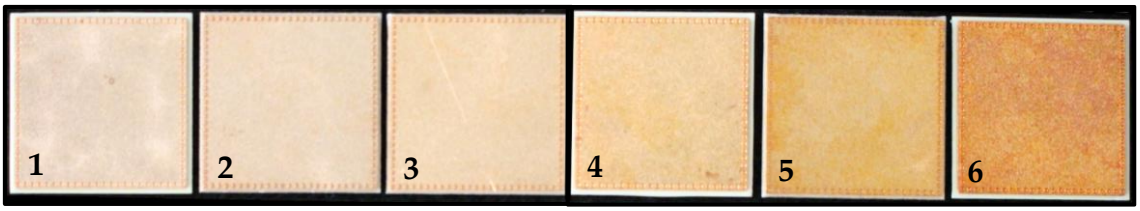


Figure 1.13: Direct bonded copper Curamik®Power substrates. (1), (2), (3), (4), (5) and (6) are different samples with different types of copper state and thickness of the copper oxide on DBC. This figure is taken and modified from reference [53].

1.4.3 Cotton fiber

Cotton is the most important natural raw material used in producing fabrics. It has been used extensively for clothing people worldwide. Cotton lint (see Figure 1.14) is considered an essential product that provides a source of high-quality fiber for the textile industry [88].

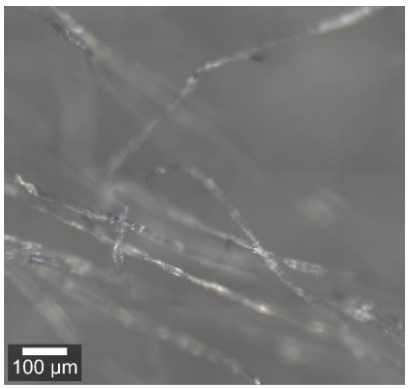


Figure 1.14: Cotton fibers.

Over 34 million hectares of land are used to grow cotton, and around 100 million households worldwide are engaged in cotton production [89]. Cotton is an essential resource in the textile industry, accounting for approximately 30% of all fibers utilized [90]. Almost all parts of the cotton plant have a range of uses. Cottonseeds are an important oil source, which is a byproduct of the plant. Additionally, cottonseeds have a high protein concentration, making them useful in animal feed. The waste material left over from the ginning process, which separates the cotton fibers from the seeds, can be used as fertilizer. The stalk of the cotton plant is also a potential resource, as it contains cellulose, which can be used to make paper and cardboard [91,92].

Cotton is the most widely produced natural textile fiber and a significant global commodity, as it is the most imported and exported raw material [89]. Cotton processing, through spinning or knitting, plays an important role in the economies of many countries. Cotton fabric production begins with the preparation of yarn, which is achieved by removing the seeds. This yarn is then woven

or knit into fabric. The fabric undergoes a series of additional processing stages, including dyeing, designing, and sewing, to produce a final product that is soft, clean, and ready for use [93]. Figure 1.15 presents the steps of the cotton manufacturing process.

Figure 1.15: Manufacture process for raw cotton; this photo is taken from reference [94].

In textile research, cotton plays a dominant role among textiles since cotton is the most important naturally occurring raw material for fabric production. It is important to guarantee the quality of the fibers in textile processing, as fiber properties highly impact the properties of finished yarns and fabrics and how easy it is to manufacture. For example, fiber strength determines yarn strength, and fiber maturity affects the dye uptake of fabrics. Raw cotton with poor fiber quality causes problems in the textile mill [91,92,95].

Cotton fibers are soft, relaxed, and breathable and have high absorbency. They are natural hollow fibers, capable of absorbing liquids such as water up to 24-27 times their weight. The elongation of the fibers begins shortly after anthesis, which is the flowering stage, and continues for three to four weeks, primarily increasing in length during this phase. Two weeks after the anthesis, deposition of cellulose fibrils with varying orientations in the secondary fiber walls commences. The growth process of cotton fibers begins shortly after anthesis, the flowering stage. The fibers increase in length during this stage and elongate for three to four weeks. Two weeks after the anthesis, the deposition of cellulose fibrils with varying orientations in the secondary fiber walls commences. This process reduces the inner space's size, known as the lumen, within the fiber (see Figure 1.16) [96-98].

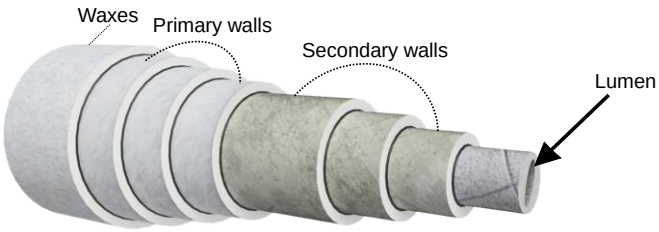


Figure 1.16: Schematic illustration of the structure of cotton fiber, showing its different layers.

Fiber growth and development primarily depend on plant photosynthesis and carbohydrate production, which are impacted by various factors. Temperature and water status are two important factors that can affect fiber growth, length, and quality. Studies have shown that temperature and plant water status impact fiber growth and length. Nighttime temperatures below 22.0 °C inhibit cellulose synthesis and deposition rate in the cotton fiber walls, leading to decreased productivity and inferior fiber quality. Low temperatures during fiber maturation can also result in sucrose accumulation, leading to "sticky cotton," a severe issue that negatively affects cotton quality and value [99].

Under normal conditions, cotton consists of approximately 95% cellulose, while the remaining 5% consists of various substances, including sugar, wax, proteins, organic acids, and pectin [100]. One major quality issue of raw cotton is the impurity content after harvesting. The impurities cause a significant economic loss because low-quality cotton is rejected during quality control. The most relevant impurities in raw cotton arise from insects and are summarized under the umbrella term “Honeydew” [101].

1.4.4 Honeydew

Cotton contaminated by sugar causes significant problems for textile equipment. These sugars are produced by the cotton plant (physiological sugars) or insects (entomological sugars). Entomological sugars, also known as honeydew, are the most common source of sticky cotton (see Figure 1.17 a,b) [102,103].

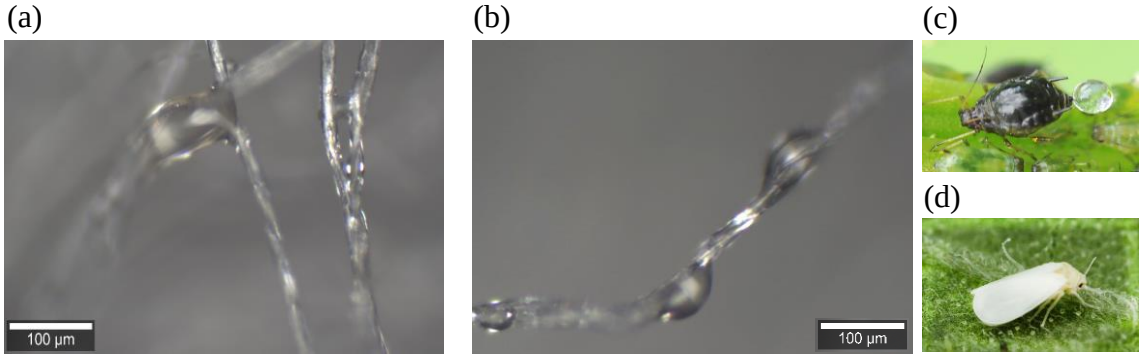


Figure 1.17: (a) and (b) Cotton fiber contaminated by sugar cause of (c) aphid and (d) Whiteflies insects these photos are taken from [104,105].

Honeydew is excreted by aphids and whiteflies (Figure 1.17 c,d) and contains mainly trehalulose, trehalose, melezitose, sucrose, fructose, and glucose (see Figure 1.18). These sugars vary in stickiness, such as sucrose, melezitose, and trehalulose are significantly stickier when deposited on fiber than glucose or fructose. Furthermore, fibers contaminated with trehalulose are more sticky than fibers contaminated with an equivalent amount of melezitose [101,106,107].

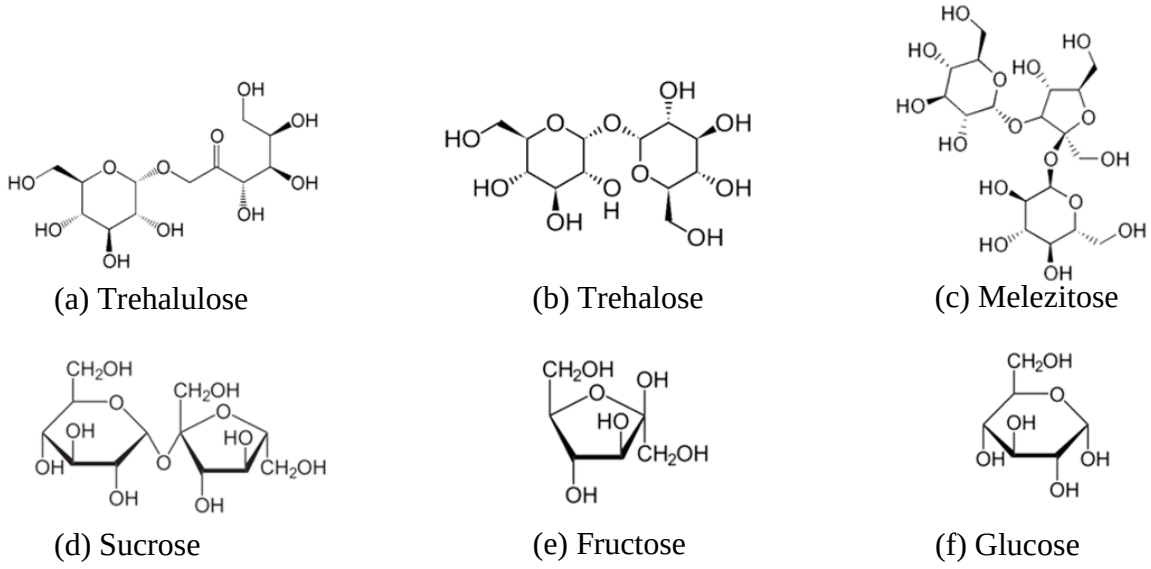


Figure 1.18: Honeydew chemical structure contents (a) Trehalulose (b) Trehalose (c) Melezitose (d) Sucrose (e) Fructose (f) Glucose.

Cotton contaminated with a high amount of honeydew becomes sticky. Transferring sticky cotton to a spinning machine causes severe problems by contaminating all mechanical components. This can cause damage to the machines, and the final yarn is of lower quality [101,106]. In Figure 1.19, the result of sticky cotton on a draw frame roll can be seen. Consequently, yarn from sticky raw cotton reaches a lower quality and achieves a low price on the market [108].

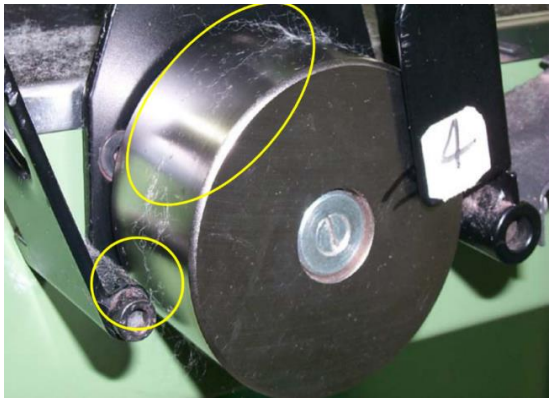


Figure 1.19: sticky cotton residue on a draw frame roll; this photo is taken from reference [102].

The stickiness of raw cotton depends not only on the amount of honeydew on the fiber but also on the ambient conditions such as humidity and machine temperature. The stickiness can be significantly reduced if the moisture content is increased [101]. One obvious idea is washing the raw cotton to reduce the sugar content and thereby reduce the stickiness. It was not a suitable solution because it consumes much water and requires an additional drying step. It makes the whole process even more difficult [109-111]. The best and most economical solution seems to be blending cotton with different honeydew contents to obtain an optimized blend [112]. Since the sugar composition determines cotton's stickiness, identifying the sugars is mandatory to optimize the blending process. If the sugar composition is known in detail, the blending process can be adapted to the climate conditions (temperature, humidity) where the following production steps will occur [101]. To reach this goal, fast on-line analytics is required to monitor the cotton quality with high resolution in honeydew content and time.

In recent years, the increase in quality and processing requirements has led to the introduction of modern techniques for processing and quality control [15,113-115]. Currently, the classification of cotton is done by the United States Department of Agriculture (USDA) system [116]. It classifies the quality by applying the High Volume Instrument (HVI). HVI is the standard instrument for measuring cotton quality in the USA and has been adopted by other countries as it provides a range of standards for measuring cotton quality, including micronaire, strength, length, color, foreign cotton, and material [117-120]. HVI cannot measure the quality of single fibers and short fibers. Therefore, another system, the Advanced Fiber Information System (AFIS), can be used. Although both HVI and AFIS measure a set of quality characteristics, they are not able to quantify and distinguish between cotton species, which is expensive and time-consuming [121]. To solve this problem, spectroscopic methods have been developed and applied to identify and classify different cotton fiber varieties in Vis/NIR region.

2 Objective

The present work aims to classify different types of cotton and determine the honeydew contents on real cotton samples by developing hyperspectral imaging in the UV region. Traditional methods, such as fiber quality index (FQI), the spinning consistency index (SCI) and the premium-discount index (PDI) used for this determination are not useful for large amounts of cotton bales. In contrast, optical spectroscopy, UV-Vis/NIR spectroscopy, and hyperspectral imaging, especially UV hyperspectral imaging, are able to distinguish between different cotton types and honeydew contaminated on cotton. The data evaluation, where a correlation between the spectroscopic and sensory data has to be established, is a special challenge. However, this can be overcome by multivariate data analysis, for example, to evaluate the spectra in connection with the results of the amount of honeydew contaminated on cotton sample measurement. Subsequently, the obtained information is linked by PCA and PLS-R. This is necessary because the information is usually superimposed and thus cannot be derived directly from the respective spectra.

The analysis method described above is already widely used for scientific purposes. However, developing a method suitable for in-line analysis and monitoring industrial manufacturing processes in the UV region has not been possible. The integration of the analysis method into the manufacturing processes offers an enormous optimization potential of the existing value chain since the time-consuming sample taking and preparation for an off-line measurement is no longer necessary due to the in-line measurement. Therefore, developing UV hyperspectral imaging is mandatory. The investigation of the effectiveness of such a prototype is tested by using standard samples such as active pharmaceutical ingredients (APIs) as well as direct bonded copper (DBC). To achieve this purpose, a calibration model is first developed using PLS-R. This model is then used to build a prediction model for estimating the APIs content and oxide layer thickness on DBC samples from the UV hyperspectral image. These experiments have demonstrated that UV hyperspectral imaging can effectively identify and categorize pharmaceutical and DBC samples. Based on this information, it is now possible to focus on the actual question. The question if different concentrations of honeydew contaminated on cotton can be determined by UV hyperspectral imaging, with the necessary accuracy and safety, on-line or in-line during manufacturing processes, should be answered. Therefore, PCA and PLS-R models are developed for honeydew content. Finally, hyperspectral imaging is compared to conventional UV-Vis spectroscopy in terms of prediction accuracy.

3 Paper I: Characterization of Pharmaceutical Tablets Using UV Hyperspectral Imaging as a Rapid In-Line Analysis Tool

Mohammad Al Ktash[1,2], Mona Stefanakis[1,2], Barbara Boldrini[1], Edwin Ostertag[1] and Marc Brecht[1,2,]*

[1]Process Analysis and Technology PA & T, Reutlingen University, Alteburgstraße 150, 72762 Reutlingen, Germany
[2]Institute of Physical and Theoretical Chemistry, Eberhard Karls University Tübingen, Auf der Morgenstelle 182, 72076 Tübingen, Germany
[*]Correspondence: Marc.Brecht@Reutlingen-University.de

This is originally published in sensors (https://doi.org/10.3390/s21134436) as

"Al Ktash, M.; Stefanakis, M.; Boldrini, B.; Ostertag, E.; Brecht, M. Characterization of Pharmaceutical Tablets Using UV Hyperspectral Imaging as a Rapid In-Line Analysis Tool. Sensors 2021, 21, 4436. https://doi.org/10.3390/s21134436"

3.1 Abstract

A laboratory prototype for hyperspectral imaging in ultra-violet (UV) region from 225 to 400 nm was developed and used to rapidly characterize active pharmaceutical ingredients (API) in tablets. The APIs are ibuprofen (IBU), acetylsalicylic acid (ASA) and paracetamol (PAR). Two sample sets were used for a comparison purpose. Sample set one comprises tablets of 100% API and sample set two consists of commercially available painkiller tablets. Reference measurements were performed on the pure APIs in liquid solutions (transmission) and in solid phase (reflection) using a commercial UV spectrometer. The spectroscopic part of the prototype is based on a pushbroom imager that contains a spectrograph and charge-coupled device (CCD) camera. The tablets were scanned on a conveyor belt that is positioned inside a tunnel made of polytetrafluoroethylene (PTFE) in order to increase the homogeneity of illumination at the sample position. Principal component analysis (PCA) was used to differentiate the hyperspectral data of the drug samples. The first two PCs are sufficient to completely separate all samples. The rugged design of the prototype opens new possibilities for further development of this technique towards real large-scale application.

3.2 Introduction

A large number of remote sensing applications have been developed over the last decade [122]. This also led to establish non-destructive imaging systems that are able to quickly identify quality problems within the scanned area [22,23]. Spectral imaging involves both spectral and spatial information of any particular sample or region within an area of interest, thus each pixel represents spectral and spatial information. Imaging systems can be realized in the modes of hyperspectral and multispectral imaging. The difference between these modes is the number and width of the recorded spectral bands. In multispectral imaging 3-10 bands are used [123]. In hyperspectral imaging hundreds or thousands of correspondingly more narrow bands are employed [18,29,86,124]. Therefore, hyperspectral imaging is also known as imaging spectroscopy, a technique that combines conventional imaging with spectroscopy [86]. Hyperspectral imaging setups produce a 3D data matrix often referred to as hypercube. Two of the dimensions are reserved for the spatial information (x, y coordinate) while the third dimension represents the spectroscopic information (λ coordinate) [21,125,126].

Hyperspectral imaging is not restricted to the visible range, nowadays high performance systems are also available for the near infrared range (NIR) [35,86]. Hyperspectral imaging is a

rapid and non-destructive method which analyzes samples without changing their physical shape. This robust technique in combination with real-time chemometric analysis can be easily integrated into an industrial production environment. This enabled chemical sensing systems for very different applications in the fields of food quality monitoring, textile classification, agriculture, detection target of military, astronomy, life science, medicine and pharmaceutical drugs [22,23,25,26,127,128]. Traditional methods such as UV-Vis spectroscopy, high performance liquid chromatography (HPLC) or mass spectrometry (MS) are, in contrast, time consuming, expensive and require sample preparation and destruction [22,86]. Very recently, Tschannerl et al. reported an interesting application of hyperspectral imaging in UV range. They were able to precisely discriminate between phenolic flavor concentrations in melted barley by using hyperspectral imaging in UV and NIR regions [86].

In 2004, the food and drug administration (FDA) in the US started the Process Analytical Technology (PAT) initiative to control manufacturing processes [22]. Hyperspectral imaging is an attractive PAT tool for the quality assurance of final products. Hyperspectral imaging, as expected, will be increasingly used as a PAT tool in the industry; it has been already applied in the industrial manufacturing of pharmaceutical drugs and quality control of pharmaceutical products [17,129]. Most drugs appear colorless to the eye, meaning that they do not absorb light in the visible region but they may absorb in the UV region according to the chemical structure [130]. Such drugs as ibuprofen (IBU), acetylsalicylic acid (ASA) and paracetamol (PAR) show certain absorbance in UV region [131]. Up to now, a variety of drug studies in the UV-Vis region have been performed; Saeed et al. investigated the active pharmaceutical ingredients paracetamol, aspirin, ibuprofen, codeine and caffeine in different formulations by UV-Vis spectroscopy [132]. Rote et al. developed a method to simultaneously quantify paracetamol and nabumetone by area under curve in bulk and tablet dosage form [133].

Hyperspectral imaging setups acquire thousands of spectra in short time resulting in a massive amount of data. Therefore, techniques for data evaluation like the principle component analysis (PCA) methods are required. PCA is one of the most common statistical methods. This technique is used for data evaluation/reduction but simultaneously minimizing information loss in spectroscopy [38,134]. In addition, it is capable of visualizing common features in the data set to detect possible groups and their heterogeneity within samples [135]. PCA combined with hyperspectral imaging data can highlight the relative distributions of different components in mixtures and reveal the spectral features in the spectroscopic data [67,127].

The aim of this study is to develop a hyperspectral imaging system in the UV wavelength range for the in-line characterization of pharmaceutical tablets. The results show that hyperspectral imaging in the UV range is a suitable technique for in-line measurements with the aim of a real-time classification at short time and low costs.

3.3 Materials and Methods

3.3.1 Samples

Two groups of samples were analyzed. Figure 3.1 shows photographs of all tablets used. Further information is listed in Table 3.1. In the following, these samples are referred to as IBU_{pure}, ASA_{pure}, PAR_{pure}, IBU_{ratio}, IBU_{beTa}, ASP_{BAYER}, PAR_{ratio} and THO. For hyperspectral imaging measurements, the coating of the commercially available painkiller tablets was removed by sandpaper manually (grain size 320, Emil Lux GmbH & Co. KG, Wermelskirchen, Germany). For each removal step a new stripe was brushed over it twice. The painkiller tablets were measured at different depths. A layer of approximately 500 µm ± 50 µm thickness was removed from the samples after each measurement. Three samples of each type were collected (painkiller samples) and created (pure API samples) for the study.

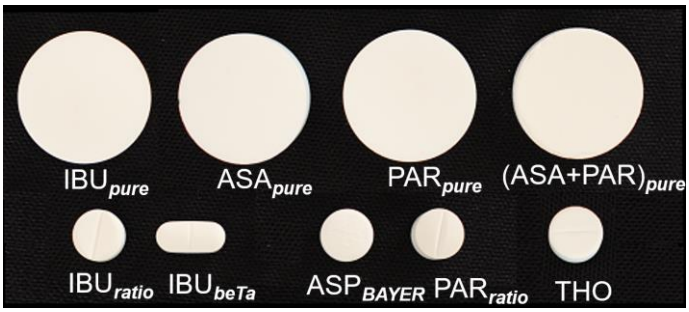


Figure 3.1: Drug samples. Reference API samples and painkiller table.

Table 3.1: Types of drug samples

Samples	Descriptions	Abbreviation	Manufacturer	CAS Number
Ibuprofen	Ibuprofen, >98%, API	IBU_{pure}	Caesar & Loretz GmbH, Hilden, Germany	15687-27-1
Acetylsalicylic acid	Acetylsalicylic acid, 99%, API	ASA_{pure}	Acros organics, New Jersey, US	50-78-2
Paracetamol	Paracetamol, 99%, API	PAR_{pure}	Hebei Jiheng (Group) Pharmaceutical Co., Ltd.	103-90-2
Ibuprofen	Ibuprofen (400 mg)	IBU_{ratio}	Ratiopharm GmbH, Ulm, Germany	-
Ibuprofen	Ibuprofen (400 mg)	IBU_{beTa}	Betapharm, Arzneimittel GmbH, Augsburg, Germany	-
Aspirin	Acetylsalicylic acid (500 mg)	ASP_{BAYER}	Bayer Vital GmbH, Leverkusen, Germany	-
Paracetamol	Paracetamol (500 mg)	PAR_{ratio}	Ratiopharm GmbH, Ulm, Germany	-
Thomapyrin	Thomapyrin (250 mg acetylsalicylic acid/ paracetamol, 50 mg coffin)	THO	Sanofi-Aventis GmbH, Frankfurt, Germany	-

In total, 4 g of ASA_{pure} and IBU_{pure} were pressed at 10 tons for 2 min by a hydraulic press (PerkinElmer, Inc., Waltham, MA, USA) into the depicted disc shape. Then, 4 g PAR_{pure} powder were dried in a vacuum oven (VACUTHERM, Thermo Scientific, Waltham, MA, USA) for 1 h at 120 °C, and pressed at 10 tons for 20 min (see Table 3.1). A mixture of 2 g ASA_{pure} and 2 g PAR_{pure} was prepared by using a SpeedMixer™ (DAC 150.1 CM41, Hauschild GmbH & Co KG, Hamm, Germany), and pressed at 10 tons for 2 min.

3.3.2 API's in Solution

A solution of ASA_{pure} (100 µg mL^{-1}) was prepared by dissolving 50 mg ASA_{pure} in 500 mL of 0.1 M HCl:methanol (1:1) in 500 mL volumetric flask with strong shaking.

For PAR_{pure} and IBU_{pure} solutions, 10 mg of each API were dissolved in 15 mL methanol by shaking. Then, 85 mL water was added to adjust the volume up to 100 mL (resulting to 100 ppm). From that, 5 mL were taken, and volume was adjusted up to 50 mL with diluent [132].

3.3.3 UV Spectroscopy

Total (specular and diffuse) reflectance spectra of all samples (pure API and painkiller tablets) were recorded in the range of 200–380 nm using a commercial UV spectrometer (Lambda 1050+, PerkinElmer, Inc., Waltham, MA, USA). Both sides of the pure API samples were measured. The UV-Vis/NIR spectrometer was equipped with a 150 mm Spectralon® integrating sphere to acquire data in reflection mode with an R6872-Photomultiplier (PMT). A deuterium lamp was used as light source in the spectrometer. The samples were placed at the reflectance port of the integrating sphere with a diffused scattering Spectralon® disk placed behind the samples. The port measuring area is approximately 4.9 cm².

Absorbance spectra were measured using the aforementioned spectrometer in the range of 200–320 nm connected to the transmittance accessory. The liquid samples were measured at 2 nm spectral resolution. A 1 mm quartz SUPRASIL® cuvette (106-1-K-40, Hellma, Müllheim, Germany) was used for measuring the API's in solution.

Fluorescence excitation spectra were recorded by using a commercial setup (Fluorolog–3, HORIBA, Kyoto, Japan). The system includes a double grating monochromator in the excitation (λ_{Ex} = 270 nm) and emission (λ_{Em} = 280 nm–380 nm) paths in an "L" configuration. A 10 mm quartz SUPRASIL® cuvette (111-10-K-40, Hellma, Müllheim, Germany) was used for measuring the samples.

3.3.4 UV Hyperspectral Imaging

Figure 3.2a shows a scheme of the hyperspectral imaging setup. The setup is based on a spectrograph (RS 50-1938, inno-spec GmbH, Nürnberg, Germany) connected to a CCD camera (Apogee Alta F47: Compact, inno-spec GmbH, Nürnberg, Germany) with 300 ms integration time. The samples were placed on a conveyor belt moving with speed 0.3 cm/s, which was positioned completely in a tunnel made of PTFE. The purpose of the tunnel design is to have an easily accessible system, which also ensures diffuse illumination of the samples and maintains a reasonable illumination strength and homogeneity. This minimizes an influence of the sample shape and roughness on the spectra. The illumination is provided by a Xenon lamp (XBO, 14 V, 75 W, OSRAM, München, Germany). Figure 3.2b–d illustrates the principle and workflow of the data acquisition. The continuous line by line collection of spectral information enables a lateral (x, y) 2D image as shown in Figure 3.2c, whereas each pixel contains

a further spectroscopic dimension (λ) as shown in Figure 3.2d. Thus, a 3D data matrix (hypercube) is recorded.

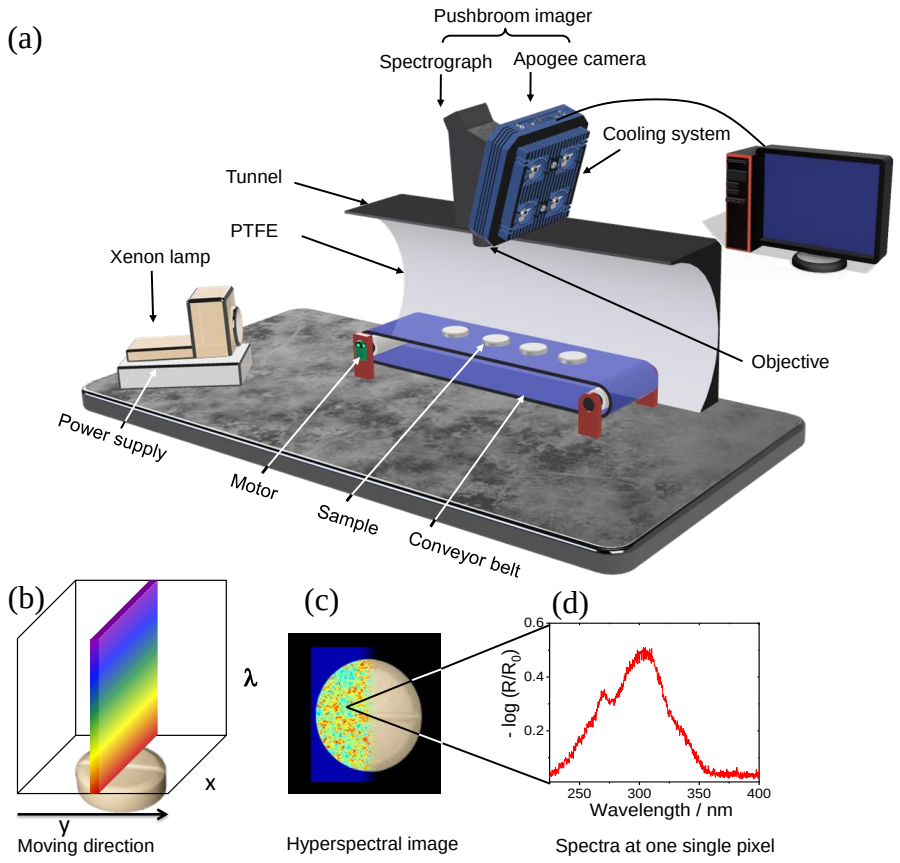


Figure 3.2: (a) Setup of a hyperspectral imaging system based on the pushbroom concept (the tunnel in the scheme was cut to show the inside). (b) Pushbroom Imager scanning principle. (c) Hyperspectral image generated immediately from the scanning of a sample. (d) UV spectrum for one single pixel extracted from the image given in (c).

3.3.5 Data Collection and Preprocessing

Figure 3.3 shows the original images of the drug samples before and after background subtraction. The UV hyperspectral images are captured by moving the drug samples at constant speed. For the collection of UV hyperspectral imaging data set one sample of each type was chosen randomly.

A distinction between the respective spectral characteristics was made first to differentiate signal and background. For this purpose, the regions assigned to the drug samples were manually selected to eliminate the signals from background. The remaining hypercube was used as input for the subsequent PCA classification.

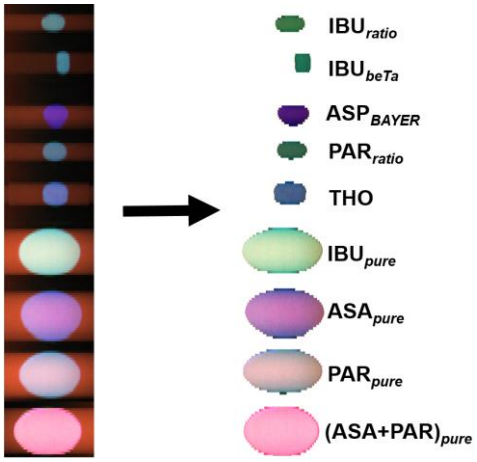


Figure 3.3: Hyperspectral raw image of nine drug samples on the left. Images after subtraction of the background on the right.

3.3.6 Data Handling and Software

The UV spectra were recorded with the Lambda 1050 UV WinLab software from PerkinElmer. The UV hyperspectral imaging data were analyzed by the SI-Cap-GB version V3.3.x.0 software (inno-spec GmbH, Nürnberg, Germany). Hyperspectral data matrices were analyzed by Prediktera Evince version 2.7.11. PLS_Toolbox (PLS Toolbox 8.5.1, Eigenvector Research, Inc., WA, USA) and MATLAB (MATLAB 9.2.0, Mathworks, MA, USA) were used for the data processing and analysis. An initial baseline correction was followed by a Savitzky-Golay 1st derivative (15 points, 2nd polynomial order). PCA models were calculated with cross validation (venetian blinds, 10 splits, 1 sample per split) and mean centering. A PCA combined with a quadratic discriminant analysis (QDA, 2 PCs) was calculated by using the software Unscrambler X 10.5 (Camo Analytics AS, Oslo, Norway) including the same spectral preprocessing.

Lighting conditions may vary between the samples and even within the samples across the scan line. A regular way to reduce this effect is to convert measured raw spectra to reflectance spectra by the following formula [35,86]:

$$\text{Reflectance} = -\log R/R_0 = \frac{I_{sample} - I_{dark}}{I_{reference} - I_{dark}} \quad (3.1)$$

where R and R_0 represent the reflected intensity by the sample and a specific reference material with high reflectance capability. I_{sample} is the intensity of the original image data, I_{dark} is the intensity of the dark current image data and $I_{reference}$ is the intensity of the white reflectance image [23]. For a better comparison of the reflectance spectra to the extinction spectra in solution (absorbance) the negative decadic logarithm is calculated as $-\log R/R_0$.

3.4 Results and Discussion

3.4.1 UV Spectroscopy

There are numerous references for the APIs in solution [101,131,132,136] in the UV range, but for solid API drug samples suitable references were not found. For this reason, first the liquid solutions of the APIs were measured and then compared to the results found in the literature. In a second step, samples in the solid phase, i.e., the pure API reference samples and the commercial painkillers, were investigated.

3.4.1.1 APIs in Liquid Phase, Transmission Spectroscopy

The absorbance of IBU_{pure}, ASA_{pure}, PAR_{pure} as well as a mixture of ASA_{pure} with PAR_{pure} in liquid solution were analyzed in the UV range. Figure 3.4 shows their absorption spectra in the UV region (200–320 nm). The smaller features of IBU_{pure} and ASA_{pure} in the range of 240–300 nm are shown in the inset in Figure 3.4. All samples show a strongly increasing absorbance below 310 nm. IBU_{pure} presents one prominent maximum at 223 nm and three weaker maxima located at approximately 258 (sh), 265 and 273 nm. ASA_{pure} exhibits a broad maximum at approximately 228 nm and a further, more pronounced but less intense maximum at around 277 nm. PAR_{pure} shows a distinct band with a maximum at 244 nm and a weak shoulder at 284 nm. The mixture of ASA_{pure} and PAR_{pure} presents a band maximum at 240 nm and a shoulder at 282 nm. These findings are listed in Table 3.2 [131,132,136]. The determined band positions are consistent with those reported by Saeed et al. (2016) and Lawson et al. (2017) [132,136].

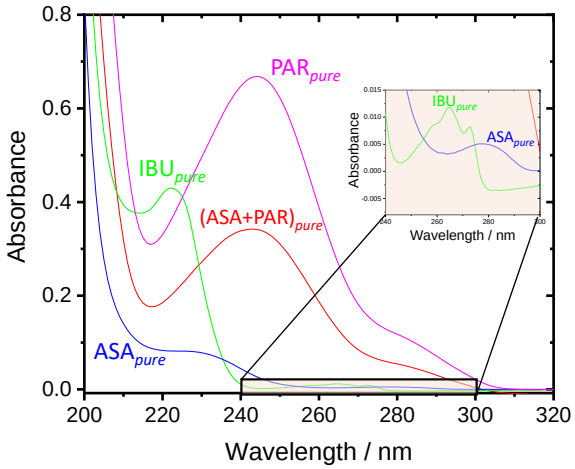


Figure 3.4: UV absorbance spectra of APIs ibuprofen (IBU), acetylsalicylic acid (ASA), paracetamol (PAR) and a mixture of acetylsalicylic acid and paracetamol (ASA+PAR) in liquid phase.

Table 3.2: UV band maxima positions of liquid and solid phase samples [131,132,137].

Drug Type	API Liquid Phase	API Solid Phase	Painkiller Tablets Solid Phase
IBU	223 nm 258 nm (sh) 265 nm 273 nm	240 nm 275 nm	238 nm (IBUratio, IBUbeTa) 275 nm (IBUratio, IBUbeTa)
ASA/ASP	228 nm 277 nm	230 nm 277–310 nm 328 nm (sh)	228 nm (ASPBAYER) 280 nm (ASPBAYER) 294 nm (ASPBAYER) 329 nm (ASPBAYER)
PAR	244 nm 284 nm	232 nm 305 nm	233 nm (PARratio) 300 nm (PARratio)
THO	-	-	238 nm 331 nm
ASA+PAR (mixture)	240 nm 282 nm	235 nm 277–332 nm	-

3.4.1.2 API and Painkiller Tablets, Total Hemispherical Reflectance Spectroscopy

Two sample sets of tablets were used to study the total hemispherical reflectance in the solid phase (see Figure 3.1 and Table 3.1). The first set consisted of pure APIs: IBU_{pure}, ASA_{pure} and PAR_{pure} and a mixture of ARA_{pure} and PAR_{pure}. The second set consisted of commercial painkiller tablets. Three samples from each API were prepared and analyzed. Figure 3.5 shows the preprocessed reflectance spectra of solid samples in the UV region (200–380 nm). Spectra were recorded from each side of the samples (Figure 3.5a).

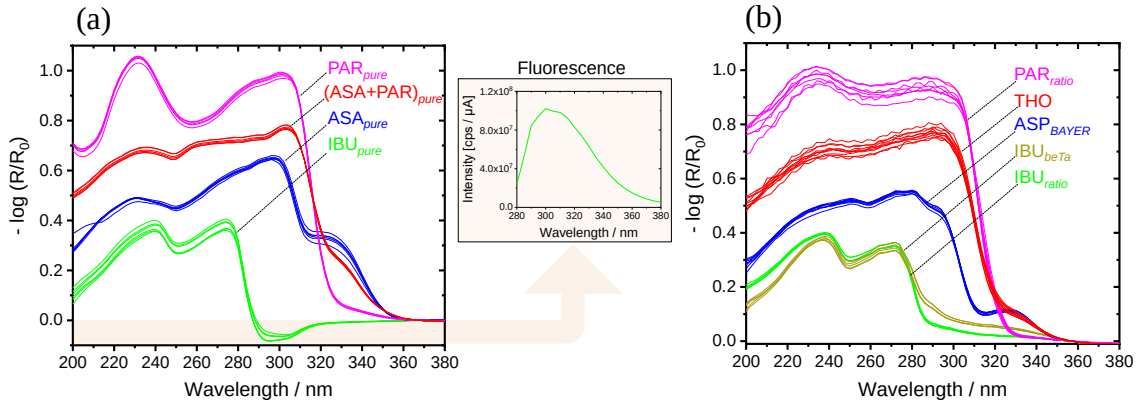


Figure 3.5: UV total hemispherical reflectance spectra of drug samples in the solid phase in the wavelength range 200–380 nm. (a) API drugs IBU_{pure}, ASA_{pure}, PAR_{pure} and a mixture of ASA_{pure} with PAR_{pure}. Upper right: Fluorescence emission of IBU sample with excitation at 270 nm. (b) Painkiller tablets IBU_{ratio}, IBU_{beTa}, ASP_{ratio}, PAR_{ratio} and THO.

The most striking feature is the negative reflectance of IBU_{pure} in the wavelength range of 288–340 nm, which is due to fluorescence emission (inset in Figure 3a). All spectra show several contributions, which are listed in Table 4.2.

The painkiller tablets were measured at different depth levels, i.e., one layer of approximately 500 µm was removed from the samples after each measurement; the resulting spectra are shown in Figure 3b. The similarity of the spectra at all depth levels inside the tablets indicates an almost regular distribution of ingredients. Although IBU_{ratio} and IBU_{beta} were manufactured from different companies, they show similar spectral characteristics. The most prominent contributions are also listed in Table 3.2.

The comparison of the spectra from the APIs and commercial painkiller tablets indicates that the overall spectral characterizations are comparable. Nevertheless, several deviations are observed. The spectral features are more pronounced in the API samples, also the negative absorbance observed in the IBU sample is absent in the commercial painkillers. The reason for these differences is mainly that the commercial tablets do not have 100% API content. For example, the IBU_{ratio} tablets contains additionally pregelatinized corn starch, hypromellose, croscarmellose sodium, stearic acid, highly dispersed silicon dioxide, macrogol 8000, titanium dioxide. Some of these substances show some absorption in selected spectral range i.e., titanium dioxide shows a pronounce absorbance [138]. Since the exact percentage of the composition is not known, a final statement on the influence of these substances on the spectra cannot be made.

3.4.2 UV Hyperspectral Imaging

3.4.2.1 API Tablets, Hyperspectral Imaging

Figure 3.6 shows the results of UV hyperspectral imaging in the range from 225 to 400 nm. Figure 4.6a shows the raw image before (left) and after subtraction of the background (right). Figure 3.6b shows a spectrum of an arbitrary but representative pixel for each API sample. The most dominant contribution for IBU_{pure} is observed around 275 nm and for PAR_{pure} at around 305 nm. For ASA_{pure}, two strong contributions at 300 and 330 nm are observed. The mixture of $(ASA+PAR)_{pure}$ shows—as expected—a combination of the spectral properties of ASA_{pure} and PAR_{pure}. In the range 255–270 nm, all API preparations show a small peak in their reflectance at around 265 nm. Towards lower wavelengths, the spectra show no additional features.

In the next step, a PCA model with cross validation (venetian blinds, 10 splits, 1 sample per split) was calculated for the spectra of all preparations. The first two PCs ex-plain 98.9% of the total variance. Figure 3.6c shows the scores plot of the PC1 and PC2. The scores plot shows that PC1 and PC2 are sufficient to separate all samples clearly from one another. PC1 yields a clear separation of IBU_{pure} from the other APIs, whereas the remaining APIs are separated with PC2. The mixture $(ASA+PAR)_{pure}$ is found almost in the middle between ASA_{pure} and PAR_{pure}.

The loadings plot for PC1 and PC2 is given in Figure 3.6d. The loading of PC1 is dominated by an overall positive contribution in the range between 280 and 350 nm, PC2 shows one more narrow negative contribution at 311 nm and one positive at 330 nm.

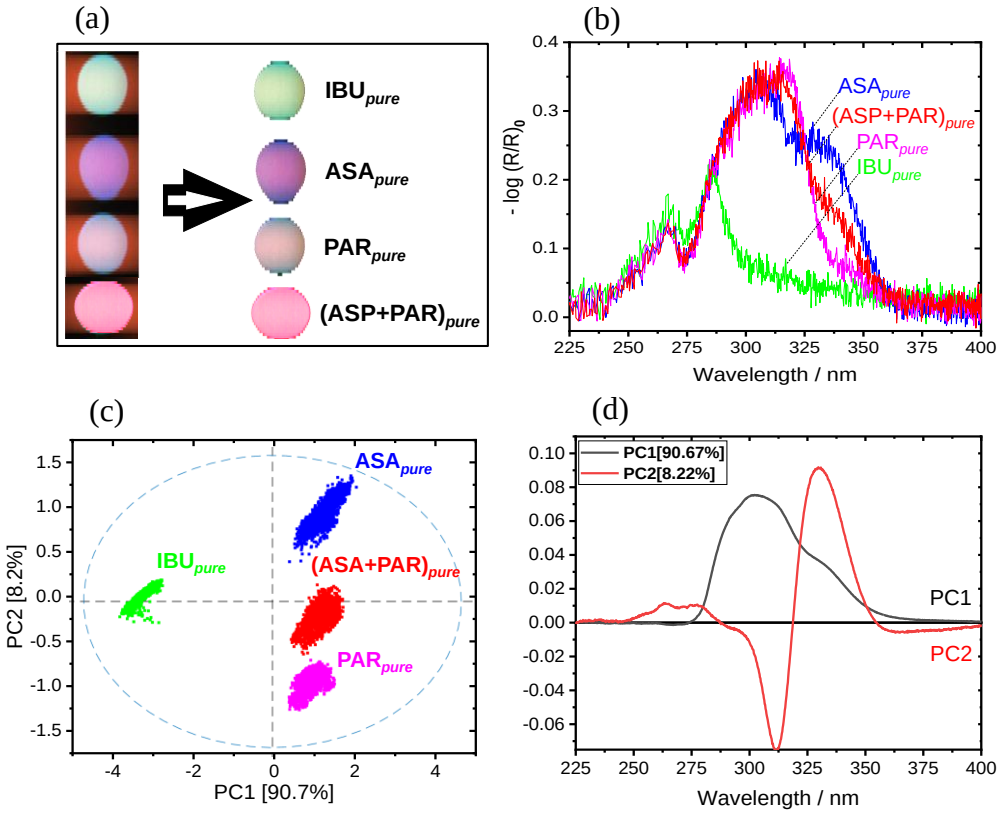


Figure 3.6: (a) Raw hyperspectral image for all API drug samples before and after subtracting the background. (b) Spectrum recorded for a single pixel of each of pure API samples in the UV range 225–400 nm. (c,d) Scores and corresponding loadings plot.

The comparison between the shape of spectra shown in Figure 3.5a or Figure 3.6b shows similarities as well as some clear deviations. The shape of the spectra of all APIs is quite well reproduced in the range above 275 nm. Most striking in this range is an intensity deviation of the different spectral bands, i.e., for the ASA$_{pure}$ the shoulder at 330 nm is much more pronounced in the hyperspectral imaging spectra. The same is valid also for the PAR$_{pure}$ sample. In the range below 275 nm clear deviations are observed. The shape of the spectra shown in Figure 3.6b is characterized by a continuously decreasing intensity, whereas the spectra of the APIs in Figure 3.5a show clear variations in their shape in this range (see also below).

Figure 3.5 presents UV spectra with a good signal-to-noise-ratio recorded with a research grade UV desktop spectrometer. Here, the measured area for one spectrum consisted of 12 × 5 mm. Figures 3.6 and 4.7 show UV spectra with a less good signal-to-noise-ratio recorded with the UV

hyperspectral imager. These spectra result from one single pixel of the detector, representing a much smaller area of the tablet which considered of 13 × 13 µm. A further reason for the low signal-to-noise-ratio is weak irradiation intensity in the hyperspectral imaging setup by an XBO lamp.

3.4.2.2 Commercial Painkiller Tablets, Hyperspectral Imaging

Figure 3.7 shows results from the hyperspectral imaging of the second sample set, consisting of commercial painkiller tablets, in the range from 225 to 400 nm. Figure 3.7a shows the raw image before (left) and after (right) subtraction of the background. Figure 3.7b shows a spectrum of an arbitrary but representative pixel for each painkiller samples. For IBU_{ratio} and IBU_{beTa}, the most dominant contribution is observed around 270 nm, and a further contribution with much lower intensity is observed at around 315 and 333 nm. The spectra of IBU_{ratio} and IBU_{beTa} are quite similar; it seems that the contributions from the further added chemical ingredients are spectroscopically comparable. The spectrum for ASP_{BAYER} is dominated by a broad intensity distribution at around 304 nm. Here, two contributions of different intensity are specifiable, a more intense with maximum at 304 nm and a weaker one at 333 nm. For PAR_{ratio}, only one strong contribution with maximum at 310 nm is observed; whereas THO shows two contribution of different intensity, a more in-tense with maximum at 304 nm and a weaker one at 333 nm. In the range 225–275 nm, all painkiller samples show a minor peak in their reflectance at around 270 nm. Towards lower wavelength, the spectra show no additional features.

The shape of the spectra of the commercial painkiller match those of the APIs (Figure 3.6b) quite well. Slight deviations are most likely due to additional ingredients in the commercial samples.

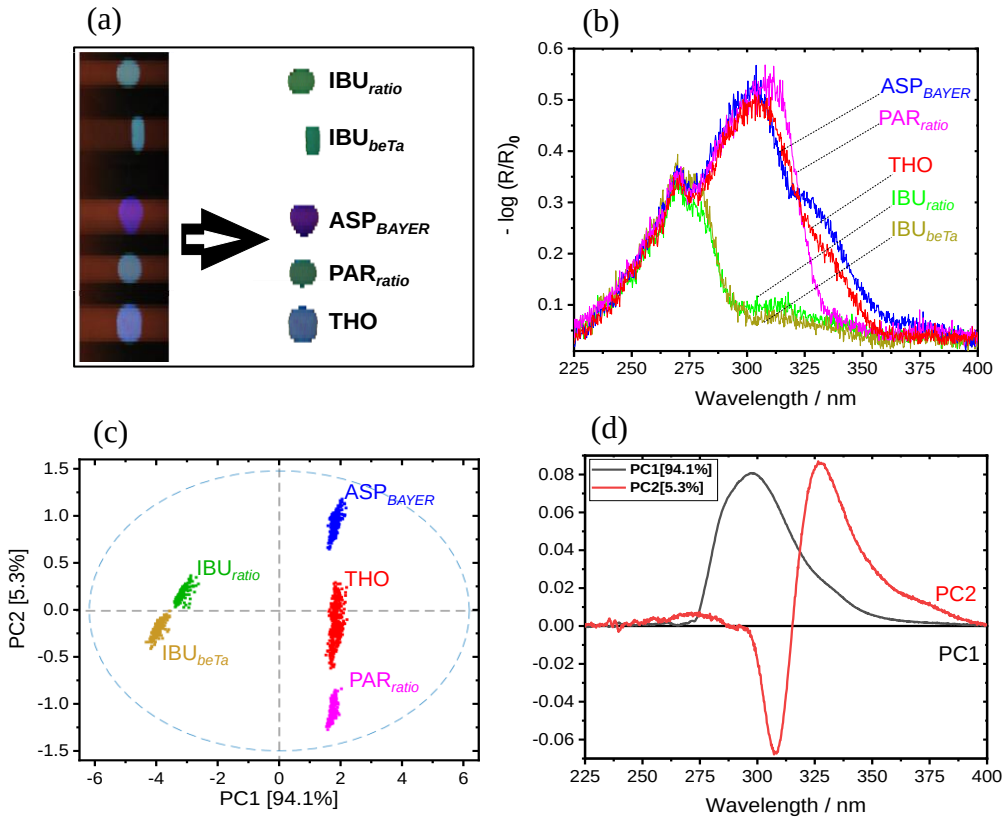


Figure 3.7: (a) Raw hyperspectral image for all commercial painkiller tablets before and after subtracting the background. (b) Spectrum recorded for a single pixel of each painkiller tablet in the UV range 200–400 nm. (c,d) Scores and corresponding loadings plot.

Figure 3.7d shows the loadings plot for PC1 and PC2. PCA model was calculated by cross validation (venetian blinds, 10 splits, 1 sample per split). The loading of PC1 is dominated by an overall positive contribution in the range between 275 and 350 nm, whereas PC2 shows one more narrow negative contribution at 308 nm and one positive at 327 nm. The distribution of the clusters in the scores plot in Figure 3.7c shows a comparable variability with the scores plot in Figure 3.6c. Only for the THO sample an increased spreading is observed along PC2. In general, such type of variability in the shape of the cluster can arise for several reasons; a change in the sample's properties on the scale of the resolution actually achieved, or changes due to shape effects of the samples or positioning within the hyperspectral imaging setup. A general reason for deviations between the hyperspectral imaging and UV-Vis spectroscopy (see Figure 3.5 vs. Figure 3.6b or Figure 3.7b) are the different geometries used for illumination and detection in the setups. In the UV spectrometer the light is collected in an almost perfectly reflecting integrating sphere, while in case of the UV hyperspectral imaging, a tunnel made of PTFE is used for illumination and collecting as shown in Figure 3.2a. As a consequence, a clear differentiation between specular

and diffuse reflection is not possible in the hyperspectral imaging setup, therefore, a mixture of both contributions will be detected here.

Comparing the hyperspectral imaging spectra in Figure 3.6b or Figure 3.7b with the spectra given in Figure 3.5 it is clear that the hyperspectral imaging data provide almost no useful spectroscopic information in the region < 275 nm. The low performance in this range is due to the efficiency of detector and the illumination in the hyperspectral imaging setup. A further consequence of this is that contributions at higher wavelengths appear more dominant as they actually are. The tendency of increasing sensitivity exists for the entire wavelength range. This is also why the shoulders observed in the spectra of ASA/ASP and IBU at > 25 nm (in both sample sets) appear much more enhanced compared to the spectra in Figure 3.5. As a consequence, the actual hyperspectral imaging setup yields valuable results for all samples, but reliable spectroscopic information is only accessible in the range above 275 nm, and there, attention must be paid to the relative intensities. Despite the spectroscopy weaknesses, the combination of UV hyperspectral imaging and chemometric modeling enables a complete separation of all samples in both sample sets. The loadings plots (Figure 3.6c or Figure 3.7c) indicate that a differentiation of all samples is possible considering only a few spectral channels, so that rapid classification is easily possible.

In order to validate the pure API PCA model (see Figure 3.6), the scores of PC1 and PC2 were used to calculate a quadratic discriminant analysis (QDA). The confusion matrix resulted from this model is listed in Table 3.3. A confusion matrix describes the performance of the classification model based on QDA. An overall accuracy for the pure API tablets of 99.8% is reached, which means the model can correctly classify approximately all spectra of the pure API tablets. The highlighted diagonal describes how many spectra were predicted by the model as true. Only 19 spectra of $(ASA+PAR)_{pure}$ were predicted as PAR_{pure} and two spectra of PAR_{pure} as $(ASA+PAR)_{pure}$. This is because $(ASA+PAR)_{pure}$ contains both API components (ASA_{pure}, PAR_{pure}).

Table 3.3: The confusion matrix of the pure API spectra.

		Predicted			
	API Samples	IBU_{pure}	ASA_{pure}	PAR_{pure}	$(ASA+PAR)_{pure}$
Actual	IBU_{pure}	2365	0	0	0
	ASA_{pure}	0	2428	0	0
	PAR_{pure}	0	0	2574	2
	$(ASA+PAR)_{pure}$	0	0	19	2586

This QDA model was used to classify all spectra of the painkiller tablets. Even 99.8% of the spectra were predicted correctly (see Table 3.4). This means approximately all painkiller tablets were predicted correctly in true API classes. Only two spectra of ASP_{BAYER} were assigned as

(ASA+PAR)$_{pure}$. This is because (ASA+PAR)$_{pure}$ contains both API components (ASA$_{pure}$, PAR-$_{pure}$).

Table 3.4: Classification of the painkiller tablets based on the pure API model.

		Predicted			
Actual	Samples	IBU$_{pure}$	ASP $_{pure}$	PAR$_{pure}$	(ASA+PAR)$_{pure}$
	IBU$_{ratio}$/IBU$_{beTa}$	469	0	0	0
	ASP $_{BAYER}$	0	283	0	2
	PAR$_{ratio}$	0	0	209	0
	THO	0	0	0	394

The UV region is often preferred in process control and quality assurance, but hyperspectral imaging in this region is rarely reported. The aim of this study was to develop a simple UV hyperspectral imaging setup capable of distinguishing between different drug samples as an example for a possible industrial application. With the prototype, a painkiller table can be measured at 4 s. This speed is adequate for scientific purposes, but too low for industrial applications. The limiting factor towards a setup for a production environment is the intensity of the illumination and the quantum yield of the pushbroom imager. With an appropriate light source and imager then this setup is capable for in-line data acquisition, process control, in-line classification/sorting, and thus real-time release testing.

3.5 Conclusions

UV hyperspectral imaging was used to characterize active pharmaceutical ingredients in tablets. Two sample sets were analyzed; sample set one consisted of tablets with 100% API content and sample set two consisted of commercially available painkiller tablets. Reference measurements were performed on the pure APIs in liquid solutions and in solid phase using a commercial UV spectrometer.

Hyperspectral imaging in combination with PCA is a promising approach for the detection and differentiation of all drug samples studied. The PCA model was able to separate all drug types with the first two principle components. The advantage of the home-built setup is a high spatial/spectral resolution and a data acquisition speed completely sufficient for scientific studies. Based on the design and the data shown, a setup fulfilling the requirements of a real industrial process can be easily realized.

Author Contributions: Conceptualization, E.O. and M.B.; methodology, M.A.K., M.S., B.B. and E.O.; software, M.A.K, M.S. and B.B.; validation, M.A.K., M.S., B.B. and E.O.; formal analysis, M.A.K. and M.S.; investigation, M.A.K. and M.S.; resources, M.B.; data curation, M.A.K., M.S. and B.B.; writing—original draft preparation, M.A.K. and M.B.; writing—review and editing, M.A.K., M.S., B.B., E.O. and M.B. ; visualization, M.A.K and B.B.; supervision, M.B.; project administration, M.B.; funding acquisition, M.A.K., E.O. and M.B.

Funding: Mohammad Al Ktash acknowledges the support of Katholischer Akademischer Ausländer-Dienst (KAAD).

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy restrictions.

Acknowledgments: The authors thank Karsten Rebner, Tim Bäuerle and Tobias Drieschner for valuable discussions.

Conflicts of Interest: The authors declare no conflict of interest.

4 Paper II: UV Hyperspectral Imaging as Process Analytical Tool for the Characterization of Oxide Layers and Copper States on Direct Bonded Copper

Mohammad Al Ktash[1,2,†], Mona Stefanakis[1,2,†], Tim Englert[3,4], Maryam S. L. Drechsel[1], Jan Stiedl[3], Simon Green[3], Timo Jacob[4], Barbara Boldrini[1], Edwin Ostertag[1], Karsten Rebner[1] and Marc Brecht[1,2,]*

[1]Process Analysis and Technology PA & T, Reutlingen University, Alteburgstraße 150, 72762 Reutlingen
[2]Institute of Physical and Theoretical Chemistry, Eberhard Karls University Tübingen, Auf der Morgenstelle 18, 72076 Tübingen, Germany
[3]Robert Bosch GmbH, Automotive Electronics, Postfach 1342, 72703 Reutlingen, Germany
4Institute of Electrochemistry, Ulm University, Albert-Einstein-Allee 47, 89081 Ulm, Germany
[*]Correspondence: Marc.Brecht@Reutlingen-University.de
[†]These authors contributed equally to the work.

This is originally published in sensors (https://doi.org/10.3390/s21217332) as

"Al Ktash, M.; Stefanakis, M.; Englert, T.; Drechsel, M.S.L.; Stiedl, J.; Green, S.; Jacob, T.; Boldrini, B.; Ostertag, E.; Rebner, K.; Brecht, M. UV Hyperspectral Imaging as Process Analytical Tool for the Characterization of Oxide Layers and Copper States on Direct Bonded Copper. Sensors 2021, 21, 7332. https://doi.org/10.3390/s21217332"

4.1 Abstract

Hyperspectral imaging and reflectance spectroscopy in the range from 200–380 nm were used to rapidly detect and characterize copper oxidation states and their layer thicknesses on direct bonded copper in a non-destructive way. Single-point UV reflectance spectroscopy, as a well-established method, was utilized to compare the quality of the hyperspectral imaging results. For the laterally resolved measurements of the copper surfaces an UV hyperspectral imaging setup based on a pushbroom imager was used. Six different types of direct bonded copper were studied. Each type had a different oxide layer thickness and was analyzed by depth profiling using X-ray photoelectron spectroscopy. In total, 28 samples were measured to develop multivariate models to characterize and predict the oxide layer thicknesses. The principal component analysis models (PCA) enabled a general differentiation between the sample types on the first two PCs with 100.0% and 96% explained variance for UV spectroscopy and hyperspectral imaging, respectively. Partial least squares regression (PLS-R) models showed reliable performance with R^2_c = 0.94 and 0.94 and RMSEC = 1.64 nm and 1.76 nm, respectively. The developed in-line prototype system combined with multivariate data modeling shows high potential for further development of this technique towards real large-scale processes.

4.2 Introduction

Copper is considered as one of the most important conductors for integrated circuit (IC) packaging and wire bonding. It has significant advantages in comparison to other materials (e.g., aluminum) and is thus a good alternative for smaller structures. Copper as a metal has a high mechanical stability and excellent electrical and thermal conductivities at low cost [139]. However, copper contact surfaces contaminate and interact with oxygen to copper (I) oxide (Cu_2O) and copper (II) oxide (CuO) layers. This process is considered a problem as it influences the conductivity efficiency. Science and engineering progress has driven the development of sensor technology in the past years [122,140]. This led to novel optical sensors, such as hyperspectral imagers, to identify quality problems [20,141].

Hyperspectral imaging is a technique that integrates a conventional spectroscopic system with imaging in order to acquire spectral and spatial information from the area of interest [21,39,86,142]. Therefore, hyperspectral imaging enables quantitative analysis with improved levels of accuracy [19,23,143]. It is considered as a rapid, non-destructive and robust method. Combining spectral

imaging with chemometric algorithms opens up new industrial applications, including manufacturing process control [144]. Such spectral imaging systems are used in different fields, such as food, pharmaceutical and textile production, as well as agriculture, military, astronomy, life sciences and medicine [20,25,26,128,141,145].

Hyperspectral imaging is able to capture images in different spectral bands, such as in the visible (Vis), infrared (NIR) and ultraviolet (UV) range. In contrast, traditional methods, such as Auger electron and X-ray photoelectron spectroscopy (XPS), which are used to analyze copper samples, are time consuming, expensive and require sample preparation and destruction [87,146]. The industry demands a high lateral resolution, which cannot be fulfilled by single-point UV-Vis spectroscopy [19,147]. Several detection methods have been developed to classify and identify the copper state and copper oxide layers. In the past, UV-Vis/NIR spectroscopic applications as well as Vis/NIR hyperspectral imaging have been preferred in the industrial environment, especially for copper and other metal conductors [26,87,146,148-150]. The detection and characterization of oxide layers on metallic copper samples was studied by Stiedl et al. using visible hyperspectral imaging and UV-Vis spectroscopy. They were able to detect the thickness of the oxide layers on the technical copper [26,87].

Recently, Tschannerl et al. have shown the application of hyperspectral imaging in the UV range to discriminate between phenolic flavor concentrations in melted barley [86]. In another recently published study, Al Ktash et al. have developed this technology in the direction of real applications. The authors were able to precisely classify between different active pharmaceutical ingredients (API) and painkiller tablets by using an UV hyperspectral imaging prototype [19].

Hyperspectral imaging collects information in three dimensions (x, y, λ), resulting in a massive number of variables. Therefore, data reduction algorithms, such as principal component analysis (PCA) and partial least squares regression (PLS-R), are required. PCA combined with hyperspectral imaging data enables the detection of spectral features in the spectroscopic data along with identifying the relative distribution of the components in mixtures [38,151]. The PLS-R is an empirical data-driven modelling approach that relies on representative model building data for two variable blocks (X and Y). It is used to search for a correlation between a simple and easily acquirable data set (X) and a labor- as well as cost-intensive second set of measurements (Y) by calculating a certain number of factors. In the present study, the X data contains the UV spectra, and the Y data the oxide layer thickness of the direct bonded copper sheets. Consequently, quantitative descriptions and calibrations are possible [152].

Despite several studies having focused on the characterization of copper oxide films, sample homogeneity remains a big challenge in the estimation of their thicknesses over the complete surface. We address this topic in the present contribution using a hyperspectral imaging system in the UV wavelength range for the in-line characterization of copper states and oxide layers thicknesses on direct bonded copper. The data were evaluated by PCA and PLS-R. The results show that hyperspectral imaging in the UV range has the potential to predict oxide layer thicknesses and copper states in a rapid and non-destructive manner.

4.3 Materials and Methods

4.3.1 2.1. Samples

In total, 28 direct bonded copper Curamik® Power substrates (Rogers Corporation, Chandler, AZ, USA) with dimensions of 21.0 mm × 21.0 mm × 1.1 mm were used for sample preparations. The samples were first ultrasonically cleaned at 50 °C for 5 min with Vigon A 200 (Zestron, Ingolstadt, Germany) as cleaning medium and then rinsed with deionized water for 3 min. The copper sheets were oxidized at five different preparation protocols (see Table 4.1). Sample type 1 was left in its initial condition. Figure 4.1 shows an example of each copper sheet type.

Table 4.1: Sample preparation protocol for the direct bonded copper substrates

Sample Type	1	2	3	4	5	6
Number of measured samples	5 *	4	5 *	5 *	5 *	4
Temperature/°C	-	110.0	142.5	142.5	175.0	175.0
110.0	-	2	11	20	11	20
Time/min	0	4.0	6.0	8.3	14.0	21.1
Mean oxide layer thickness/nm	0	5.9	3.0	4.5	7.0	8.2

* One of each sample set was used for PLS-R prediction.

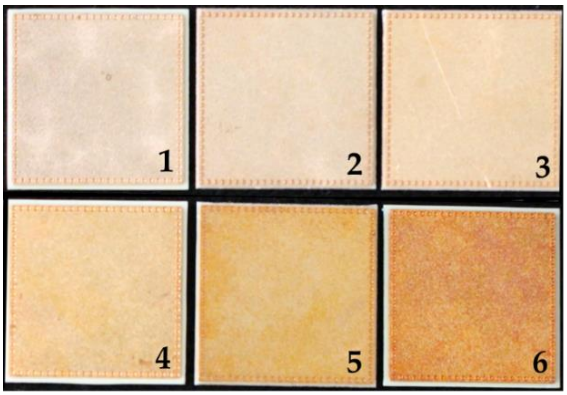


Figure 4.1: Direct bonded copper Curamik®Power substrates. (1) is an example of sample type 1, (2) sample type 2, (3) sample type 3, (4) sample type 4, (5) sample type 5 and (6) sample type 6.

4.3.2 Oxide Layer Thickness Measurement

The thicknesses of the oxide layers were determined by depth profiling using X-ray photoelectron spectroscopy (XPS). The measurements were conducted under a system base pressure of 4.0×10^{-10} mbar. A monochromatic Al Kα radiation was used and the anode tube operated at 12.5 kV with 20 mA. The take-off angle for the electrons was 0° with respect to the surface normal. The XPS core level spectra were measured with a standard X-ray source SPECS XR50 (SPECS Surface Nano Analysis GmbH, Berlin, Germany) and a concentric hemispherical analyzer Phoibos 100, SPECS (SPECS Surface Nano Analysis GmbH, Berlin, Germany). The pass energy of the concentric hemispherical analyzer was 50 eV for the survey and 20 eV for the high-resolution spectra. The data acquisition was performed with 0.5 eV; 0.1 eV per step, respectively.

4.3.3 UV Spectroscopy

Total (specular and diffuse) reflectance spectra were recorded in the range of 200–380 nm using a UV spectrometer (Lambda 1050+, PerkinElmer, Inc., Waltham, MA, USA). The 150 mm inte-grating sphere module functioned as a detection unit and was deployed in reflectance with a R6872-Photomultiplier (PMT). A deuterium lamp was used as light source in the spectrometer. The samples were placed at the reflectance port of the inte-grating sphere with a diffused scattering Spectralon® disk placed behind the samples. The port measuring area is approximately 0.42 cm^2. Three spectra were recorded for each direct bonded copper type while the sample was rotated in different angles (see Figure 4.2). The UV spectra were recorded with the Lambda 1050 UV WinLab software from PerkinElmer.

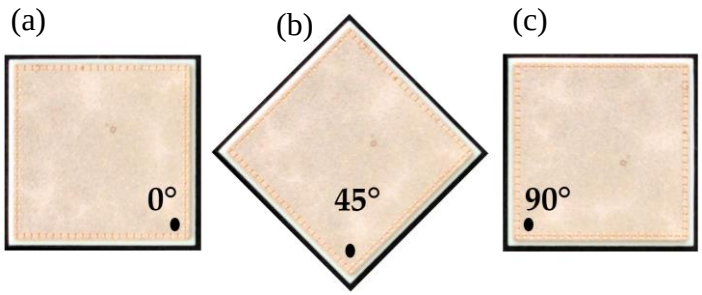


Figure 4.2: An example of a direct bonded copper sheet rotated according to the three different measurement angles (a) 0°, (b) 45° and (c) 90°.

4.3.4 Data Collection and Preprocessing

The hyperspectral imaging setup was optimized compared to our previous work [19]. The pushbroom imager is a BlueEye Tec (inno-spec GmbH, Nürnberg, Germany), consisting of a

spectrograph (RS 50–1938, inno-spec GmbH, Nürnberg, Germany), with a slit width of 80 µm, connected to a back-illuminated CMOS camera with total size 2048 × 2048 pixel (spatial × spectral) and pixel size of 6.5 µm × 6.5 µm. Additionally, the dispersion is approximately 0.1 nm/px [153]. The quantum efficiency of the CMOS camera is between 30 and 50% [154]. The optimal integration time was 10 ms. The samples were placed on a black conveyor belt (700 mm × 215 mm × 60 mm, Dobot Magician, Shenzhen Yuejiang Technology Co., Ltd., Shenzhen, China) moving with a constant speed of 0.15 mm/s, which was positioned completely in a tunnel made of PTFE. The illumination was provided by two ozone producing Xenon lamps (XBO, 14 V, 75 W, OSRAM, München, Germany). The ozone was eliminated by a laboratory vacuum system (AirTracker, TEKA Absaug- und Entsorgungstechnologie GmbH, Coesfeld, Germany). Another xenon lamp was added to the setup to increase the intensity and optimize the integration time. In combination with the black conveyor belt and a state-of-the-art UV pushbroom imager a more industrial-like prototype was created.

The principal and workflow of the data acquisition remained [19]. The UV hyperspectral imaging data were acquired by the FluxRecorder version 4.2.1.17 (inno-spec GmbH, Nürnberg, Germany). The reflectance was calculated by the FluxRecorder automatically according to the radiometric calibration [19,21,39,155]. PTFE was used as white reference. For collecting the dark reference, the objective was closed by its cover and the illumination was turned off.

Figure 4.3 shows the original images of the direct bonded copper samples before and after background subtraction. Hyperspectral data matrices were analyzed by Evince version 2.7.11 (Prediktera AB, Umeå, Sweden). While importing the raw data in Evince, a data reduction was performed by binning four columns and rows (x,y) and six channels (λ).

The background was removed by calculating a PCA and selecting the corresponding background scores. Therefore, some edges and borders of the samples were also eliminated, resulting in different sample shapes (Figure 4.3). The reduced hypercube was then used as input for the subsequent PCA and PLS-R. In the end, approximately 2.0 million spectra remained from the initially obtained 4.0 million spectra.

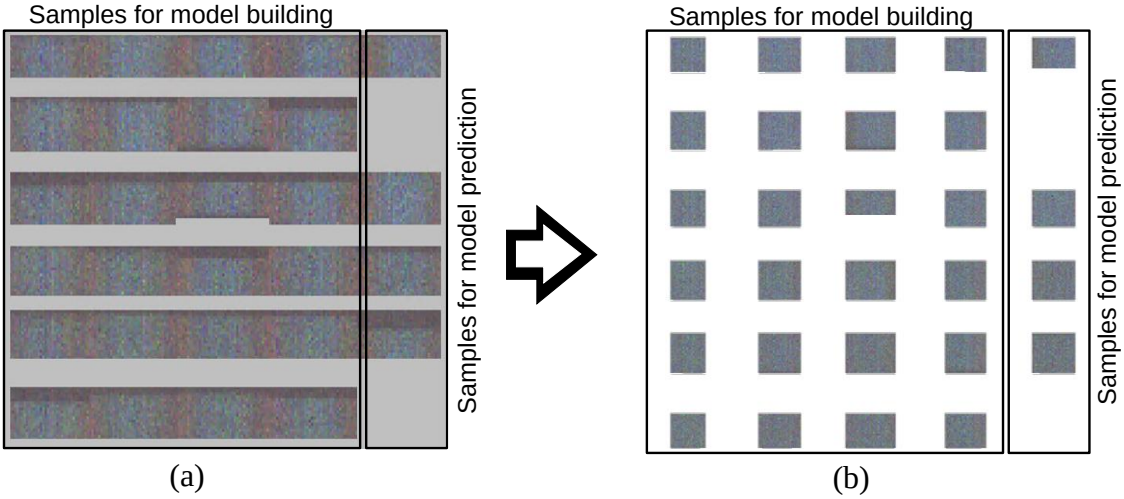


Figure 4.3: Hyperspectral raw images of 28 direct bonded copper samples on the left (a). Images after subtraction of the background on the right (b). In total, 24 samples were used for building the PLS-R model and four samples were used for prediction.

4.3.5 Multivariate Data Analysis and Data Handling

Multivariate data analysis (MVA) was performed with "The Unscrambler X 10.5" (Camo Analytics AS, Oslo, Norway). All spectra recorded by UV hyperspectral imaging and commercial spectroscopy were preprocessed in the same way: Gaussian smoothing with 15 points reduction in the range from 200 nm to 380 nm. The spectral resolution of the hyperspectral imaging data was further reduced to 1 nm by averaging to ensure comparability to the UV spectra of the single-point spectrometer. The principal component analysis (PCA) was calculated with mean centering, cross-validation and the NIPALS algorithm to distinguish between the direct bonded copper sample types.

Partial least square regression (PLS-R) models for the oxide layer thickness prediction were created with mean centering, full cross-validation and the Kernel algorithm. Four direct bonded copper sheets of each preparation type were used to develop the PLS-R model. Additionally, the remaining samples of copper type 1, 3, 4 and 5 were used as prediction samples to test the final PLS-R model. The predicted values were compared to the determined oxide layer thicknesses by XPS. Finally, the oxide layer thickness of each pixel of the remaining samples was predicted by the hyperspectral imaging PLS-R model. The distribution map thus generated was visualized by MATLAB (R2020b 9.9.0, Mathworks, Natick, MA, USA). The samples were binned by factor 5 in the x and y direction due to the large amount of data and noise.

4.4 Results and Discussion

4.4.1 UV Spectroscopy

Direct bonded copper substrates were investigated using diffuse reflectance spectroscopy in the UV region (200–380 nm). In total, 28 samples were measured. Generally, the thickness of the oxide layers increases with the oxidation time and temperature. During the oxidation process, copper is oxidized first to copper (I) oxide (CuO_2) and then to copper (II) oxide (CuO). Figure 4.4a shows the preprocessed reflectance spectra. Based on the shape of the spectra, the different steps of the oxidation process can be observed. Sample type 1 is representing copper in its initial condition. The other samples have undergone an oxidation process, as detailed in Table 4.1. A band minimum is detected approximately at 220 nm. A pronounced band maximum for all copper samples occurs in the wavelength range from 315 to 320 nm. Weak shoulders at 243 nm (sh) and 266 nm (sh) are observed. Sample types 1, 2 and 3 present one prominent maximum at 295 nm. Sample types 4, 5 and 6 show a distinct band with maximum at 378 nm. The band at 220 nm could be ascribed to Cu_2O. Increasing Cu_2O pronounces the minimum. The band at 295 nm is assigned to the copper material (see Appendix Figure 10.1). This band started to fade away due to the increase in the maximum band at 320 nm. This band is absent in sample types 4, 5 and 6. For these sample types a band at 378 nm appears. These spectral differences were due to different oxide layer thicknesses and copper states (Cu^0, Cu_2O and CuO) on the copper sheets. The remaining small differences among the spectra were attributed to the roughness, measuring angles and sample positions.

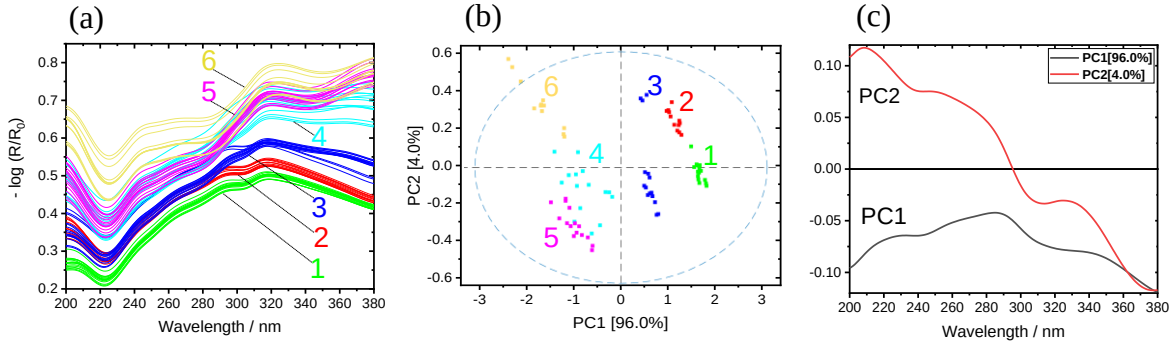


Figure 4.4: (a) UV reflectance spectra of copper sheets. Copper with initial condition type 1 (green), 2 (red), 3 (blue), 4 (light blue), 5 (pink) and 6 (yellow) represent the oxidation layer thicknesses 0 nm, 4 nm, 8.3 nm, 14 nm and 21.1 nm, respectively. (b) PCA with scores and (c) the corresponding loadings plot.

Figure 4.4b shows the scores plot of the first two principal components (PC). The first two PCs explain nearly 100.0% of the total variance. The scores of different sample types are clearly distinguished. Every copper sample type with a corresponding copper state and oxide layer thickness

appear as a distinct group. PC1 yields a clear separation of copper in the initial condition (type 1) from the other copper types. The groups move below the average in PC1, with increasing oxide layer thickness and conversion of copper states. Copper types 4 and 5 are slightly overlapped as their oxide layer thicknesses are almost comparable (see Table 4.1). The variance in each cluster results from the different samples for each type. The differences between the samples could be due to temperature profiles in the oven while preparing the samples, roughness variation, or sample positioning during the measurements.

The loadings plot for PC1 and PC2 is given in Figure 4.4c. The shape of PC1 resembles the Cu^0 spectrum (see Appendix Figure 10.1 and Table 10.1). This indicates that an increasing amount of Cu^0 on a sample results in a more positive sample arrangement on PC1. Vice versa, the less Cu^0 is present in the samples because of the growing oxide layer thickness, the more the samples are shifted in the negative range of PC1. The influence of the oxidation state (Cu_2O, CuO) is expressed by PC2 (see Appendix Figure 10.1 and Table 10.1); these results are comparable with previous studies [26].

4.4.2 UV Hyperspectral Imaging

All samples were analyzed by a UV hyperspectral imaging prototype, as described in Materials and Methods. In order to make the data more comparable to the UV spectroscopy, the average spectra were calculated to reduce the number of spectra. A total of 25 spectra was determined from the hyperspectral imaging data for each of the 28 samples. Figure 4.5a shows the results of the UV hyperspectral imaging in the range from 200 to 380 nm.

The comparison between the shapes of the spectra is given in Figures 4.4a and 4.5a, showing similarities as well as a small deviation. They are due to the type of the illumination source and the design of the experimental setups. For reflectance spectroscopy, a deuterium lamp was used, while for hyperspectral imaging, two xenon lamps were available. Deuterium lamps have higher spectral irradiances in the deep UV range compared to xenon lamps [22]. However, the xenon illumination was sufficient for the characterization of direct bonded copper sheets. Therefore, the interferences < 270 nm are more pronounced compared to the higher wavelengths. As a result, the spectra shown in Figure 4.5a provide almost no clearly recognizable spectroscopic information in the region <270 nm. The detector's efficiency and illumination provide low performance in this wavelength range. Therefore, the easily accessible tunnel design for hyperspectral imaging was developed to ensure a diffuse illumination of the samples. As a result, a reasonable illumination strength and homogeneity were reached.

As discussed before, the spectra were influenced by the copper states and thicknesses of the oxide layers on the copper sheets. Copper in the initial condition is represented in the spectra originating from sample type 1 (see Table 4.1, Figure 4.4a). The most dominant contributions for all copper sample types are observed in the wavelength range 324–328 nm and at 241 nm. Copper types 1, 2 and 3 present a weak shoulder at 292 nm.

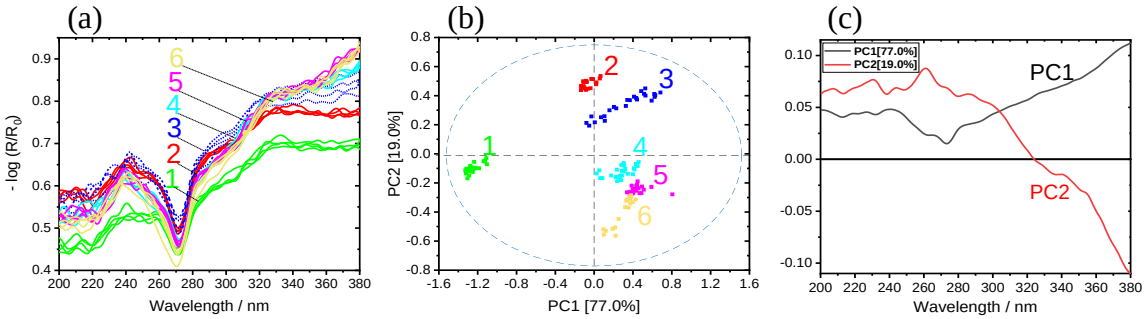


Figure 4.5: (a) Average UV hyperspectral imaging spectra of copper sheets. Copper with initial condition type 1 (green), 2 (red), 3 (blue), 4 (light blue), 5 (pink) and 6 (yellow) represent the oxidation layer thicknesses 0 nm, 4 nm, 8.3 nm, 14 nm and 21.1 nm, respectively. (b) PCA with scores and (c) the corresponding loadings.

In the next step, a PCA model with a cross-validation was calculated for the average spectra of all samples. Figure 4.5b shows the scores plot of PC1 and PC2. The first two PCs explain nearly 96.0% of the total variance. The scores of different sample types are clearly distinguished. Every copper sample type with a corresponding copper state and oxide layer thickness appears as a distinct group. PC1 yields a clear separation of copper with initial condition (type 1) from the other copper types. A discrimination of the copper state and oxide layer thickness is observed on PC2. Beginning from the positive to the negative scores on PC2, the samples are arranged in the order copper type 2, 3, 4, 5 and 6, respectively. Again, copper type 2 and 3 (positive scores) can be separated from the other samples 4, 5 and 6 (negative scores).

The loadings plot for PC1 and PC2 is given in Figure 4.5c. PC1 shows the differences between Cu^0 and the oxidation states (Cu_2O, CuO). The most dominant contribution is observed in the range from 260 to 280 nm and the increasing shape > 280 nm. The loadings plot of PC2 mainly shows increasing oxide layer thickness. The most prominent contribution is observed in the range from 250 to 280 nm and the decreasing shape > 280 nm. Compared to PC1, PC2 has a positive maximum at 263 nm. The minimum on PC1 is located at 273 nm. This region could include the information about the copper state. The influence of the oxidation state (Cu_2O, CuO) and oxide layer thickness is observed by PC2.

Figure 4.4a presents UV spectra with a good signal-to-noise-ratio recorded by a UV spectrometer, which collected one single spectrum over an area of 0.42 cm2. Figure 4a shows the UV spectra recorded by the hyperspectral imaging setup. The spectra were averaged over an area size comparable to the UV spectrometer. The UV hyperspectral imager recorded raw spectra with a less good signal-to-noise-ratio. These spectra result from one single pixel of the detector, representing a much smaller area of the direct bonded copper, which is estimated to be 6.5 µm × 6.5 µm. Additional reasons for the low signal-to-noise-ratio are the weak irradiation intensity by the xenon illumination and the quantum efficiency of the camera in this spectral range of approximately 30–50% [154]. Furthermore, ozone-producing xenon lamps were used. With the help of a vacuum system, the influence of the ozone absorption at 250 nm was minimized.

The benefit of hyperspectral imaging is lateral information in real time. To get a visual impression of the inhomogeneity of the copper states and oxide layer thickness, the thickness for every pixel from the first two PCs was plotted as a distribution map, shown in Figure 4.6. A sample with high absorbance has a high proportion of blue in the score image (e.g., Cu^0), while one with low absorbance shows a higher proportion of red (e.g., Cu(II)). Clear differences between the samples are observed according to the oxidation time and temperature. As discussed before, PC1 yields a clear separation of copper in the initial condition from the other copper types. A discrimination of the copper state and oxide layer thickness can be observed on PC2. The regular distribution of the pattern in PC2 indicates a common origin; this could be the variability of the temperature inside the oven among each sample. Additionally, in the distribution maps, it is possible to clearly identify oxidation hotspots on the direct bonded copper.

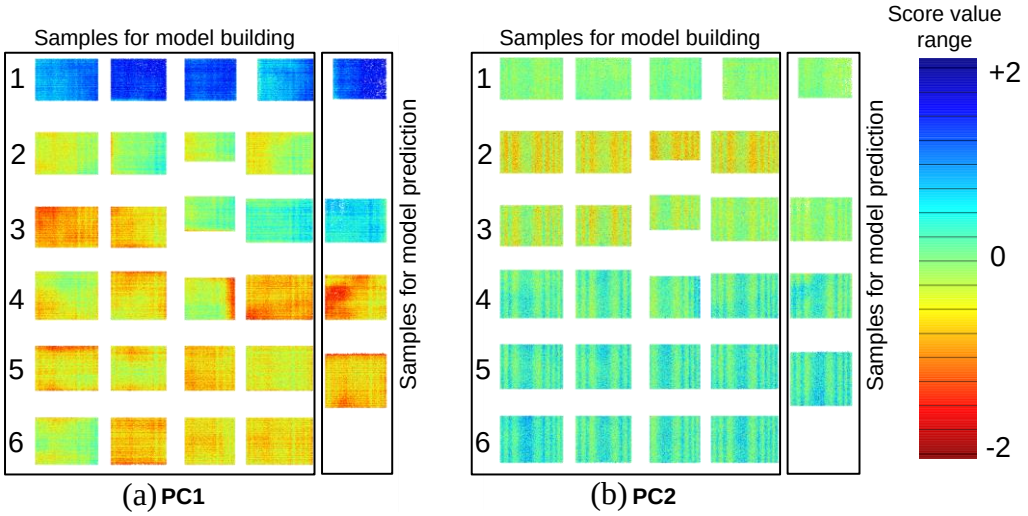


Figure 4.6: Distribution maps of the oxide layer PC1 (a) and PC2 (b). Each rectangle represents a single copper sheet. The sample type for each row corresponds to Table 1. The samples are divided into two sets: model building and model prediction for PLS-R. The colored pixels (the score value range) represent the oxide content, from low (blue) to high (red).

4.4.3 PLS-R

A PCA structures data sets according to their maximum variance, whereas PLS-R searches for the optimal correlation between spectral characteristics and an external target value. PLS-R models of the direct bonded copper for each method have been established and compared, by using the spectra of the UV reflectance spectroscopy and UV hyperspectral imaging. In this study, spectral features were extracted from the spectral datasets and correlated to determine the oxide layer thickness via XPS.

Gaussian smoothing with 15 points was performed to minimize the noise. A PLS-R model was developed with a calibration set of n = 24 samples (see Figure 4.6), three factors, the Kernel algorithm and full cross-validation. A prediction sample set was used to test the PLS-R model performance with an external validation to assess the predictive ability. The prediction sample set consisted of four samples with mean oxide layer thicknesses of 0 nm, 6 nm, 8.3 nm and 14 nm (see Figure 4.6). Table 4.2 summarizes the overall chemometric model results for both the UV spectroscopy and hyperspectral imaging.

The number of factors for each PLS-R model was optimized according to a high coefficient of determination (R^2) and a low root mean square error of calibration (RMSEC) and cross-validation (RMSECV). This approach was applied to both the calibration (R^2_c) and cross-validation (R^2_{cv}) model for each method (Table 4.2).

Table 4.2: Model statistics for the calibration and full cross-validation models for oxide layer thickness on the direct bonded copper.

Method	**Number of Factors**	**Parameters Calibration**		**Parameters Validation**	
		R^2_c	**RMSEC/nm**	R^2_{cv}	**RMSECV/nm**
UV spectroscopy	3	0.94	1.64	0.93	1.74
UV hyperspectral imaging	3	0.94	1.76	0.93	1.88

The variances explained by the UV reflectance model for the X and Y variables were 99.0% and 95.0%, respectively, by using three factors. The variances of the X and Y variables were 98% and 94% for the UV hyperspectral imaging model, by using three factors as well. This indicated that three PLS components (factors) were sufficient to describe most of the variance in the data according to the spectral information.

The results show that the PLS-R models are very effective in correlating the oxide layer thickness with both spectroscopic data sets. This is indicated by a high R^2_c and a low RMSEC and a high R^2_{cv} with a low RMSECV (see Table 4.2). Figure 4.7a,b show the correlation between the refer-

ence and predicted values of the UV reflectance spectra and UV hyperspectral imaging, respectively. The deviation and the variance within a sample type are increasing according to the oxide layer growth on the direct bonded copper. Copper sample types 1, 2 and 3 have a smaller variance within the sample type. In contrast, sample types 4, 5 and 6 have more variance in the UV reflectance spectra model. For UV hyperspectral imaging all samples have nearly the same variance. This variance is probably due to the efficiency of the detector and the illumination in both setups.

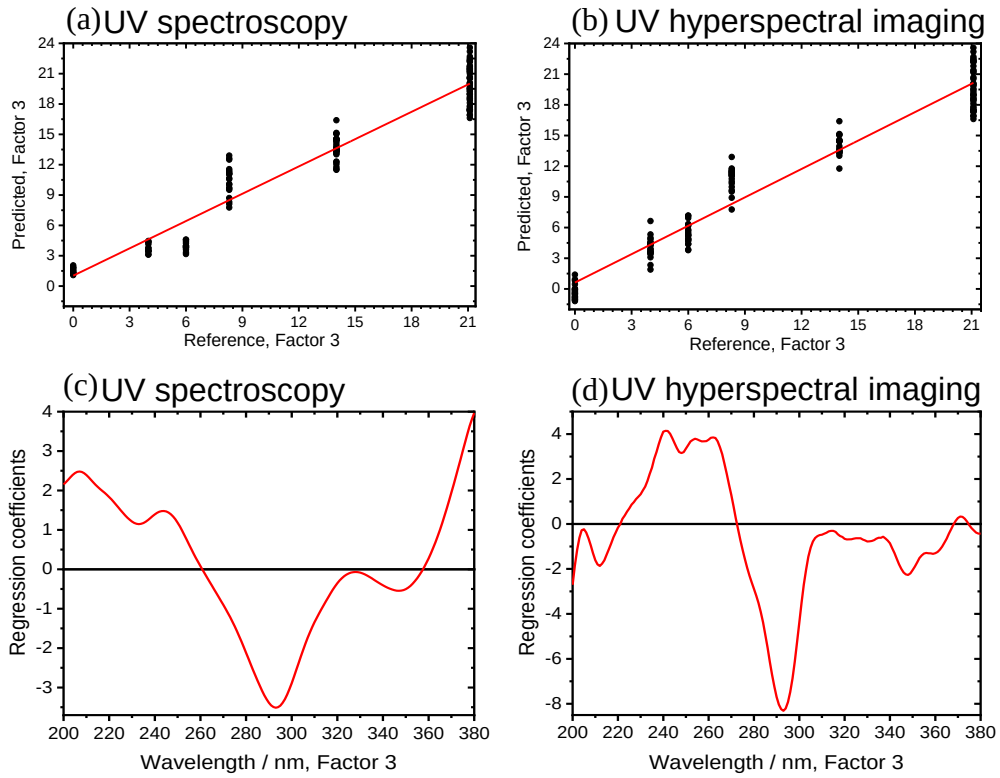


Figure 4.7: Three-factor PLS-R models for the oxide layer thicknesses of direct bonded copper in the UV region (200–380 nm). (a) Predicted vs. reference of UV spectra. (b) Predicted vs. reference of UV hyperspectral imaging. (c) Regression coefficients of the UV spectra. (d) Regression coefficients of the UV hyperspectral imaging.

The regression coefficients of the three-factor UV spectroscopy PLS-R model are shown in Figure 4.7c. Again, absorbance bands around 210 nm, 245 nm, 293 nm and 330 nm emerge, as displayed in the spectra. Above 360 nm, an increasing baseline in the regression coefficient plot is registered. In Figure 4.7d, the corresponding regression coefficients of the UV hyperspectral imaging PLS-R model are displayed. They have a comparable shape, but more details can be detected. For example, in the range <260 nm and from 310 to 340 nm, more spectral features are pronounced. At 370 nm, a defined band appears for the UV hyperspectral imaging model, while an increase >360 nm in the UV spectroscopy model is registered.

Correlated to the bands at <260 nm, 320 nm, 335 nm and >360 nm, the oxide layer thickness increases in the UV spectra, which is also comparable to UV hyperspectral imaging in the range 230–265 nm and 306–340 nm for increasing the oxide layer thickness. At 293 nm, the oxide layer thickness decreases for both setups. Differences in the beginning and ending of the regression coefficients between both methods could be due to the detectors limits with the UV hyperspectral imaging setup, as already discussed in the literature [19,26].

In order to evaluate the PLS-R models, four samples of type 1, 3, 4 and 5 were used to test the model's performance by predicting the oxide layer thickness. The sample set contains four samples with mean oxide layer thicknesses of 0 nm, 6 nm, 8.3 nm and 14 nm (see Figure 4.6).

In addition, the results indicated that the PLS-R was very effective in predicting the oxide layer thickness with three factors, $R^2_p = 0.90$ with RMSEP = 1.62 nm and bias = 0.51 for UV spectra, and $R^2_p = 0.85$ with RMSEP = 1.98 nm and bias = 0.61 for UV hyperspectral imaging.

In Table 4.3, the results for the mean value of the predictions and deviations are given. The predicted values are matched well with the references.

Table 4.3: Prediction of the oxide layer thicknesses for direct bonded copper from PLS-R models.

Method	Sample Type	Reference/nm	Predicted/nm	Deviation/nm
UV spectroscopy	1	0	1.59	0.93
	3	6	6.00	1.02
	4	8.3	7.86	1.44
	5	14	15.25	1.53
UV hyperspectral imaging	1	0	-0.87	1.49
	3	6	5.51	2.08
	4	8.3	11.74	1.91
	5	14	14.35	1.79

Mean values with a high standard deviation were measured by XPS (see Table 4.1), as reference values for the PLS-R models. This average of one sample type is comparable to the UV spectra recording. However, hyperspectral imaging enables to recognize different oxide layer thicknesses among the samples. Therefore, the oxide layer thickness in nm of each pixel was calculated. The UV hyperspectral imaging PLS-R model was applied to the four prediction samples of type 1, 3, 4 and 5. In Figure 4.8, the resulting distribution map is shown. The pixels represent the oxide layer thicknesses in nm, from low (blue) to high (red). Sample type 1 displays the initial direct bonded copper sheet without induced oxidation, while the other samples show an increasing oxide layer thickness.

Although, PLS-R is a robust model to describe the majority of the variance of the data according to the spectral information. Compared to the results of the Vis hyperspectral imaging [26], the UV

hyperspectral imaging models seems to be more robust. This is indicated by the fact that less factors are necessary to achieve a model with better statistic parameters (higher R^2, lower RMSE) by using a new UV prototype.

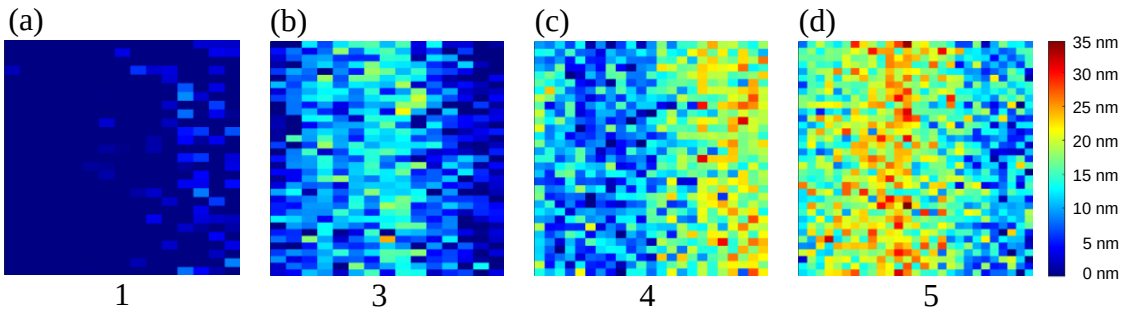


Figure 4.8: Distribution map predicted from the three-factor PLS-R model of the UV hyperspectral imaging data. The oxide layer thicknesses for each pixel of samples (a) sample type 1, (b) sample type 3, (c) sample type 4 and (d) sample type 5 were calculated for model prediction.

Hyperspectral imaging in the UV range is rarely reported, although it is often chosen for process control and quality assurance [19,86]. The aim of this study was to characterize the copper states and oxide layer thicknesses by using a single-point UV spectrometer and a UV hyperspectral imaging setup that can serve as an example for a possible real-time industrial application. With our hyperspectral imaging prototype, a whole direct bonded copper sheet can be measured and processed within 10 s. With the implemented pushbroom imager, hardware binning is also possible, and can decrease the measuring and processing time. The scan speed for the determination of the oxide layer thicknesses on direct bonded copper can be optimized by selecting a few relevant variables instead of the complete UV spectrum. The intensity and type of the illumination are the limiting factors towards a setup for a production environment. This study opens a novel possibility for further development of this method capable of rapid in-line data acquisition, process control and in-line classification/sorting, which meets the requirements of a real-time process with industrial standard and precision.

4.5 Conclusions

UV hyperspectral imaging and UV reflectance spectroscopy (200–380 nm) were used to characterize 28 direct bonded copper samples. UV reflectance spectroscopy, as a well-known method, was utilized to compare the quality of the UV hyperspectral imaging results.

Hyperspectral imaging in combination with PCA and PLS-R is a promising approach for the laterally resolved detection and differentiation of copper states and the determination of oxide layer

thickness in the UV region. The PCA models were able to separate all direct bonded copper types according to the copper states and oxide layer thicknesses, using only the first two principal components. PLS-R models with three factors provided a high R^2 and low RMSE for calibration, validation (n_{cv} = 24) and prediction (n_p = 4). To the best of our knowledge, this is the first work reporting the identification and quantification of copper oxide thin films by UV hyperspectral imaging. The advantage of the home-built setup is the high spatial and spectral resolution and a relatively high data acquisition speed under laboratory conditions. Starting from the presented design and data given in this contribution a setup fulfilling the requirements of a real industrial process can be easily realized.

Supplementary Materials: The following are available online at https://doi.org/10.3390/s21217332/s1, Appendix Figure 10.1: Reference spectra for the copper Cu^0, Cu_2O and CuO by using UV spectrometer, Appendix Table 10.1: Description of the direct bonded copper substrates and their sample preparation.

Author Contributions: Conceptualization, M.A.K., M.S. and K.R.; methodology, M.A.K., M.S., T.E., M.S.L.D., J.S. and B.B.; software, M.A.K. and M.S.; validation, M.A.K., M.S., M.S.L.D. and E.O.; formal analysis, M.A.K. and M.S.; investigation, M.A.K., M.S. and T.E.; resources, T.J., S.G., E.O., K.R. and M.B.; data curation, M.A.K., M.S., T.E. and B.B.; writing—original draft preparation, M.A.K., M.S. and T.E.; writing—review and editing, M.A.K., M.S., T.E., M.S.L.D., J.S., S.G., T.J., B.B., E.O., K.R. and M.B.; visualization, M.A.K. and M.S.; supervision, S.G., T.J., K.R. and M.B.; project administration, K.R., M.B.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing Ph.D. thesis.

Acknowledgments: We especially thank the company inno-spec GmbH (Nürnberg, Germany) for the possibility to test the BlueEye Tec hyperspectral imaging system. Within these test measurements the UV data were generated. We would like to thank Frank Wackenhut for his guidance in programming and presenting the data. The authors also thank Bayan Ayyad for helping in the data handling.

Conflicts of Interest: The authors declare no conflict of interest.

5 Paper III: UV-Vis/NIR Spectroscopy and Hyperspectral Imaging to Study the Different Types of Raw Cotton

Mohammad Al Ktash[1,2], Otto Hauler[1], Edwin Ostertag[1] and Marc Brecht[1,2,]*

[1]Lehr- und Forschungszentrum Process Analysis and Technology (PA&T) der Hochschule Reutlingen, Alteburgstraße 150, 72762 Reutlingen, Germany
[2]IPTC and LISA+ center, University of Eberhard Karls Tübingen, Auf der Morgenstelle 18, 72076 Tübingen, Germany
[*]Correspondence: Marc.Brecht@Reutlingen-University.de

This is originally published in the Journal of Spectral Imaging (https://doi.org/10.1255/jsi.2020.a18) as

"M. Al Ktash, O. Hauler, E. Ostertag and M. Brecht, "Ultraviolet-visible/ near infrared spectroscopy and hyperspectral imaging to study the different types of raw cotton", J. Spectral Imaging 9, a18 (2020). https://doi.org/10.1255/jsi.2020.a18 © 2020"

5.1 Abstract

Different types of raw cotton were investigated by a commercial ultraviolet-visible/near-infrared (UV-Vis/NIR) spectrometer (210 nm - 2200 nm) as well as on a home-built setup for NIR hyperspectral imaging (NIR-hyperspectral imaging) in the range 1100 nm - 2200 nm. UV-Vis/NIR reflection spectroscopy illustrates a dominant role of proteins, hydrocarbons and hydroxyl groups. A similar result was revealed with NIR- hyperspectral imaging. Experimentally obtained data in combination with principle component analysis (PCA) provides a general differentiation of different cotton types. For UV-Vis/NIR spectroscopy, the first two principal components (PC) represent 82 % and 78 % of the total data variance for UV-Vis and NIR regions respectively. Whereas, for NIR- hyperspectral imaging due to the large amount of data acquired, two methodologies for data processing were applied in low and high lateral resolution. In the first method, the average of the spectra from one sample was calculated and in the second method the spectra of each pixel were used. Both methods are able to explain ≥ 90 % of total variance by the first two PCs. The results show that it is possible to distinguish between different cotton types based on a few selected wavelength ranges. The combination of hyperspectral imaging and multivariate data analysis has a strong potential in industrial applications due to its short acquisition time and low cost development. This study opens a novel possibility for a further development of this technique towards real large-scale processes.

5.2 Introduction

Hyperspectral imaging is an imaging technology that combines spatial information with spectroscopy. It is a fast and non-destructive method, which has evolved into a powerful analysis tool for product inspection. Thereby, spatial images with very detailed spectral information for each pixel of an object are collected simultaneously [21,24,26]. In the past, spectroscopical applications as well as hyperspectral imaging in the UV-Vis and NIR range are more frequently found in the textile research and in industrial applications [156]. In textile research, cotton plays a dominant role among textiles, since cotton is the most important naturally occurring raw material for the production of fabrics [157,158]. More than 34 million hectares of land are used to grow cotton, and around 100 million households worldwide are engaged in cotton production [159]. Cotton is considered as a key resource in the textile industry and accounts for about 30 % of all fibers used in this sector [160]. In recent years, the increase in quality and processing requirements has led to the introduction of modern techniques for processing and quality control [15,113-115]. Nevertheless, distinguishing between different cotton species is still a demanding task.

Several detection methods have been developed and applied to identify and classify different cotton varieties [158,161]. Most of them are off-line techniques such as thermogravimetric analysis and optical spectroscopy [162-164]. Only little information is expected in the visible range, since most raw cotton and residuals are reflective (or transparent) [165,166]. Valuable information can be expected in the NIR region from characteristic molecular vibration, e.g. CH_n and OH groups of cotton which are omnipresent [167]. Unfortunately, the overall sensitivity for small variations of the sample as well as for small amounts of contaminations in the NIR range is low and they are hard to detect [36,168]. Therefore, numerous studies in the NIR region used a combination of spectroscopy and chemometric modeling [167,169-175].

With NIR- hyperspectral imaging system, a complete optical spectrum with innumerable spectra are collected at all image pixels. This is in contrast to multispectral systems such as RGB cameras where only a limited number of wavebands are collected [86,176].

Most of the hyperspectral imaging applications were focused on remote sensing systems such as satellites or aircrafts to gather information for agricultural, geological inspections and military purposes. Nowadays, hyperspectral imaging is evolving into a standard for in-line and on-line inspection in process analytics and quality control. Prominent technical applications can be found in quality control for medicine, food and agricultural products [86,177,178].

In industrial applications, a hyperspectral imaging system is based on a combination of a pushbroom scanner and a conveyor belt. Pushbroom scanner is fixed over the conveyor belt as shown in Figure 5.1. Such inspection systems require a minimum of sample preparation and are able to scan several samples swiftly with high spectral resolution[179]. Here, the pushbroom scanner captures the complete spectral information line by line. The data is collected with the camera placed perpendicular to the conveyor belt. As the conveyor belt moves, images are continuously captured by the pushbroom scanner, resulting in a three dimensional (3D) data matrix with dimensions x,y and λ and is often referred to as hypercube [28].

For cotton research, hyperspectral imaging was used in the UV-Vis range to detect foreign matter with differentiation and classification of lint in cotton samples [161]. The results showed great potential using a hyperspectral imaging system for the classification of foreign matter [15,115,156].

In this study, we used optical reflection spectroscopy in the UV-Vis/NIR range as well as hyperspectral imaging in the NIR range for the differentiation of cotton sample sets. For both methods,

a chemometric model was developed that is based on PCA. Using this model, we were able to distinguish between the different cotton types of our sample sets.

5.3 Materials and Methods

5.3.1 UV-Vis/NIR spectroscopy

Reflectance spectra of the samples were recorded in the range from 210 nm to 2200 nm using a UV-Vis/NIR spectrometer (Lambda 1050, Perkin Elmer Ltd). It was used to compare the data from the NIR- hyperspectral imaging and validate to another device. The UV-Vis/NIR spectrometer was equipped with an Ulbricht sphere covered by polytetrafluorethylen (PTFE) to acquire data in diffusion reflection mode with two detectors: one is an indium gallium arsenide (InGaAs) detector and the second one is photomultiplier inside the sphere. The samples were placed on this rear of the sphere, and a diffused scattering PTFE as a white reference disc was placed behind the sample. The complete measuring aperture area is approximately 4.9 cm^2. From every cotton sample disc, a spectrum was acquired on each side. In total, three discs were measured for each sample and, thus, for each cotton sample disc, six spectra were recorded.

5.3.2 NIR hyperspectral pushbroom online imaging system

Figure 5.1a shows the setup of the used hyperspectral imaging system. The hyperspectral system is based on a pushbroom imager connected to a Xencis, Xeva 2.5 – 320 camera equipped with a mercury cadmium telluride (HgCdTe) detector of 8 nm of spectral resolution, and has a 30 µm slit width. Two halogen lamps illuminate the sample area. PTFE is used as a white reference while the dark reference is acquired by imaging without any light exposure to the sensor. Figure 5.1 (b-d) illustrates the principle and workflow for hyperspectral imaging. Figure 5.1b shows complete spectroscopic information acquired for each line. Thus, a continuous line by line collection of spectral information forms a two dimensional (2D) image as shown Figure 5.1c. It is also possible to extract a single spectrum from a given pixel or point in the 2D image as shown in Figure 5.1d.

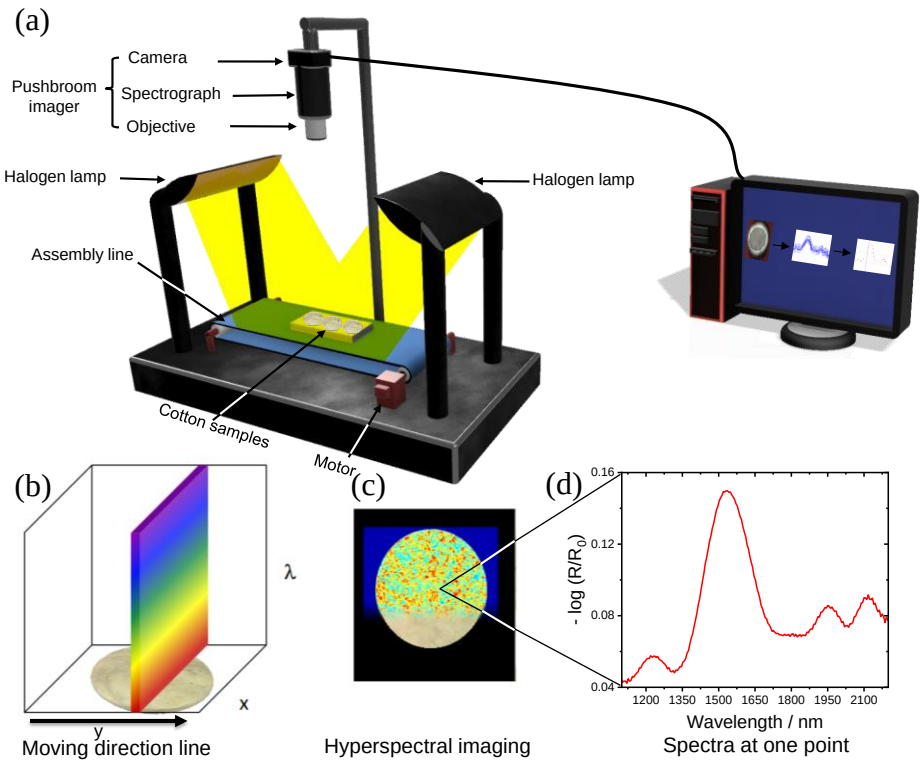


Figure 5.1: (a) Setup of a hyperspectral imaging system based on the pushbroom concept. (b) hyperspectral imaging scanning principle. (c) hyperspectral imaging generated immediately from the scanning of a cotton sample disc. (d) NIR-Spectrum for one single pixel extracted from the image.

5.3.3 Samples

Figure 5.2 show 5 types of raw cotton and one hemp sample which were investigated. The samples are organic raw material cotton (RoB), hemp plant from China (HC), recycled cotton (RcO), standard raw material cotton (RoSt), recycled organic bright cotton (RcBH) and mechanically cleaned cotton sample (CLN). Three samples of the aforementioned cotton types were collected from the bulk, amounting to 0.75 g form each sample. The samples were pressed at 10 tons for 2 minutes to have same physical properties by a hydraulic press into a disc shape. The hydraulic press was cleaned after each sample to reduce the chance of any impurities.

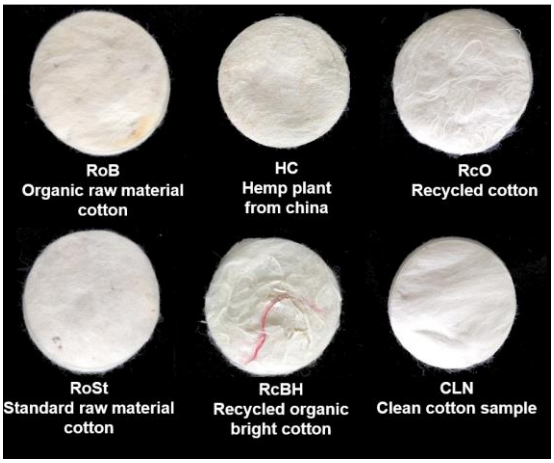
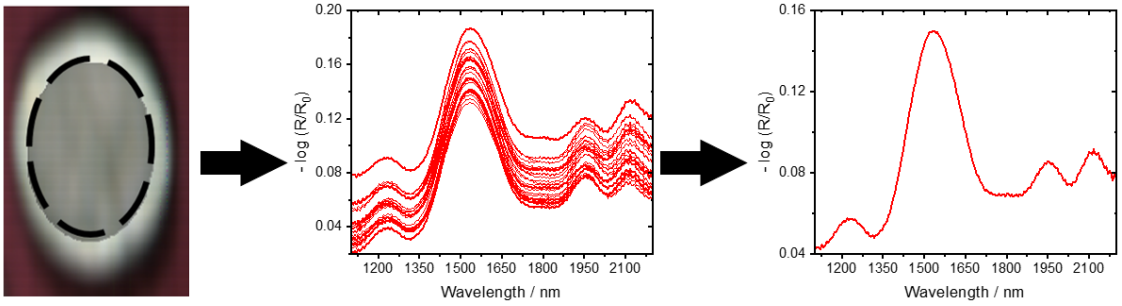


Figure 5.2: Raw cotton sample discs.

5.3.4 Data collection and preprocessing of hyperspectral data

The following two methods for data pre-processing are described resulting in low and high lateral resolution. Matlab (MATLAB 9.2.0, Mathworks, MA, USA) scripts were written for per-processing of the hyperspectral data cube.

Figure 5.3 shows the workflow for calculating the mean spectrum of each sample. The hyperspectral image is collected by moving any cotton sample disc at a constant speed, approximately 50 spectra were collected manually within the indicated area of interest, as shown in Figure 5.3a (marked as dash line) and plotted as shown in Figure 5.3b. The average of these spectra is calculated and shown in Figure 5.3c.



(a) Hyperspectral imaging (b) Extracted spectra (c) Average Spectra

Figure 5.3: (a) hyperspectral imaging of a cotton sample disc with area of interest (dash line) with a diameter of 2.5 cm. (b) Spectra extracted from the selected area. (c) Average spectrum of all spectra shown in (b).

Figure 5.4 shows the workflow for the second preprocessing method. The hyperspectral image is captured by moving the 18 cotton sample discs at a constant speed. To differentiate

between signal and background, a distinction is first made between the respective spectral characteristics.

For this purpose, two parallel planes are fitted into each spectral channel, one for the background and one for the samples. The distance between these planes is then selected as the parameter for the spectral difference between the sample and the background. The color channel with the highest value is used as mask for all other color channels. Half of this difference is set as threshold value. All lateral points of the color channel whose intensity value is above this threshold value are classified as background and removed. This clipping mask is applied to the entire hyperspectral data cube. The remaining data corresponds to the spectral contributions from the samples. These are converted from the 3D hyperspectral data set into a 2D format by joining the lateral points of the X and Y dimensions. This creates a matrix in which each row corresponds to a pixel with a complete spectrum. This matrix is used as input for the PCA.

Figure 5.4a shows the image obtained through the hyperspectral camera. The color channel with the highest differential value is displayed in Figure 5.4b. Figure 5.4c shows a single color channel of the cotton sample discs' hyperspectral data cube after removing all lateral components associated with the background, the removal of outliers like dead pixels or cosmic events, and the application of a PCA filter, which removes all contributions of higher PCA components. The PCA filter works as follows: the first three of the resulting PCs explain about 88 % of the variance. The 4^{th} and higher components, while contributing less than 7 % to the overall variance, contain mainly noise and were therefore discarded for further analysis. The remaining 5% of the total variance is found within the residuals, and do not contribute significant information. Figure 5.4d shows an image, where the red-green-blue (RGB) value corresponds to scores of the first (R), second (G) and third (B) components.

In the next step, all score values that are 90 % similar to another score in all the main components considered are removed from the data set. From the remaining score values a reduced data set with the load values of the considered main components was generated. The reduced data set is then converted back into a 3D hyperspectral data cube by separating the combined lateral information. Figure 5.4c shows the reduced data as lateral information for one spectral channel. The principal component analysis of this data again shows a significant grouping of the different types of cotton. In the end, approximately 120,000 spectra remain from initially obtained 1.7 Million spectra.

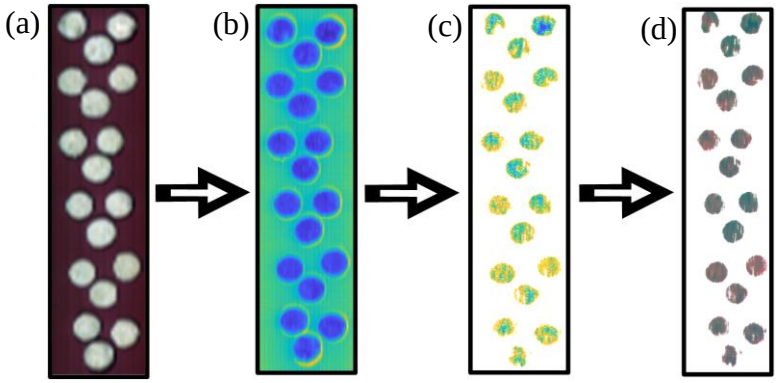


Figure 5.4: (a) Hyperspectral raw imaging of 18 cotton sample discs with a diameter of 3.1 cm. (b) Image of the color channel with the highest variance between cotton disks and background. (c) Images after subtraction of the background, removal of outliers, and application of filters. (d) Image of RGB value corresponds to scores of the first (R), second (G) and third (B) components.

5.3.5 Data handling and software

The UV-Vis/NIR spectra are recorded with the Lambda 1050 UV WinLab software from PerkinElmer. The NIR hyperspectral pushbroom images are analyzed by the Prediktera software from Evince 2.7.9. PLS Toolbox 8.5.1 (Eigenvector Research, Inc., USA) which is used for the data processing and analysis. Lighting conditions may vary between the samples and even within the samples across the scan line. A regular way to calculate this effect is to convert measured raw spectra to reflectance spectra by the following formula [35,86]:

$$\text{Reflectance} = -\log R/R_0 = \frac{I_{sample} - I_{dark}}{I_{reference} - I_{dark}} \quad (5.1)$$

Where R and R_0 represent the transmitted and incident intensity. I_{sample} is the intensity of the original image data, I_{dark} is the intensity of the dark current image data and $I_{reference}$ is the intensity of the white reflectance image. Pre-processing of the mean center, smoothing (Sav-Gol) with filter width 15 and polynomial order one, and generalized least squares (GLS) are applied to the data. GLS is used to achieve an efficiency by transforming variance covariance matrix into a homoscedastic one[36]. It works as a filter that calculate the differences between the samples. The differences are considered as interference or clutter and GLS aims to reduce these interferences [180-182].

5.4 Results and discussion

5.4.1 UV-Vis/NIR spectroscopy

Figure 5.5a shows UV-Vis/NIR spectra (210 nm – 2200 nm) from all samples. Six spectra were recorded for each cotton sample type, three on each side. As expected, the spectra show a high similarity. All spectra show the strongest reflectance at 280 nm which can be attributed to proteins on the samples, see Table 5.1 [165]. In the visible range from 400 nm – 750 nm, the spectra do not show any distinct features since most of the raw cotton is reflective. In the NIR region, several spectral features can be observed. Dominant contributions are found at 1500 nm, 1933 nm and 2100 nm corresponding to the functional groups CH, ROH and OH, respectively.

Due to the high similarity of the spectra, a differentiation of the samples is demanding. As a consequence, PCA is used to further differentiate the samples and was applied for processed spectra.

The processing of spectra is described in the Materials and Methods section. Figure 5.5b shows the scores plot of the first two principal components PC1 and PC2 for UV-Vis region (210 nm – 1100 nm). The PCA model explains 70.1 % and 82.3 % of the spectral information with the first two PCs respectively. The scores plot shows that PC1 and PC2 are sufficient to separate all samples. In this representation, the hemp (HC) sample shows the most distinct separation from the cotton group, as expected. Figure 5.5c shows the corresponding loadings plot for PC1 and PC2. The most significant differences between those loadings are found in the regions from 210 nm – 350 nm, 450 nm – 700 nm. In the UV range (210 nm – 350 nm), the strongest influence on PC1 is found at 280 nm, 300 nm and for PC2 at 290 nm. They can be assigned to proteins and amino acids (see Table 5.1)[165]. The contributions in the visible range (450 nm – 700 nm) show a maximum/minimum at 680 nm, it can be assigned to the color of the RcO samples (see also the inset in Figure 5.5a).

Figure 5.5d shows the scores plot of the first two principal components PC1 and PC2 for NIR region (1100nm – 2200 nm). The PCA model explains 63.5 % and 78.0 % of the spectral information with the first two PCs respectively. The scores plot shows that the first two PCs are sufficient to separate all samples clearly from one another in NIR range. In the scores plot, the hemp (HC) and CLN sample shows the most distinct separation from the cotton group.

Figure 5.5e shows the loadings plot for PC1 and PC2. The most significant differences between those loadings are found in the regions from 1100 nm – 1200 nm, 1350 nm – 1500 nm, 1600 nm – 1700 nm and 1850 nm – 2100 nm. In the NIR region (1100 nm – 2200 nm), several spectral features are variable which are assigned to the hydrocarbons and hydroxides oscillation (see Table 5.1).

With UV-Vis/NIR, a separation of the analyzed cotton sample discs has been successfully demonstrated. However, the large deviations between PC1 and PC2 are mainly found in the UV-Vis and NIR region. Therefore, the application of an online method for characterization is the most suitable for these spectral regions.

Table 5.1: UV-Vis/NIR reflectance maxima [165],[85,183].

Reflectance (nm)	Functional groups
1240 nm	CH
1525 nm	ROH
1790 nm	CH_3, CH_2
1955 nm	OH
2117 nm	ROH
2342 nm	CH, CH_2, CH_3

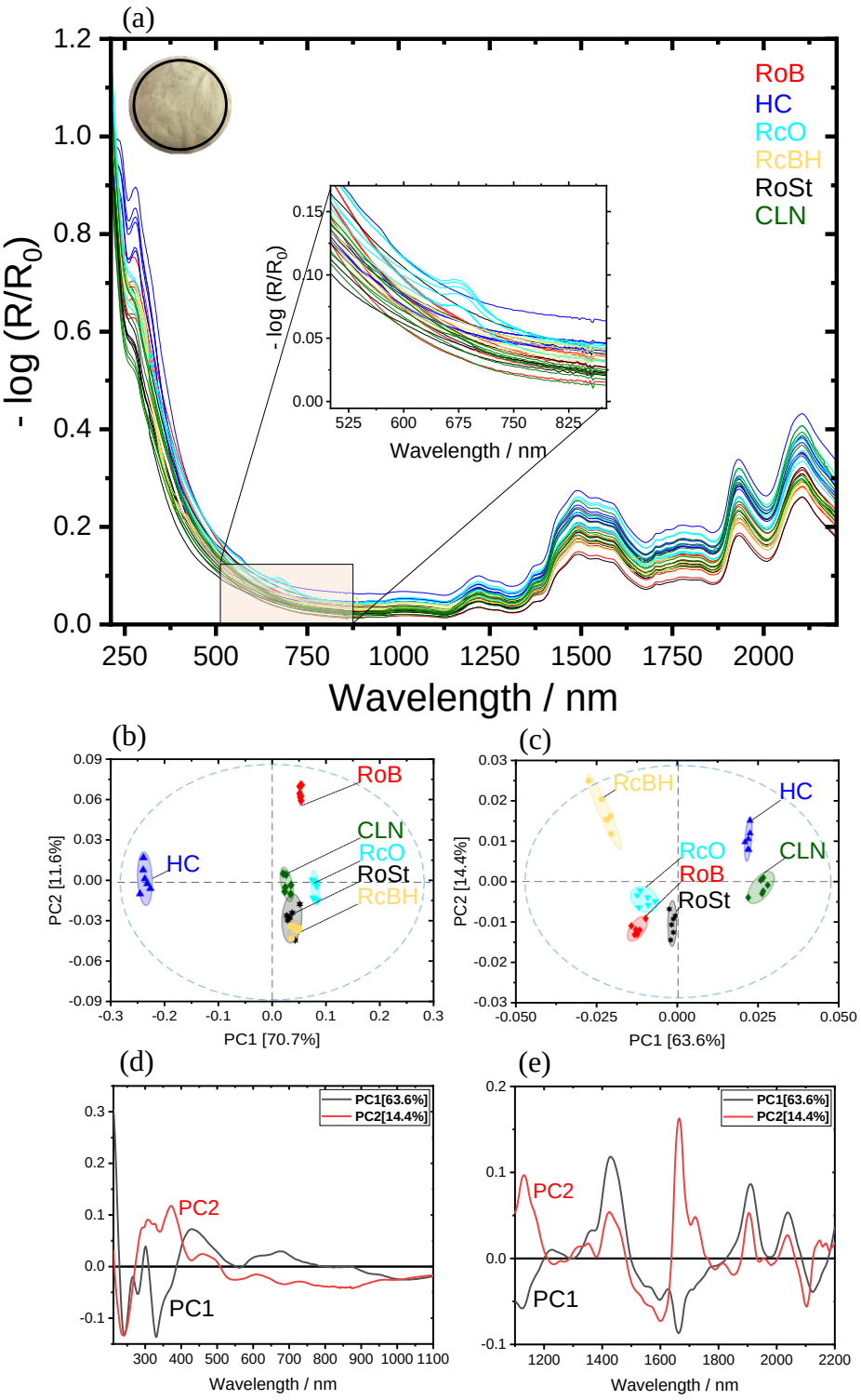


Figure 5.5: (a) UV-Vis/NIR spectra of cotton sample discs including one HC sample in the wavelength range 210 nm – 2200 nm. Upper left: Image of a cotton sample disc where the region of integration for determining the average spectra is indicated by a black area with a diameter of 2.5 cm. (b) Scores plot for the processed spectra in the UV-Vis. The 2D projection of the 95 % confidence ellipse of the data collected from each type of cotton is included to facilitate visualization of the obtained results. (c) Loadings plot for the UV-Vis. (d) Scores plot for the NIR. (e) Loadings plot for the NIR.

5.4.1 NIR Hyperspectral Imaging

Two data processing techniques were applied to the NIR hyperspectral images to calculate PCA models. As before, three samples of each raw fibers were analyzed. The setup for hyperspectral imaging as well as for determination of the spectra from the hyperspectral data matrix is described in the Materials and Methods section.

In the first method, the mean value of the spectra was calculated for each cotton sample disc. A total of six spectra are determined from the hyperspectral imaging data for each cotton sample type.

Figure 5.6a shows hyperspectral NIR spectra in the range from 1100 nm – 2200 nm. The most dominant contributions are observed around 1525 nm, they can be attributed to the presence of OH groups. Four weaker peaks are observed around 1340 nm, 1790 nm, 1955 nm and 2117 nm, their assignment is given in Table 5.2[36].

The PCA of these spectra explains 93.7 %, 97.0 % or 98.3 % of the spectral information with the first two, three or four PCs respectively. Figure 5.6b shows the results for the first three PCs. In the scores plot it can be seen that the first three PCs are sufficient to separate all samples clearly from one another.

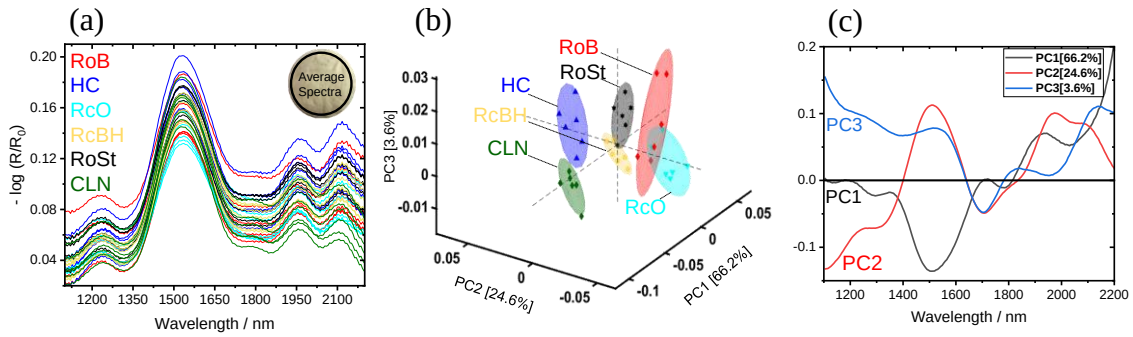


Figure 5.6: (a) Spectra recorded by hyperspectral imaging of cotton sample discs including one HC sample in the NIR range from 1100nm - 2200nm. Upper right: Image of a cotton sample disc where the region of integration for determining the average spectra for each sample is indicated by a black circle with a diameter of 2.5 cm. (b) Scores plot for the processed spectra in NIR- hyperspectral imaging. The 2D projection of the 70 % confidence ellipse of the data collected from each type of cotton is included to facilitate visualization of the obtained results. (c) Loadings plot for the NIR- hyperspectral imaging.

Figure 5.6c shows the loadings plot for the first three PC's. In the range from 1340 nm – 1663 nm the reflectance around 1508 nm can be assigned to the presence of ROH (see Table 5.2).

The reflectance in the range from 1789 nm - 2100 nm, 1973 nm can be assigned to the OH group. The contribution approximately 2270 nm is due to CH [85,183].

Table 5.2: NIR Hyperspectral imaging reflectance maxima [85,183].

Reflectance (nm)	Functional groups
1240 nm	CH
1525 nm	ROH
1790 nm	CH3, CH2
1955 nm	OH
2117 nm	ROH
2342 nm	CH, CH2, CH3

In the second method, several thousand spectra from every cotton sample disc were used to calculate the PCA model. The preprocessing and workflow of the spectra from the hyperspectral data matrix is described in the section Materials and Methods. Figure 5.7a presents examples of hyperspectral NIR spectra from a single pixel of each of the six cotton types in the range 1100 nm – 2200 nm.

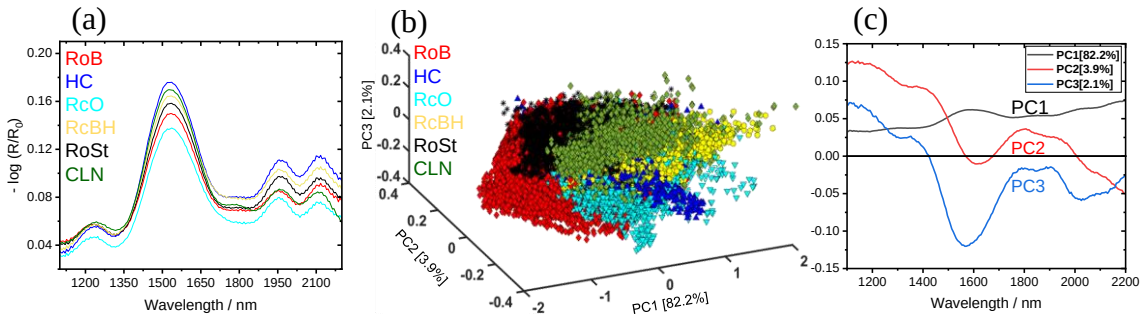


Figure 5.7: Hyperspectral of cotton sample discs including one HC sample in the NIR range from 1100nm - 2200nm. (a) Six example spectra recorded at individual pixels. (b) Scores plot calculated for the whole data set including several thousand processed spectra. (c) Loadings plot for the NIR-Hyperspectral imaging.

The PCA of these spectra explains 86.0 %, 88.2 % or 89.0 % of the spectral information with the first two, three or four PCs respectively. Figure 5.7b shows the results for the first three PCs, the first three PCs are sufficient to separate all samples from one another. A clear separation is observed for RoB and CLN, while the HC, RoSt, RcO and RcBH are slightly overlapping. Nevertheless, these samples can be separated only if one pair e.g. HC and RoSt is included in a separate model (data not shown). Figure 5.7c shows the loadings plot of the first

three PCs. Overall, the loadings are comparable with the loadings shown in Figure 5.6c, except a change of the sign. In the range from 1350 nm – 1700 nm, the reflectance around 1550 nm can be assigned to the presence of ROH (see Table 5.2). The reflectance in the range from 1800 nm - 1990 nm can be assigned to OH groups. The signal around 2302 nm is due to CH [85,183].

The first three PCs explain a significant amount of the NIR hyperspectral data for both preprocessing methods. Calculating the PCA model at each pixel or deriving it from the mean spectra does not significantly change the data behavior of the model (Figure 5.6 and Figure 5.7). The advantage of using average spectra instead of the complete data set is fast data processing. However, this method is limited to recognize or spectrally separate background from the samples automatically. Therefore, a certain time is required to select the samples manually and calculate the average spectra for each cotton sample disc. On the other hand, when applying a filter (see Figure 5.4) the separation of the sample from the background works automatically, but here the quantity of data hampers a fast processing. The scattering in the scores plot in Figure 5.7 shows the huge variability of the properties of the samples, theses only become visible if the spectra are taken with hyperspectral imaging. Compared with the scattering where the spectral information is averaged over a larger area (Figure 5.5 and Figure 5.6) this is remarkably reduced. The large variability of the score values from the hyperspectral imaging indicates a change of the samples properties on the scale of the resolution actually achieved. For the hyperspectral imaging setup this is about 13µm. The high lateral resolution achieved here shows that sample properties on this scale vary and are therefore relevant, as new insights into the heterogeneity of fiber samples can be gained. As a consequence, the data show the high potential for hyperspectral imaging which is beyond the differentiation of fiber types.

In the next step, a filter is required that combines the advantages of both methods to speed up the data handling. Together with this, a simplified model can be developed that meets the requirements of real online applications.

5.5 Conclusions

UV-Vis/NIR reflection spectroscopy and hyperspectral imaging in combination with PCA is a promising approach for the detection and differentiation of raw cotton types. The most relevant information for the differentiation of cotton types was found in both the UV and NIR range (see Figure 5.5c).

The results obtained with UV-Vis/NIR spectroscopy revealed that the contribution in the UV can be assigned to the presence of protein at 280 nm. The most dominant contribution to absorbance in the NIR range can be assigned to CH_3 for the most prominent band at 1775 nm and to ROH vibrations at 1500 nm. The spectral data were analyzed with PCA in order to achieve a differentiation of different cotton types. The PCA model was able to classify all types with the first two PCs explaining the maximum variance of the data.

NIR- hyperspectral imaging results reveal the most dominant absorbance assigned to CH_3 and ROH at 2270 nm and 1525 nm respectively. Two methods were used for processing the large amount of data. Both approaches resulted in a differentiation of all types. The advantages of the rugged online home-built setup is a high spatial/spectral resolution and a rapid data acquisition. With this method, several samples can be measured in a short time and at low cost.

Based on the data shown it is reasonable to develop a simplified chemometric model, which meets the requirements of a real process with industrial standards and precision.

6 Paper IV: Prediction of Honeydew Contaminations on Cotton Samples by In-Line UV Hyperspectral Imaging

Mohammad Al Ktash[1,2,†], Mona Stefanakis[1,2,†], Frank Wackenhut[1], Volker Jehle[3], Edwin Ostertag[1], Karsten Rebner[1] and Marc Brecht[1,2,]*

[1]Center of Process Analysis and Technology (PA&T), School of Life Sciences, Reutlingen University, Alteburgstraße 150, 72762 Reutlingen, Germany
[2]Institute of Physical and Theoretical Chemistry, Eberhard Karls University Tübingen, Auf der Morgenstelle 18, 72076 Tübingen, Germany
[3]Texoversum Faculty Textile, Reutlingen University, Alteburgstraße 150, 72762 Reutlingen, Germany
[*] Correspondence: Marc.Brecht@Reutlingen-University.de.
[†]These authors contributed equally to the work.

This is originally published in sensors (https://doi.org/10.3390/s23010319) as

"Al Ktash, M.; Stefanakis, M.; Wackenhut, F.; Jehle, V.; Ostertag, E.; Rebner, K.; Brecht, M. Prediction of Honeydew Contaminations on Cotton Samples by In-Line UV Hyperspectral Imaging. Sensors 2023, 23, 319. https://doi.org/10.3390/s23010319"

6.1 Abstract

UV hyperspectral imaging (225 nm – 410 nm) was used to identify and quantify the honeydew content on real cotton samples. This study presents the implementation and application of UV hyperspectral imaging as a non-destructive, high-resolution, and fast imaging modality. For this novel approach a reference sample set, which consists of sugar and protein solutions that were adapted to honeydew, was set up. In total, 21 samples with different amount of added sugars/porteins were measured to calculate multivariate models to predict and classify the amount of sugar and honeydew at each pixel of a hyperspectral image. The principal component analysis models (PCA) enabled a general differentiation between different concentrations for sugar and honeydew, respectively. A partial least squares regression (PLS-R) model was built based on the cotton samples soaked in different sugar and protein concentrations. The result shows a reliable performance with R^2_{cv} = 0.84 and low RMSECV = 0.009 g for the validation. The PLS-R reference model was able to predict the honeydew content laterally resolved in gram on real cotton samples for each pixel with light, strong and very strong honeydew contaminations. Therefore, in-line UV hyperspectral imaging combined with chemometric models can be a future effective tool for the quality control of industrial processing of cotton fibers.

6.2 Introduction

Hyperspectral imaging is an imaging technology that combines video image analysis with spectroscopy [141,184]. Precisely, it is a series of images acquired by moving the object or the imager. It is a fast and non-destructive technique, which has developed into a robust analysis tool for product screening. Such systems are able to capture spectral and spatial information with high resolution. As a result, a spectrum of an object can be obtained for each hyperspectral image pixel simultaneously [20,185].

Hyperspectral imaging and spectroscopic applications are widely used in industrial environments [19,53,186-189]. The importance of hyperspectral imaging is steadily growing in the textile industry. For example, the visible (Vis) and near infrared (NIR) range are often applied for quality control and sorting processes [190,191]. Where the UV range is rarely used so far.

In textile research, cotton is considered as one of the most important natural fibers for fabric production [192,193]. It provides approximately 50 % of the world's textile fibers [194]. Cotton consists of approximately 95 % cellulose and 5 % sugar, wax, proteins, organic acids and pectin

[195]. The process ability is affected and degraded by the sugar content. Sugar is a naturally excretion of aphids and whiteflies on the cotton through metabolic processes and is specifically called honeydew [196]. This contamination on the raw cotton causes stickiness, which causes problems in the processing stage. This leads to economic loss because the sticky raw cotton is rejected during quality control [195,197]. In-line detection and subsequent removal of sticky cotton would lead to an uninterrupted production and thus higher profit [191,198,199].

Several detection methods have been developed and applied to identify cotton contaminations and stickiness in ultraviolet (UV)-Vis/NIR spectroscopic applications as well as Vis/NIR hyperspectral imaging [103,115,141,200]. Most of them are off-line such as optical spectroscopy [103,201]. Identification of cotton and cotton trash components was studied by Fortier et al. [202] using FT-NIR spectroscopy. Mustafic et al. [161] examined the applicability of hyperspectral imaging to detect and classify cotton foreign matter in the visible spectral region, whereas the Vis/NIR region was studied by Jiang et al. [203]. Other methods, such as thermogravimetric analysis [162], high-pressure liquid chromatography (HPLC) and minicard [204], require an elaborate sample preparation, are time-consuming and expensive compared to optical spectroscopy. In-line detection and quantification of stickiness on cotton samples using NIR hyperspectral images was investigated by Severino et al. [199]. They were able to discriminate between glucose on cellulose and melezitose, trahalose, glucose, fructose, and sucrose at each pixel.

Tschannerl et al. [86] compared hyperspectral imaging in UV and NIR regions to precisely discriminate between phenolic flavor concentrations in melted barley. The rarely used UV region showed interesting results despite the illumination wasn't optimal. Previously, our group reported a hyperspectral imaging setup for the UV spectral region, this setup was used to distinguishing between different pharmaceutical drugs [19] as well as, for characterizing oxide layers thickness and copper states on direct bonded copper [53]. The results clearly showed that a spectral imager based on pushbroom technology has many advantages in terms of achieving fairly short UV operational wavelengths and a high spectral resolution. However, hyperspectral imaging rapidly scans samples resulting in a large amount of spectral data within a short time period. Therefore, multivariate data analysis, such as principal component analysis (PCA) and partial squares regression (PLS-R), are required to reduce the amount of data without losing important information. PCA reveals the most relevant information of a data matrix [38]. A combination between PCA and quadratic discriminant analysis (QDA) enables data classification and investigation of model quality parameters [65,205]. The information about the relation between a number of predictor variables and independent variables can be extracted by using PLS-R [206].

The aim of this work is to develop a chemometric model able to identify and quantify the amount of honeydew on real cotton samples based on UV hyperspectral imaging. For this approach a reference sample set, which consists of honeydew typical sugars and proteins, was prepared. Mechanically cleaned cotton was soaked with solutions with different sugar concentrations. Chemometric models, especially PCA and PLS-R, were developed based on UV hyperspectral imaging. PCA is used to classify the cotton samples according to their sugar concentration and PLS-R is applied to correlate the UV spectra with the sugar concentration. This PLS-R model successfully predicts the amount of honeydew in gram on real cotton samples. This work is considered as the first scientific work to identify and quantify the amount of honeydew contents at each pixel. Therefore, hyperspectral imaging is a suitable technique for in-line environment applications in a rapid and non-destructive manner.

6.3 Materials and Methods

6.3.1 Chemicals and preparation of solutions

0.2 g of each macronutrient 1-6 was weighted and dissolved in 10 mL of deionized water (see Table 6.1). A six-fold serial dilution was prepared in 50 mL volumatric flasks. For each diluting step 25 mL of the previous solution and 25 mL of deionized water were mixed for 2 min (see Table 6.2).

Table 6.1: Description of the macronutrients and natural materials.

Macronutrients and natural materials	Samples	Description	Manufacture	CAS Number
1	Glucose	D-Glucose anhydrous Laboratory reagent grade	Fisher Scientific GmbH, Leics, UK	50-99-7
2	Fructose	D-Fructose, 99.0 %	ThermoFisher GmbH, Kandel, Germany	57-48-7
3	Sucrose	D-Sucrose, >=99.9 % For Molecular Biology	Fisher Scientific GmbH, New Jersey, USA	57-50-1
4	Melezitose	D-(+)-Melezitose monohydrate, >=99.0 %	Sigma-Aldric Chemie GmbH, Steinheim, Germany	10030-67-8
5	Trehalose	D- Trehalose anhydrous, 99.0 %	Acros Organics, New Jersy, USA	99-20-7
6	Protein	Bovine Serum Albumin (BSA) fraction V, lyophilized powder	PAN-Biotech GmbH, Aidenbach, Germany	9048-46-8

6.3.2 Sample set and sample preparation

The sample set consists of cleaned cotton to build the model and cotton samples contaminated by honeydew to test the model.

The cotton samples for model building were collected from a bulk of cotton mechanically cleaned [207,208] from the Texoversum Faculty Textile at Reutlingen University. The cotton is a blend of different long staple Pima qualities. In total, 21 cotton samples were prepared with a weight of 0.3 g ± 0.0001 g (XSE205 DualRange, Mettler Toledo GmbH, Switzerland). The samples were dried in a vacuum oven (Vacutherm VT 6130 M, Termo Fisher Scientific Inc., Waltham, Massachusetts, USA) at 30 °C and 50 mbar for 8 h to remove absorbed humidity. The humidity was estimated by a commercially available sensor (Humidity-Detector MD, H. Brennenstuhl GmbH & Co. KG, Tübingen, Germany). The weight loss is documented in Appendix Table 10.2. 4 mL macronutrients solution were used for each sample. Three samples per concentration were made (see Table 6.2). The samples were soaked in an aluminum plate (28 mL, Carl Roth GmbH+Co. KG, Karlsruhe, Germany). The samples were dried again in a vacuum oven at 30°C and 50 mbar for 44 h. By determining the remaining weight the average macronutrient content can be calculated for each sample (see Appendix Table 10.2). The humidity and temperature of the labatory were monitored (BL30, Klima-Datenlogger, Trotec GmbH, Heinsberg, Germany) during the whole workflow.

Comparable real cotton samples were collected by ICA Bremen GmbH (Bremen, Germany) to test the predictive power of the model. The samples were chosen according to their honeydew content in the steps light, strong and very strong. The samples origin is Sudan Acala (see Table 6.3).

Figure 6.1 shows the samples pressed in the sample holder prepared for measuring. The sample types are named from A to F and one mechanically cleaned (CLN) sample, where A has the highest concentration (2 wt%), and F has the lowest concentration (0.0625 wt%) (see Table 6.2). The average sugar content remaining on the samples after 44 h was calculated (see Table 6.2). For ease of reading, we omit the term macronutrients for the description of the solution of various sugars and the protein in the following, and replace it with the short-term "sugar" for the sample nomenclature.

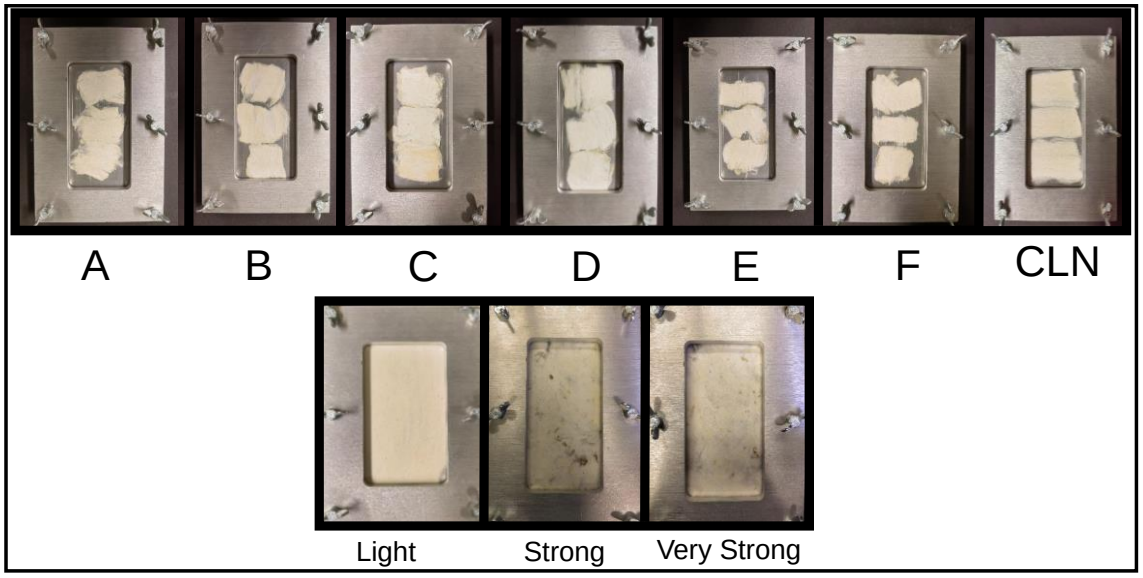


Figure 6.1: Overview of the samples pressed in the sample holder. For each concentration, three samples were prepared and measured at once (A to F and CLN). Real cotton samples with different honeydew contents (light, strong and very strong.

Table 6.2: The concentration of the sugar solutions and the weighted averaged sugar amount on cotton samples.

Sample type	Sugar concentration / wt %	Ratio of: sugar / g dried cotton / g
A	2	0.2593
B	1	0.1331
C	0.5	0.0743
D	0.25	0.0386
E	0.125	0.0326
F	0.0625	0.0322
CLN	-	-

Table 6.3: The number of honeydew stickiness points on cotton samples.

Stickiness Type	Single measurments	Average number of sticky points	Sample
Light	2, 11, 5	6	4301
Strong	47, 45, 47	46	Sudan Girba Acala 3SG
Very strong	60, 69, 80	70	Sudan Gezira Acala type 3SG

6.3.3 UV hyperspectral imaging setup

Figure 6.2a shows a scheme of the hyperspectral imaging setup. The pushbroom imager consists of a back-illuminated CCD camera (Apogee Alta F47: Compact, inno-spec GmbH, Nürnberg, Germany) and a spectrograph (RS 50-1938, inno-spec GmbH, Nürnberg, Germany) with a slit width of 30 µm. The CCD camera has a resolution of 1024 x 1024 pixel (spatial x spectral) and a pixel size of 13 µm x 13 µm. The optimal integration time was 300 ms. The conveyor belt (700 mm × 215 mm × 60 mm, Dobot Magician, Shenzhen Yuejiang Technology Co., Ltd., Shenzhen, China) moves with a constant speed of 0.15 mm/s. The conveyor belt was located totally in a polytetrafluoroethylene (PTFE) (Sphereoptics GmbH, Herrsching, Germany) tunnel. The illumination was achieved by two Xenon lamps (XBO, 14 V, 75 W, OSRAM, München, Germany). Figure 6.2d shows a sample holder developed to reduce the influence of the topography of the samples. Therefore, a quartz glass made of suprasil 2 grade B with the dimension of 140 mm x 80 mm x 1 mm (Aachener Quarzglas-Technologie Heinrich GmbH & Co.KG, Aachen, Germany) was used. PTFE was used as white reference.

Figure 6.2(c-e) illustrates the principle and workflow of the data acquisition. Figure 6.2c presents the principle of hyperspectral imaging line scanning method which collects one line at a time, with all of the pixels in a line being measured simultaneously. The continuous line by line collection of spectral information results in a lateral (x, y) 2D image as shown in Figure 6.2d, whereas each location contains a further spectroscopic dimension (λ) as shown in Figure 6.2e. Thus, a 3D data matrix (hypercube) was recorded.

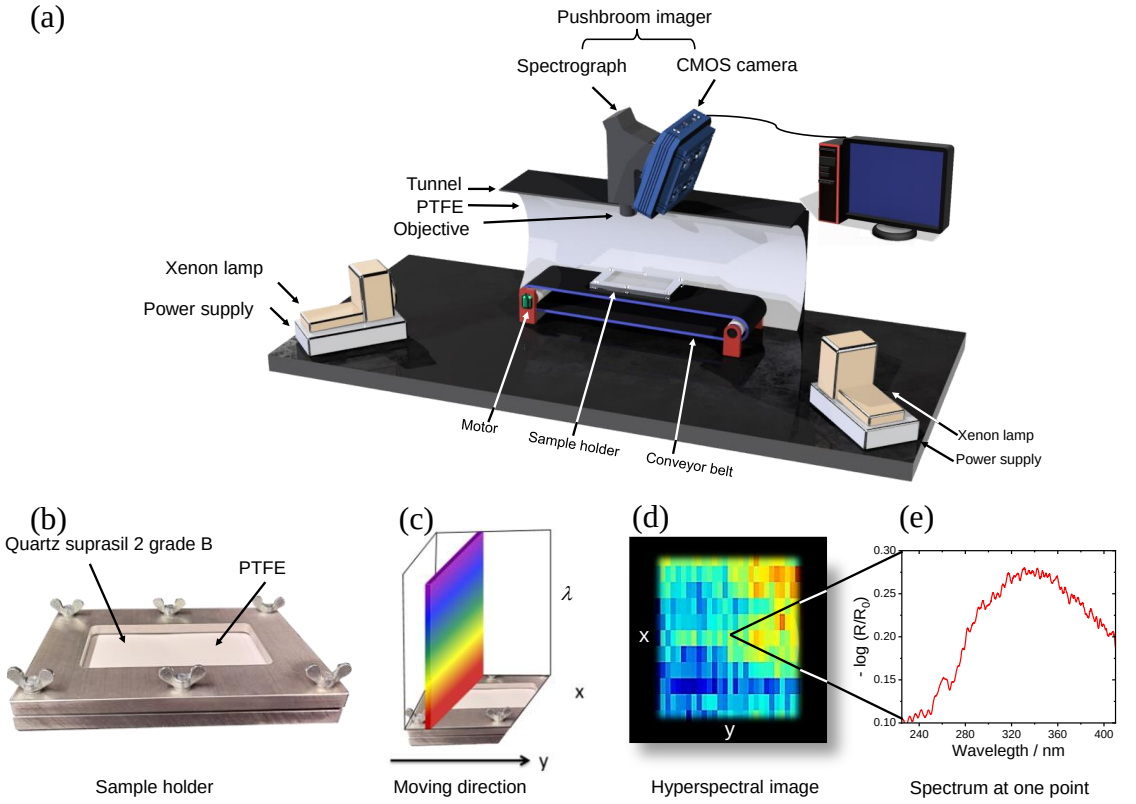


Figure 6.2: (a) Setup of a hyperspectral imaging system based on the pushbroom concept (the tunnel in the scheme was cut to show the inside). (b) Custom made sample holder consisting of quartz glass as sample cover and PTFE as reference. (c) Pushbroom imager scanning principle. (d) Hyperspectral image generated immediately from the scanning of a sample. (e) UV spectrum after preprocessing for one point extracted from the image given in (d).

6.3.4 Data collection and preprocessing

The UV hyperspectral imaging data were acquired by the SI-Cap-GB version V3.3.x.0 software (inno-spec GmbH, Nürnberg, Germany). The reflectance was calculated by the SI-Cap-GB automatically after recording $I_{reference}$ and I_{dark}. Illumination conditions will vary between samples and even within samples across the scan line, especially for heterogeneous samples such as cotton with high scattering due to sample topography. A common method to reduce the influence of the sample topography is to convert the raw spectra of each pixel into reflectance spectra (radiometric calibration) using the following formula [19,21,37-39]:

$$\text{Reflectance} = -\log R/R_0 = \frac{I_{sample} - I_{dark}}{I_{reference} - I_{dark}} \quad (6.1)$$

R and R_0 are the intensities reflected from the sample and a specific reference material with high reflectivity, in this case PTFE. The intensity of the original image is represented in I_{sample}. Accordingly, the intensity of the dark current image is given by I_{dark} and the intensity of the PTFE image is $I_{reference}$ [20]. In order to enhance the absorption bands the negative decadic logarithm is calculated as -log (R/R_0).

Hyperspectral data matrices were analyzed by Evince version 2.7.13 (Prediktera AB, Tvistevägen, Sweden). It is used for data handling and extracting the spectra of each pixel. For model building the sample set contains of 4896 pixel × 1024 pixel which represents approximately 5.0 million spectra. One spectrum ranges from 225 nm to 410 nm including 1024 variables. With 4464 pixel × 1024 pixel which represents approximately 4.6 million spectra the model was tested. In total, this results in approximately 15 GB of data size. Figure 6.3a shows an example of RGB hyperspectral images of cotton samples sprayed with different concentrations of sugar solution. A region of interest was selected by using a rectangular shape to extract the spectra (see Figure 6.3b).

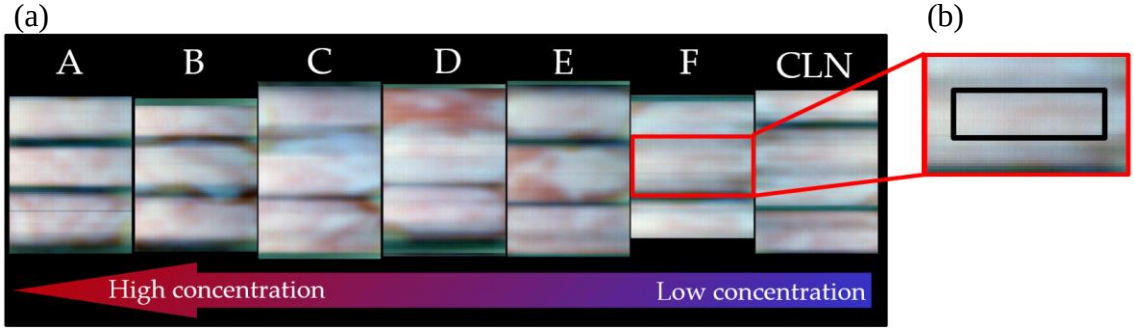


Figure 6.3: Example of data extraction. (a) Hyperspectral raw images of 18 cotton samples sprayed with different concentrations of sugar (A highest to F lowest) and one cleaned cotton sample (CLN). For model building, all spectra were extracted manually. (b) Zoom-in-image of a cotton sample with the region of interest marked by a black rectangle.

6.3.5 Multivariate data analysis and model building

Multivariate data analysis (MVA) was performed with "Aspen UnscramblerTM, version 10.5.1" (Aspen Technology Inc., Bedford, MA, USA). The UV spectra were pretreated prior to the multivariate data analysis in the following way: Base line correction followed by a Savitzky-Golay smoothing (8 points, symmetric, 2nd polynomial order). The principal component analysis (PCA) models were calculated with mean centering, cross-validation, and NIPALS-algorithm. A partial least square regression (PLS-R) model for the sugar concentrations was created with mean centering, full cross-validation and Kernel-algorithm. All cotton samples of each concentrations have been used to develop the PLS-R model. The PLS-R model was tested by predicting the honeydew

content on the real cotton samples. Three different areas from these real cotton samples were also investigated by PCA with aforementioned settings. To show the quality of the model, PCA was combined with a quadratic discriminant analysis (QDA, 5 PCs). A fourth area was predicted by the PCA-QDA model.

MATLAB (MATLAB 9.2.0, Mathworks, MA, USA) and PLS_Toolbox (PLS Toolbox 8.5.1, Eigenvector Research, Inc., Wenatchee, WA, USA) were used for presenting the data.

6.4 Results and Discussion

6.4.1 Cotton samples impregnated with sugar

Cotton samples were investigated using hyperspectral imaging in the UV region (225 nm – 410 nm). In total, 21 samples were measured. The reference sample set was created to get a proper model to predict the honeydew content on real cotton samples. Figure 6.4a shows the averaged absorbance spectra in terms of reflectance. A baseline correction was applied to eliminate the spectral offset due to scattering. In general, the spectral shapes of all samples are quite similar. The most dominant band is pronounced approximately at 332 nm and weak shoulder can be recognized at 346 nm (sh). An absorption band is observed at 261 nm. Two weak shoulders are remarked at 291 nm. Despite the efficiency of the detector and the weak intensity light source in the spectral region between 250-270 nm [19], the signal is less intense, but nevertheless contain useful spectroscopic information for the actual application. A prove of the remaining performance of the setup in this range is given in Appendix Figure 10.3.

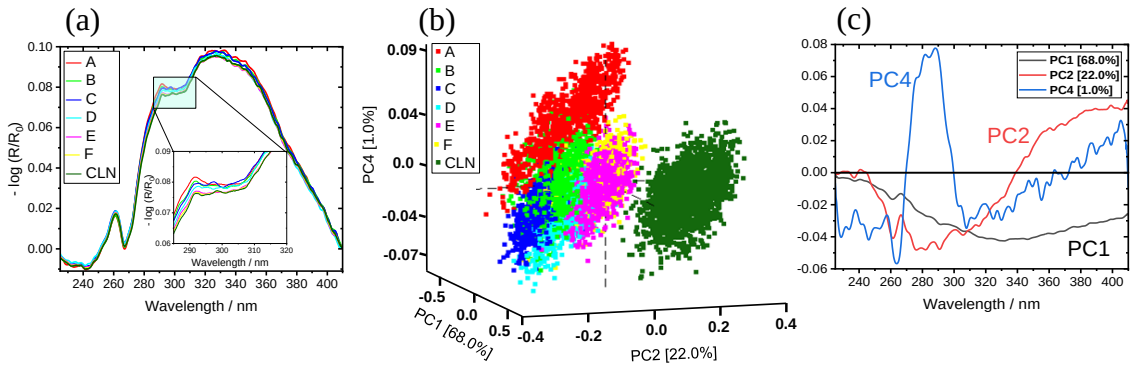


Figure 6.4: (a) Averaged UV spectra of cotton samples with sugar solutions in different concentrations: A (2 wt %, red), B (1 wt %, light green), C (0.5 wt %, blue), D (0.25 wt %, light blue), E (0.0125 wt %, pink), F (0.0625 wt %, yellow) and CLN (mechanically cleaned, dark green). PCA sugar model for the cotton samples with (b) scores and (c) corresponding loadings (PC1 black, PC2 red and PC4 blue).

Figure 6.4b and Figure 6.4c present the PCA model of the cotton samples with different concentrations of sugar. Figure 6.4b shows the scores plot for the first (68.0 %), second (22.0 %) and fourth (1.0 %) principal component (PCs). These PCs explain nearly 91.0 % of the total variance. The variance on PC3 is not necessary to distinguish between different sugar concentrations. For completeness, PC3 is displayed in supplementary materials Appendix Figure 10.2. Different sugar concentrations on cotton can clearly be distinguished by the PCA scores. The mechanically cleaned sample (CLN) is separated on PC2. PC4 shows the separation between the highest sugar concentration and the lowest concentration. Slight overlapping is observed due to inhomogeneity of the impregnation procedure for the samples with sugar. The overlap tendency increases from higher to lower concentrations as well as the variance within a sample increases with the concentration. Each cluster overlaps with the two closest sugar concentrations (higher and lower).

Figure 6.4c shows the loadings plot for PC1, PC2 and PC4. The strongest influence on PC1 is at 330 nm. Most of PC1 describes the morphology of the fiber itself. PC2 has a minimum at 280 nm and a maximum at 380 nm. These bands are responsible for the separation of the CLN sample from the others and are distinguishing between the different concentrations. For PC4 a maximum contribution is observed at 285 nm. These bands can be assigned to the presence of protein (see also Appendix Figure 10.3) [141]. The most significant differences between those loadings are found in the spectral region between 290 nm and 380 nm.

PLS-R is used for quantitative spectroscopic analysis. A PLS-R model was developed with a calibration sample set $n = 21$ to correlate the spectral information with the sugar concentration. Cotton samples (Table 6.2) were used for testing the performance of the model with a cross validation.

The variance explained by the model for the X- and Y-variables was 84 % by using five factors. Accordingly, the five PLS factors were sufficient to describe the correlation between the spectra and the sugar content. The accuracy of the calibration and validation were evaluated using the coefficient of determination (R^2) for the calibration ($R^2_c = 0.84$) and cross validation ($R^2_{cv} = 0.84$) model. The root mean square error of calibration (RMSEC = 0.009 g) and cross validation (RMSECV = 0.009 g) indicates the model performance. A high R^2_c and R^2_{cv} was achieved with an extremely low RMSEC and RMSECV.

Figure 6.5 shows the PLS-R model for the cotton soaked in different concentrations of sugar in the UV region (200 nm – 380 nm). Figure 6.5a presents the correlation between reference vs. predicted, while the regression coefficients for the five factor model is shown in Figure 6.5b. For

model building and understanding the PLS-R factor loadings, loading weights for all five factors are displayed in the Appendix Figure 10.4. Sample E and sample F have similar ratios 0.0326 and 0.0322 (sugar / g per dried cotton / g), respectively, due to the preparation procedure's limit. Therefore, they are overlapping in the reference vs. predicted plot. A negative band at 263 nm and a positive band at 284 nm can be assigned to protein absorbance. An average spectrum of pure dried protein is shown in the supplementary materials (Appendix Figure 10.3). The protein information is pronounced in the spectra and mandatory for the model, even though the illumination and detector should be optimized [19]. From 300 nm – 400 nm several features are observed that cannot be related to a common reason.

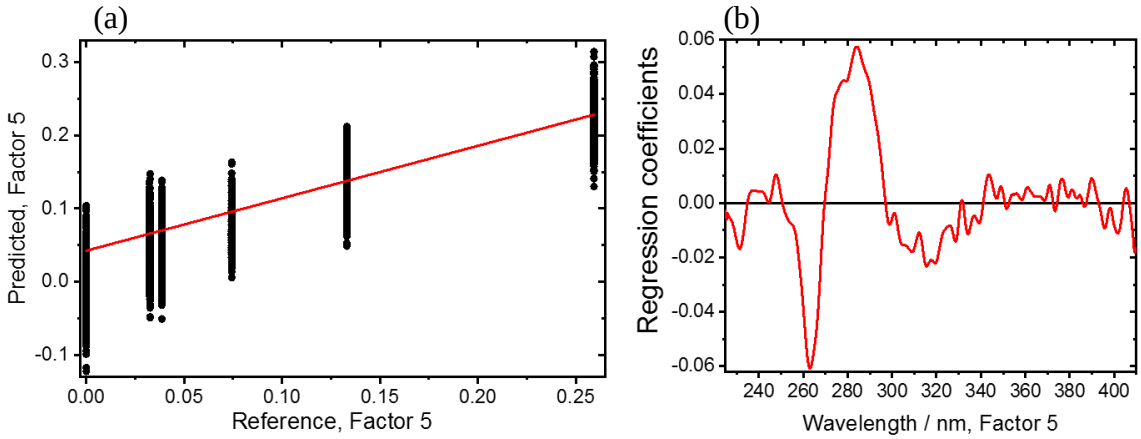


Figure 6.5: PLS-R model for different sugar concentrations in the UV region (225 nm – 410 nm). (a) Predicted vs. reference plot and (b) corresponding regression coefficients for the sugar content with a five factor PLS-R model.

6.4.2 Predicting the amount of sugar and honeydew based on the sugar PLS-R model

The performance of the PLS-R model was tested by two methods. First, cleaned cotton samples were manually sprayed with aforementioned sugar concentrations to get a distribution of sugar droplets on the cotton surfaces. One benefit of hyperspectral imaging is to get the lateral information. Therefore, the PLS-R model was used to predict the sugar content on the different samples, the result is shown in Figure 6.6. In the distribution map a clear lateral classification of the different sugar concentrations resulting in different ratios of sugar / g per dried cotton / g was achieved. From sample A to sample F the ratios decrease, CLN samples are not soaked in sugar. Each sample was prepared three times which are shown in the rows. Again, sample E and sample F are indistinguishable, due to the preparation procedure's limit. Overall, the averaged predicted ratios decrease from samples A to samples CLN. The lateral inhomogeneities become visible through hyperspectral imaging.

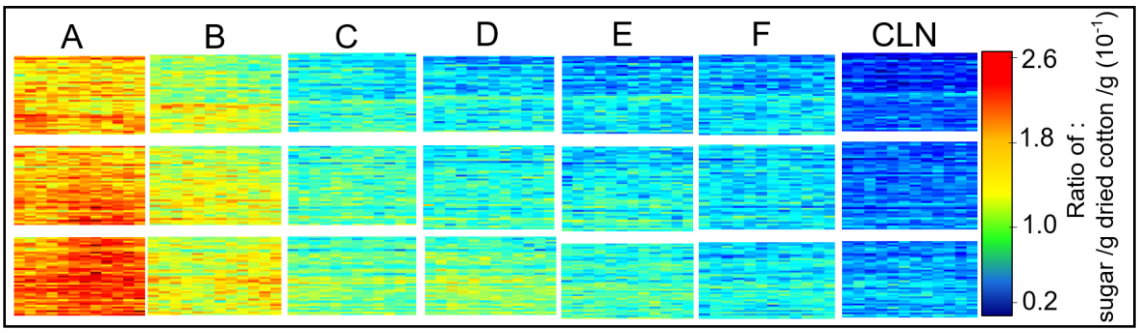


Figure 6.6: Distribution maps of the sugar content predicted on the mechanically cleaned cotton samples, which are manually sprayed by sugar solution. The prediction of each pixel is based on the PLS-R sugar model. Each rectangle represents a single cotton sample: A (2 wt %), B (1 wt %), C (0. 5 wt %), D (0. 25 wt %), E (0.125 wt %), F (0.0625 wt %) and CLN (mechanically cleaned). The colored pixels (see the score value range) represent the sugar content, from low (blue) to high (red).

Second, the PLS-R model was used to predict the honeydew content for each pixel of the real cotton types labeled: light, strong and very strong. In Figure 6.7, the resulting distribution maps are shown. As described for Figure 6.6, the distribution map shows a clear lateral classification of different ratios of sugar / g per dried cotton / g. The amount of sugar highly correlates with the amount of honeydew. Honeydew consists of different types of sugars and proteins [209]. From the very strong samples to the light samples the ratios decrease. Three samples per type were collected are given in the rows.

Therefore, in Figure 6.7 red and blue pixels represent high and low concentrations of honeydew, respectively. As expected, the light sample displays a low honeydew concentration, while the other two samples show an increase in the laterally resolved honeydew concentration. Compared to the samples shown in Figure 6.6 the real samples show a more heterogeneous distribution of honeydew on the samples. Even in the strong and very strong samples regions can be found where almost no honeydew is present. This can be seen in the presence of blue pixels on the distribution map of the very strong and strong samples.

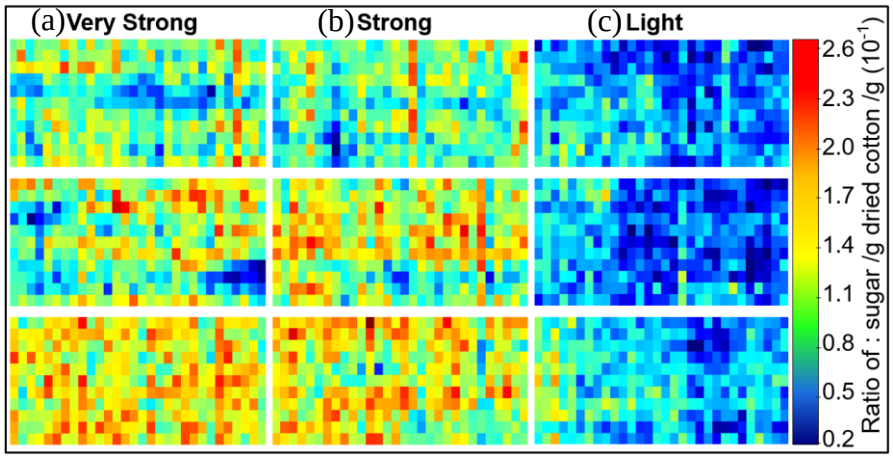


Figure 6.7: Distribution maps of the sugar content predicted on the real cotton samples, which are contaminated by honeydew. The prediction of each pixel is based on the PLS-R sugar model. Each rectangle represents a single cotton sample ((a) very strong, (b) strong, (c) light). The colored pixels (see the score value range) represent the sugar content, from low (blue) to high (red).

6.5 Conclusions

In summary, this proof of principle study has successfully demonstrated the identification and quantification of honeydew on real cotton samples by combining UV hyperspectral imaging (225 nm – 410 nm) with multivariate data analysis. For this novel approach, a reference sample set was created based on mechanically cleaned cotton which has been impregnated with honeydew typical sugar and protein solutions for further UV hyperspectral imaging investigations. The PCA model enables to classify the cotton samples according to their sugar concentration. A PLS-R model was created that is able to predict laterally resolved the sugar/honeydew content pixel by pixel. This is shown for reference samples and for real cotton samples that were labeled as light, strong and very strong contaminated by honeydew. The lateral distribution of the ratio of sugar / g per dried cotton / g per pixel giving a deeper insight into the distribution of honeydew on real cotton samples.

To the best of our knowledge, this is the first scientific work reporting the identification, quantification and distribution of the amount of honeydew content by UV hyperspectral imaging. This approach may provide an advantage in the industrial environment in the practical process application and commercialization in the future. It enables to control the honeydew contamination in the industrial processing of cotton fibers in real time. Hence, each cotton batch, independent of the honeydew amount, can be manufactured to minimize waste and costs.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Appendix Table 10.2: Description of the cotton sample preparation; Appendix Figure 10.2: PCA sugar model for the cotton samples with (a) scores and (b) corresponding loadings (PC1 black, PC3 red); Appendix Figure 10.2: Mean spectrum of pure dried protein on PTFE.

Author Contribution: Conceptualization, M.A.K, M.S.; methodology, M.A.K and M.S.; software, M.A.K, M.S. and F.W.; validation, M.A.K and M.S.; formal analysis, M.A.K., M.S. and F.W.; investigation, M.A.K and M.S.; resources, E.O., K.R., and M.B.; data curation, M.A.K., M.S. and F.W.; writing—original draft preparation, M.A.K and M.S.; writing—review and editing, M.A.K, M.S., F.W., E.O., V.J., K.R. and M.B.; visualization, M.A.K and M.S.; supervision, K.R. and M.B.; project administration, M.B.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing Ph.D. thesis.

Acknowledgments: We especially thank Thomas Blum (Reutlingen University) for his advises and valuable discussion for this study. Additionally, our thanks goes to the company innospec GmbH (Nürnberg, Germany) for helpful discussions and their expertise in this field. We would like to thank Gunther Pehl (Reutlingen University) for manufacturing the sample holder. Last but not least, we thank Barbara Boldrini and Tobias Drieschner (Reutlingen University) for her supporting contributions as an expert in hyperspectral imaging.

Conflicts of Interest: The authors declare no conflict of interest.

7 Paper V: Applying UV Hyperspectral Imaging for Quantification of Honeydew Content on Raw Cotton via PCA and PLS-R Models

Mona Knoblich[1,2,†], Mohammad Al Ktash[1,2,†], Frank Wackenhut[1], Volker Jehle[3], Edwin Ostertag[1] and Marc Brecht[1,2,]*

[1]Center of Process Analysis and Technology (PA&T), School of Life Sciences, Reutlingen University, Alteburgstraße 150, 72762 Reutlingen, Germany, mona.knoblich@reutlingen-university.de (M.K.); mohammad.alktash@reutlingen-university.de (M.A.K.); frank.wackenhut@reutlingen-university.de (F.W.); edwin.ostertag@reutlingen-university.de (E.O.)
[2]Institute of Physical and Theoretical Chemistry, Eberhard Karls University Tübingen, Auf der Morgenstelle 18, 72076 Tübingen, Germany
[3]Texoversum Faculty Textile, Reutlingen University, Alteburgstraße 150, 72762 Reutlingen, Germany, volker.jehle@reutlingen-university.de
[*] Correspondence: Marc.Brecht@Reutlingen-University.de.
[†]These authors contributed equally to the work.

This is originally published in Textiles (https://doi.org/10.3390/textiles3030019) as

"Knoblich, M.; Al Ktash, M.; Wackenhut, F.; Jehle, V.; Ostertag, E. and Brecht, M. Applying UV Hyperspectral Imaging for the Quantification of Honeydew Content on Raw Cotton via PCA and PLS-R Models. Textiles 2023, 23, 287-293. https://doi.org/10.3390/textiles3030019"

7.1 Abstract

Cotton contamination by honeydew is considered one of the significant problems for quality in textiles as it causes stickiness during manufacturing. Therefore, millions of dollars in losses are attributed to honeydew contamination each year. This work presents the use of UV hyperspectral imaging (225–300 nm) to characterize honeydew contamination on raw cotton samples. As reference samples, cotton samples were soaked in solutions containing sugar and proteins at different concentrations to mimic honeydew. Multivariate techniques such as a principal component analysis (PCA) and partial least squares regression (PLS-R) were used to predict and classify the amount of honeydew at each pixel of a hyperspectral image of raw cotton samples. The results show that the PCA model was able to differentiate cotton samples based on their sugar concentrations. The first two principal components (PCs) explain nearly 91.0% of the total variance. A PLS-R model was built, showing a performance with a coefficient of determination for the validation (R^2_{cv}) = 0.91 and root mean square error of cross-validation (RMSECV) = 0.036 g. This PLS-R model was able to predict the honeydew content in grams on raw cotton samples for each pixel. In conclusion, UV hyperspectral imaging, in combination with multivariate data analysis, shows high potential for quality control in textiles

7.2 Introduction

Cotton is widely regarded as an essential natural material in various textile products, from fabrics to clothing [210,211]. It is considered one of the most imported and exported materials worldwide [141]. Therefore, the assessment of the cotton quality is needed. Cotton contamination is one of the most significant problems for quality [93,95,97,108,114,115,210,212]. The most relevant impurities in raw cotton arise from insects producing honeydew. Honeydew is sugar-rich, excreted by whiteflies and aphids, causing stickiness during manufacturing [213,214]. Therefore, it can cause problems during processing, and the final product shows low quality. Modern techniques and methods have appeared due to the increasing demand for higher processing and quality control. These include off-line methods such as thermogravimetric analysis and single point spectroscopy. However, these techniques are slow and time-consuming [36,91,103,162,168,173,201]. In contrast, in- and on-line methods, such as hyperspectral imaging, are non-destructive and rapid, enabling real-time data acquisition and analysis [141]. Hyperspectral imaging is a type of spectroscopic imaging that allows for collecting and analyzing massive data spanning a wide wavelength range. It involves both spectral and spatial information with high resolution. Hyperspectral imaging generates large amounts of data, requiring multivariate data analysis techniques such as principal

component analysis (PCA) and partial least squares regression (PLS-R) [54]. PCA can identify and visualize groups within data clusters, while PLS-R is used to build quantitative models and generate data clusters. It is also helpful for evaluating the robustness of these models, making it a powerful tool for data analysis. Combining these two techniques is often required to analyze and interpret the results of high-resolution hyperspectral imaging effectively [19,21,22,53,115,141]. In a previous study, we developed a method using UV imaging to predict honeydew quantity on cotton samples. The approach involves using a xenon-arc lamp to quantify the amount of honeydew in the UV-A and UV-B ranges. However, it could not accurately detect it in UV-C due to the lamp's intensity limitations [54]. In this study, we overcome this limitation by using a deuterium lamp as light source. Mechanically cleaned cotton was soaked with a sugar and protein containing solution at different concentrations that are typical for honeydew. Chemometric models such as PCA and PLS-R were established using UV hyperspectral images. Cotton samples were categorized by sugar concentration using PCA, while PLS-R was used to correlate UV spectra with sugar concentration. The PLS-R model accurately predicted the amount of honeydew in grams on the raw cotton samples.

7.3 Materials and Methods

7.3.1 Chemicals and preparation of solutions and samples

The sugar and protein solutions applied to the cotton samples were formulated to mimic natural honeydew [208,215,216]. 0.2 g of each macronutrient (glucose, fructose, sucrose, melezitose, trehalose, and protein) was weighed and dissolved in 10 mL of deionized water. A six-fold serial dilution was prepared in 50 mL volumetric flasks by mixing 25 mL of the previous solution with 25 mL of deionized water for 2 minutes at each dilution step (Table 7.1).

In total, 24 mechanically cleaned cotton samples were prepared with a weight of 0.3 g ± 1 mg of each sample. The samples were dried in a vacuum oven (Vacutherm VT 6130 M, Thermo Fisher Scientific Inc., Waltham, MA, USA) at 30 °C and 50 mbar for 8 h to remove absorbed humidity. 4 mL of the aforementioned solution was used for soaking three samples per concentration. The samples were dried again in a desiccator at room temperature for one month.

Raw cotton samples were collected by ICA Bremen GmbH (Bremen, Germany) to test the model's predictive power. The samples were chosen according to their honeydew content in the steps light, strong, and very strong [54,211]. The sample types are named from A to F, and one mechanically cleaned (CLN) sample, where A has the highest concentration of sugar and protein solution

(4 wt%), and G has the lowest concentration (0.0625 wt%) (Table 7.1). The average ratio of sugar mass to dried cotton mass (m_{sugar}/m_{cotton}) remaining on the samples was calculated after drying the sample for one month (Table 7.1). The term macronutrients is omitted to describe the solution and replaced with the short-term "sugar" for the sample nomenclature.

Table 7.1: The sugar solution concentration and the weighted average sugar on cotton samples.

Sample type	Sugar concentration / wt %	m_{sugar}/m_{cotton}
A	4	0.4249
B	2	0.2413
C	1	0.1194
D	0.5	0.0609
E	0.25	0.02313
F	0.125	0.0126
G	0.0625	0.0143
CLN	-	-

7.3.2 UV hyperspectral imaging setup and data processing

Compared to our previous studies [19,53,54], the illumination of the hyperspectral imaging setup was modified; now a deuterium lamp (SL 3, StellarNet Inc, 24 V, 65.04 W, Tampa, Florida, USA) is used, providing a higher illumination strength in UV-C region compared to the xenon-arc lamp (e.g around 230 nm the intensity difference for deuterium illumination higher than xenon-arc lamp). Thus, the PTFE tunnel covering the convey belt for increasing the illumination strength was no longer necessary.

Multivariate data analysis was acquired with “Aspen Unscrambler™, version 10.5.1” (Aspen Technology Inc., Bedford, MA, USA). The PCA model was calculated with mean centering, cross-validation, and the NIPALS algorithm. A PLS-R model for the sugar concentrations was processed with mean centering, category variable with eight segmented cross-validations, and the Kernel algorithm.

7.4 Results and Discussion

The averaged absorbance spectra in terms of reflectance after a linear baseline correction are shown in Figure 7.1a. The spectra show an almost linear decrease in the reflectivity for all sugar concentrations. In the range of 275 nm and 295 nm are broad bands showing clear dependences on the sugar concentration. These bands can be assigned to protein, cellulose, and lignin. A much weaker band between 230–255 nm corresponds to the presence of pectin and DNA [217-221].

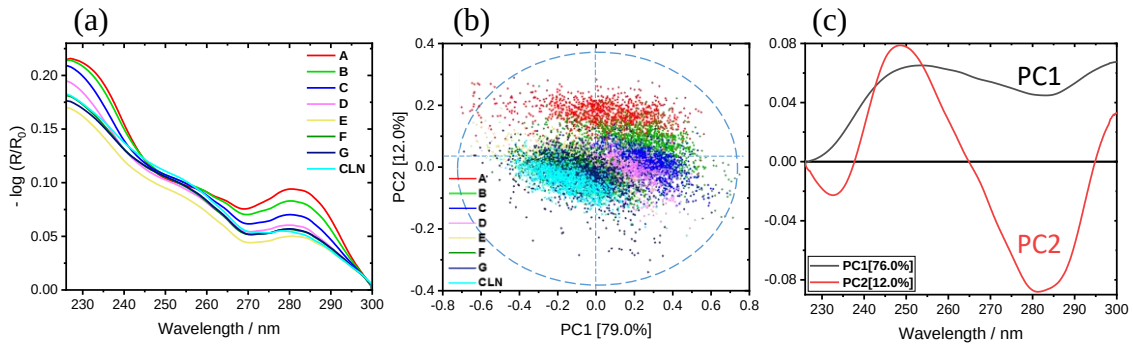


Figure 7.1: (**a**) Averaged spectra recorded via UV hyperspectral imaging of raw cotton samples with sugar solutions in different concentrations: A (4 wt %), B (2 wt %), C (1 wt %), D (0.5 wt %), E (0.25 wt %), F (0.125 wt %), G (0.0625 wt %), and CLN (mechanically cleaned). PCA sugar model for the cotton samples with (**b**) scores on the first principal component (PC1) and second principal component (PC2) and (**c**) corresponding loadings.

Figure 7.1b,c present the cotton samples' PCA model at each sample pixel with different sugar concentrations. Figure 1b shows the scores plot for the first (79.0%) and second (12.0%) principal components (PCs). These PCs explain nearly 91.0% of the total variance.

The PCA scores enable to distinguish different sugar concentrations on cotton. On PC1, high sugar concentrations are separated from low concentrations, while on PC2, the mechanically cleaned sample (CLN) shows distinct separation from the sample with high sugar concentrations. Moreover, different sugar concentrations on cotton can clearly be distinguished by the PC2. An overlap naturally results from the preparation method chosen which results in a certain inhomogeneity. With decreasing concentration, the degree of overlap between samples increases together with the variance within the sample. Each cluster shows overlap with the two nearest sugar concentrations (higher and lower). Figure 7.1c shows the loadings plots for PC1 and PC2. The most significant differences between those loadings are found between 250 nm - 280 nm in the spectral region. The maximum influence on PC1 at 250 nm and the minimum at 283 nm. Most of PC1 describes a clear dependence on the concentrations of sugar on the cotton samples. PC2 has a maximum at 249 nm and a minimum at 282 nm. These bands represent the chromophores, pectin, and DNA in the cotton fibers [217,222].

PLS-R was utilized as a technique for quantitative spectroscopic analysis. A PLS-R model was developed using a calibration sample set of 24 samples to establish a correlation between the spectral information and the sugar content. The PLS-R model's performance was tested using cotton samples (Table 7.1) with different concentrations of sugar solutions.

The PLS-R model for the *X*- and *Y*-variables explained 91% of the variance. Five PLS-R factors were sufficient to describe the correlation between the spectra and sugar content. The accuracy of

the calibration and validation were evaluated using the coefficient of determination (R^2) for the calibration (R^2_c = 0.9) and validation (R^2_{cv} = 0.91) model. The quality of the models were evaluated according to values of the error of calibration, RMSEC = 0.03 g, and the error of cross-validation, RMSECV = 0.036 g. High R^2_c and R^2_{cv} values are achieved with extremely low RMSEC and RMSECV values.

Figure 7.2 presents the PLS-R model for cotton samples soaked with different concentrations of sugar. For model building and understanding the PLS-R factor loadings, loading weights for all three factors are displayed in the supplementary materials Appendix Figure 10.5. Figure 7.2a displays the correlation between the predicted and reference values, whereas the regression coefficients for the three-factor model are illustrated in Figure 7.2b. Samples F and G have similar ratios, 0.0324 and 0.0321 (m_{sugar}/m_{cotton}), hence they overlap in the regression coefficients plot. Two negative bands at 235 nm and 282 nm and one positive band at 250 nm can be assigned to protein and pectin absorbance [141,217].

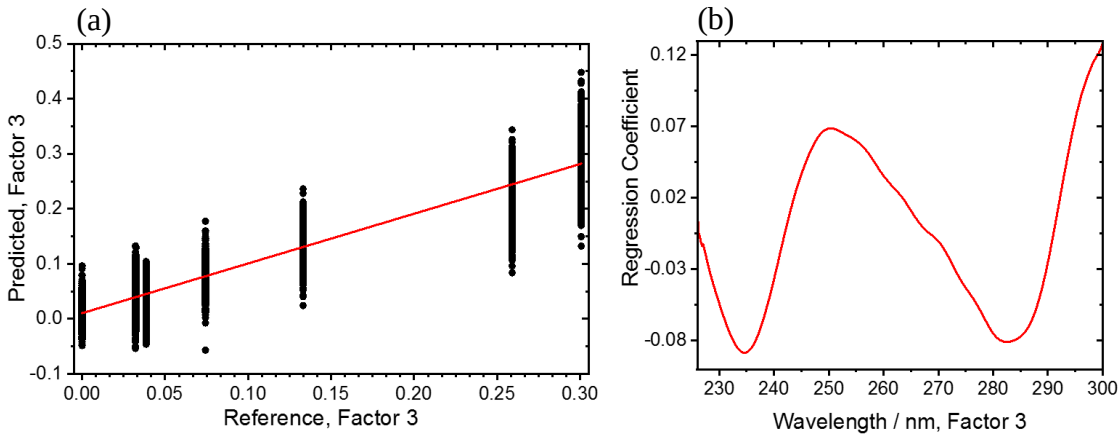


Figure 7.2: Five-factor PLS-R model for different sugar contents in the UV region (225 nm – 300 nm). (a) Predicted vs. reference plot and (b) corresponding regression coefficients.

The PLS-R model was used to predict the honeydew content for each pixel of a hyperspectral image. Three raw cotton samples of three grades of honeydew contamination (very strong, strong, and light) were collected, and the resulting distribution maps are shown in Figure 7.3. The distribution maps present a clear lateral classification of different ratios of m_{sugar}/m_{cotton}, and the predicted ratios decrease from the very strong samples to the light samples. The sugar content is highly correlated with the honeydew amount [217]. The analysis reveals a highly variable distribution of honeydew across all samples. Some regions present minimal contamination, while others, including areas/pixels in the light samples, exhibit up to 0.1 m_{sugar}/m_{cotton} ratio, comparable to those found in very strong samples. The observed inhomogeneity in honeydew distribution sug-

gests that our soaking method for the sugar solution is a realistic approach, as it induces a comparable level of variability. However, the inhomogeneity seems to be even higher in the raw samples, as shown in Figure 7.3.

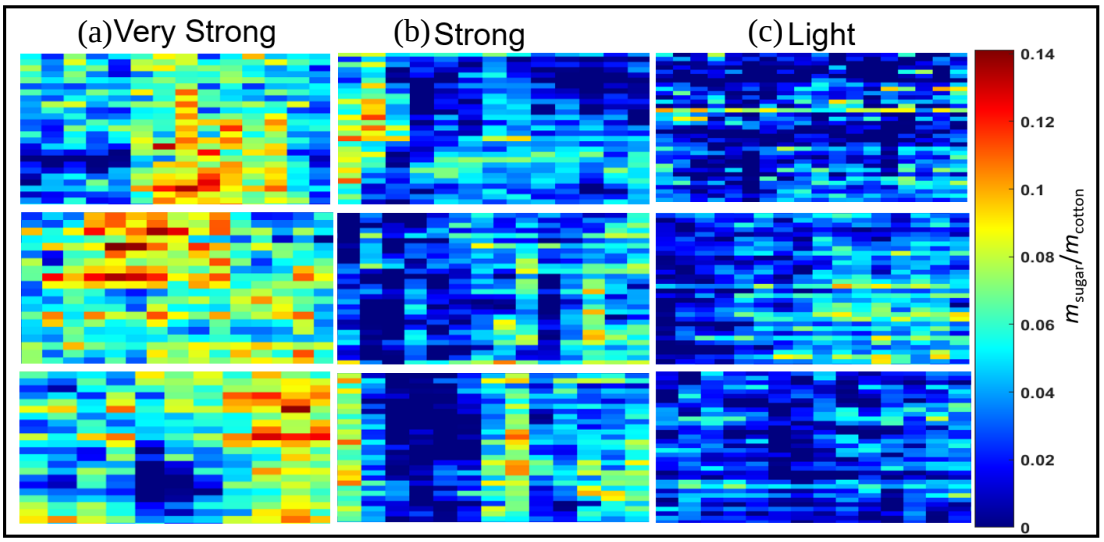


Figure 7.3: Distribution maps of the sugar content predicted for each pixel of the UV hyperspectral imaging data from the five-factor PLS-R model on the raw cotton samples contaminated by honeydew. Each rectangle represents a single cotton sample ((a) very strong, (b) strong, (c) light). The colored pixels (see the score value range) represent the sugar content, from low (blue) to high (red).

7.5 Conclusions

UV hyperspectral imaging (225–300 nm) was combined with multivariate data analysis to successfully identify and quantify honeydew on raw cotton samples. Therefore, a reference sample set based on cotton samples was prepared and imaged in UV.

The samples were soaked with solutions containing sugar and proteins at different concentrations to mimic honeydew. A PCA model enabled the classification of the cotton samples according to their sugar concentrations. The PLS-R model was able to predict laterally resolved honeydew content pixel by pixel in grams on raw cotton samples. The analysis reveals that the raw cotton samples have an inhomogeneous distribution of honeydew. Therefore, the chosen soaking method closely approximates the distribution patterns observed in the raw samples. The results were obtained by analyzing samples labeled as light, strong, and very strong contaminated with honeydew. This combination of hyper-spectral imaging with multivariate data analysis represents a high potential technique for detecting honeydew contamination in real-time.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Figure S1. X-loading weights and x-loadings for factor 1 (a, b), factor 2 (c, d), and factor 3 (e, f), respectively.

Author Contribution: Conceptualization, M.A.K, M.S.; methodology, M.A.K and M.S.; software, M.A.K, M.S. and F.W.; validation, M.A.K and M.S.; formal analysis, M.A.K., M.S. and F.W.; inves-tigation, M.A.K and M.S.; resources, E.O. and M.B.; data curation, M.A.K., M.S. and F.W.; writ-ing—original draft preparation, M.A.K and M.S.; writing—review and editing, M.A.K, M.S., F.W., E.O., V.J. and M.B.; visualization, M.A.K and M.S.; supervision M.B.; project administration, M.B.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing Ph.D. thesis.

Conflicts of Interest: The authors declare no conflict of interest.

8 Conclusion and Summary

In the first part of this thesis presented the application of UV hyperspectral imaging for the characterization of active pharmaceutical ingredients (APIs) in tablets. Two sets of samples were analyzed: one set containing tablets with 100% API content and another set consisting of commercially available painkiller tablets. The results demonstrated that UV hyperspectral imaging, combined with PCA, is a promising approach for detecting and differentiating different drug samples. The PCA models generated were able to effectively separate and classify the various drug types based on their spectral characteristics. In particular, the first two PCs captured most of the spectral variance and allowed for clear discrimination between the samples. The advantage of the home-built UV hyperspectral imaging setup used in the study was its spatial and spectral resolution and data acquisition speed sufficient for scientific purposes. The results obtained from this setup suggest that it can be easily adapted to meet the requirements of real industrial processes. The comparison between the hyperspectral imaging results and reference measurements performed using a commercial UV spectrometer indicated that while the hyperspectral imaging data provided valuable information, there were some limitations. The sensitivity and efficiency of the hyperspectral imaging setup were lower in the wavelength range below 275 nm, resulting in less reliable spectroscopic information in that region. However, despite this limitation, combining UV hyperspectral imaging and chemometric modeling enabled accurate classification and separation of the different drug samples. Overall, the study demonstrated the potential of UV hyperspectral imaging as a tool for rapid and non-destructive characterization of pharmaceutical tablets. The technique has implications for quality control, process monitoring, and real-time release testing in industrial settings. Further advancements in UV hyperspectral imaging technology and data analysis methods are expected to enhance its applications in the pharmaceutical industry.

To confirm the effectiveness of hyperspectral imaging, 28 direct bonded copper samples were characterized to determine the oxidation states of copper and measure the thickness of the oxide layer in the UV region. Single-point UV spectroscopy was used as a basis of comparison for the results obtained from the UV hyperspectral imaging setup. The results showed that UV hyperspectral imaging, in combination with PCA and PLS-R, offered a promising approach for detecting and differentiating copper states and measuring oxide layer thickness with lateral resolution. PCA models effectively separated the different types of direct bonded copper based on their copper states and oxide layer thicknesses using only the first two PCs. With three factors, the PLS-R

models exhibited high coefficients of determination (R^2) and low RMSE for calibration, cross-validation, and prediction. The models successfully correlated the spectral features with the oxide layer thickness. Notably, this study is the first to report the identification and quantification of copper oxide thin films using UV hyperspectral imaging. The UV hyperspectral imaging setup utilized in this study offered a high spatial and spectral resolution, along with relatively fast data acquisition under laboratory conditions. The design and data presented in this study provide a foundation for developing a UV hyperspectral imaging setup that meets the requirements of real industrial processes. Overall, this study demonstrated the efficacy of UV hyperspectral imaging and UV reflectance spectroscopy in characterizing direct bonded copper samples, paving the way for future applications in industrial processes for real-time data acquisition, process control, and in-line classification.

In the second part of this thesis, it was demonstrated that the integration of UV-Vis/NIR reflectance spectroscopy and hyperspectral imaging, coupled with PCA, offers a powerful technique for detecting and differentiating raw cotton types. It was determined that the most crucial information for distinguishing cotton types was present in the UV and NIR spectra. In the UV range, the most significant factor was identified as the protein contribution at 280 nm, while the key contributors to absorbance in the NIR range were CH_3 vibrations at 1775 nm and ROH vibrations at 1500 nm. PCA analysis of the spectral data successfully classified all cotton varieties, with the first two PCs accounting for the maximum data variance. Moreover, the investigation revealed that the most dominant absorbance corresponding to CH_3 and ROH in the case of NIR hyperspectral imaging occurred at 2270 nm and 1525 nm, respectively. Two methods were employed to process the substantial data generated, facilitating the differentiation of all cotton varieties. The rugged online home-built setup used in the study presented benefits such as spatial/spectral resolution and rapid data collection, enabling the assessment of multiple samples in a brief period and at a low cost. In conclusion, based on the provided data, it is feasible to develop a streamlined chemometric model that adheres to the demands of real-world industrial processes while maintaining suitable standards and accuracy. This approach could offer a practical and cost-effective means of distinguishing between various raw cotton types, potentially having considerable implications for the textile industry.

In the third part, UV hyperspectral imaging and multivariate data analysis techniques were successfully utilized to identify, quantify, and spatially resolve honeydew content on cotton samples. The aim is to address the need for real-time control of honeydew contamination in the industrial processing of cotton fibers. A total of 21 cotton samples labeled A to F and one mechanically

cleaned (CLN) were prepared by soaking in sugar with different concentrations ranging from 2 wt % to 0.0625 wt %. These samples were used as reference samples to create a dataset for the UV hyperspectral imaging investigations. Additionally, real cotton samples were chosen according to honeydew contents in the steps of light, strong, and very strong. These samples were utilized to assess the performance of the developed models. The honeydew contamination was achieved by exposing the cotton samples to honeydew typical sugar and protein solutions in real-world simulating conditions. The results showed that UV hyperspectral imaging, in combination with PCA and PLS-R, offered a promising approach for predicting the amount of honeydew contaminated on real cotton. PCA is applied to classify the cotton samples based on their sugar concentrations. The PCA model effectively separates samples with different sugar concentrations, allowing for accurate classification.

Additionally, the model reveals the spectral regions that significantly contribute to the classification, such as those associated with fiber morphology and protein presence. To achieve quantitative analysis, a PLS-R model is developed. The PLS-R model establishes a strong correlation between spectral information and sugar concentration, enabling accurate sugar content prediction on reference and real cotton samples. This predictive capability provides valuable insights into the distribution of honeydew on cotton samples.

The study emphasizes the novelty of its findings, as it is the first scientific work to report the identification, quantification, and distribution mapping of honeydew content using UV hyperspectral imaging. The spatial mapping of sugar concentrations and honeydew contamination on cotton samples provides a deeper understanding and control over honeydew contamination in the industrial processing of cotton fibers. Overall, this research holds promise for practical applications in the industry by enabling real-time control of honeydew contamination, optimizing manufacturing processes, and minimizing waste. The combination of UV hyperspectral imaging and multivariate data analysis techniques offers a powerful approach for efficiently and accurately assessing honeydew contamination on cotton samples. Further exploration of this methodology on a larger scale and its potential application to other agricultural products could enhance productivity and sustainability across various industries.

Finally, The UV hyperspectral imaging prototype underwent modifications compared to previous studies. Specifically, a deuterium lamp was utilized as the illumination source, providing higher illumination strength in the UV-C region compared to the previous xenon-arc lamp, eliminating the need for a PTFE tunnel to increase illumination strength. These modifications enhanced the performance and accuracy of the hyperspectral imaging system for honeydew detection on cotton

samples. Therefore, applying UV hyperspectral imaging combined with multivariate data analysis techniques, PCA and PLS-R, proved an effective method for identifying and quantifying honeydew contamination on raw cotton samples. By creating a reference sample set that mimicked honeydew using sugar and protein solutions at different concentrations, the UV hyperspectral imaging system was able to capture and analyze the spectral information of the cotton samples.

The PCA model showed the ability to differentiate cotton samples based on their sugar concentrations, providing a means for categorizing the contamination levels. On the other hand, the PLS-R model demonstrated strong performance in predicting and quantifying the honeydew content in grams on a pixel-by-pixel basis. The model achieved a high accuracy with an R^2 of 0.91 and low RMSECV = 0.036 g.

The results revealed the presence of distinct spectral bands within the UV range (225 nm – 300 nm) that were correlated with sugar concentration and indicative of honeydew contamination. The distribution maps generated by the PLS-R model showed lateral classification of different sugar ratios and revealed the inhomogeneity of honeydew distribution on the raw cotton samples. The spraying and soaking method employed to simulate honeydew contamination proved realistic and induced comparable levels of variability in the samples.

Overall, this study demonstrates the potential of UV hyperspectral imaging combined with multivariate data analysis as a rapid and non-destructive technique for quality control in textiles. Accurately identifying and quantifying honeydew contamination on raw cotton samples can significantly improve the quality and processing of cotton in the textile industry, thereby reducing the financial losses associated with such contamination. Further research and optimization of the technique may enhance its applicability and broaden its potential for other quality control applications in textiles and beyond.

9 Bibliography

1. Atascientific. Spectrometry and spectroscopy: what's the difference? Available online: https://www.atascientific.com.au/spectrometry/ (accessed on 17.01.2020).
2. Wikipedia. Spectroscopy. Available online: https://en.wikipedia.org/wiki/Spectroscopy (accessed on 19.04.2023).
3. Penner, M.H. Basic principles of spectroscopy. *Food analysis* **2017**, 79-88, doi:10.1007/978-3-319-45776-5_6.
4. Smith, G.S. *An introduction to classical electromagnetic radiation*; Cambridge University Press: 1997.
5. LibreTexts. The Electromagnetic Spectrum. Available online: https://chem.libretexts.org/@go/page/2398 (accessed on 23.03.2023).
6. Skrabal, P.M. *Spectroscopy: An interdisciplinary integral description of spectroscopy from UV to NMR*; vdf Hochschulverlag AG: 2012.
7. Sindhu, P. *Fundamentals of Molecular Spectroscopy*; New Age International: 2006.
8. Gribov, L. Introduction to molecular spectroscopy. *Moscow Izdatel Nauka* **1976**.
9. Hardesty, J.H.; Attili, B. Spectrophotometry and the Beer-Lambert Law: An important analytical technique in chemistry. *Collin College Department of Chemistry* **2010**.
10. Parnis, J.M.; Oldham, K.B. Beyond the Beer–Lambert law: The dependence of absorbance on time in photochemistry. *Journal of Photochemistry Photobiology* **2013**, *267*, 6-10, doi:10.1016/j.jphotochem.2013.06.006.
11. Fellers, T.; Davidson, M. Introduction to the Reflection of Light. Available online: https://www.olympus-lifescience.com/en/microscope-resource/primer/lightandcolor/reflectionintro/ (accessed on 01.04.2023).
12. Blanco, M.; Villarroya, I. NIR spectroscopy: a rapid-response analytical tool. *TrAC Trends in Analytical Chemistry* **2002**, *21*, 240-250, doi:10.1016/S0165-9936(02)00404-1.
13. Manso, M.; Carvalho, M.L. Application of spectroscopic techniques for the study of paper documents: A survey. *Spectrochimica Acta* **2009**, *64*, 482-490, doi:10.1016/j.sab.2009.01.009.
14. Haroon, K.; Arafeh, A.; Martin, P.; Rodgers, T.; Mendoza, Ć.; Baker, M. Use of inline near-infrared spectroscopy to predict the viscosity of shampoo using multivariate analysis. *International journal of cosmetic science* **2019**, *41*, 346-356, doi:10.1111/ics.12536.
15. Ni, C.; Li, Z.; Zhang, X.; Sun, X.; Huang, Y.; Zhao, L.; Zhu, T.; Wang, D. Online Sorting of the Film on Cotton Based on Deep Learning and Hyperspectral Imaging. *IEEE Access* **2020**, *8*, 93028-93038, doi:10.1109/ACCESS.2020.2994913.
16. Haven, J.J.; Junkers, T. Online monitoring of polymerizations: current status. *European Journal of Organic Chemistry* **2017**, *2017*, 6474-6482, doi:10.1002/ejoc.201700851.
17. Lewis, E.; Schoppelrei, J.; Lee, E.; Kidder, L. Near-infrared chemical imaging as a process analytical tool. *Process analytical technology* **2008**, p.187-225.
18. Li, Q.; He, X.; Wang, Y.; Liu, H.; Xu, D.; Guo, F. Review of spectral imaging technology in biomedical engineering: achievements and challenges. *Biomedical optics* **2013**, *18*, 100901, doi:10.1117/1.JBO.18.10.100901.
19. Al Ktash, M.; Stefanakis, M.; Boldrini, B.; Ostertag, E.; Brecht, M. Characterization of Pharmaceutical Tablets Using UV Hyperspectral Imaging as a Rapid In-Line Analysis Tool. *Sensors* **2021**, *21*, 4436, doi:10.3390/s21134436.

20. Gowen, A.A.; O'Donnell, C.P.; Cullen, P.J.; Downey, G.; Frias, J.M. Hyperspectral imaging–an emerging process analytical tool for food quality and safety control. *Trends in food science & technology* **2007**, *18*, 590-598.
21. Boldrini, B.; Kessler, W.; Rebner, K.; Kessler, R.W. Hyperspectral imaging: a review of best practice, performance and pitfalls for in-line and on-line applications. *Journal of near infrared spectroscopy* **2012**, *20*, 483-508, doi:10.1255/1003.
22. Gowen, A.; Odonnell, C.; Cullen, P.; Downey, G.; Frias, J. Hyperspectral imaging – an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology* **2007**, *18*, 590-598, doi:10.1016/j.tifs.2007.06.001.
23. Al Ktash, M.; Hauler, O.; Ostertag, E.; Brecht, M. Ultraviolet-visible/near infrared spectroscopy and hyperspectral imaging to study the different types of raw cotton. *Journal of Spectral Imaging* **2020**, *9*, a1, doi:10.1255/jsi.2020.a18.
24. Lu, G.; Fei, B. Medical hyperspectral imaging: a review. *Biomed Opt* **2014**, *19*, 10901, doi:10.1117/1.JBO.19.1.010901.
25. Manolakis, D.; Shaw, G. Detection algorithms for hyperspectral imaging applications. *IEEE signal processing magazine* **2002**, *19*, 29-43, doi:10.1109/79.974724.
26. Stiedl, J.; Boldrini, B.; Green, S.; Chassé, T.; Rebner, K. Characterisation of oxide layers on technical copper based on visible hyperspectral imaging. *Journal of Spectral Imaging* **2019**, *8*, doi:10.1255/jsi.2019.a10.
27. Grahn, H.; Geladi, P. *Techniques and applications of hyperspectral image analysis*; John Wiley & Sons: 2007.
28. Amigo, J.M.; Santos, C. Preprocessing of hyperspectral and multispectral images. In *Data Handling in Science and Technology*; Elsevier: 2020; Volume 32, pp. 37-53.
29. Vasefi, F.; MacKinnon, N.; Farkas, D. Hyperspectral and multispectral imaging in dermatology. In *Imaging in Dermatology*; Elsevier: 2016; pp. 187-201.
30. Gao, L.; Smith, R.T. Optical hyperspectral imaging in microscopy and spectroscopy–a review of data acquisition. *Journal of biophotonics* **2015**, *8*, 441-456, doi:10.1002/jbio.201400051.
31. Bassler, M.C.; Stefanakis, M.; Sequeira, I.; Ostertag, E.; Wagner, A.; Bartsch, J.W.; Roessler, M.; Mandic, R.; Reddmann, E.F.; Lorenz, A.; et al. Comparison of Whiskbroom and Pushbroom darkfield elastic light scattering spectroscopic imaging for head and neck cancer identification in a mouse model. *Anal Bioanal Chem* **2021**, *413*, 7363-7383, doi:10.1007/s00216-021-03726-5.
32. Kamruzzaman, M.; ElMasry, G.; Sun, D.-W.; Allen, P. Application of NIR hyperspectral imaging for discrimination of lamb muscles. *Journal of food engineering* **2011**, *104*, 332-340, doi:10.1016/j.jfoodeng.2010.12.024.
33. Türker-Kaya, S.; Huck, C.W. A review of mid-infrared and near-infrared imaging: principles, concepts and applications in plant tissue analysis. *Molecules* **2017**, *22*, 168, doi:10.3390/molecules22010168.
34. Ariana, D.P.; Lu, R. Quality evaluation of pickling cucumbers using hyperspectral reflectance and transmittance imaging: Part I. Development of a prototype. *Sensing Instrumentation for Food Quality Safety* **2008**, *2*, 144-151, doi:10.1007/s11694-008-9057-x.
35. Barnaby, J.Y.; Huggins, T.D.; Lee, H.; McClung, A.M.; Pinson, S.R.; Oh, M.; Bauchan, G.R.; Tarpley, L.; Lee, K.; Kim, M.S. Vis/NIR hyperspectral imaging distinguishes sub-population, production environment, and physicochemical grain properties in rice. *Scientific Reports* **2020**, *10*, 1-13.
36. Rodgers, J.; Beck, K. NIR characterization and measurement of the cotton content of dyed blend fabrics. *Textile research journal* **2009**, *79*, 675-686, doi:10.1177/0040517508090884.

37. Schlapfer, D.R.; Kaiser, J.W.; Brazile, J.; Schaepman, M.E.; Itten, K.I. Calibration concept for potential optical aberrations of the APEX pushbroom imaging spectrometer. In Proceedings of the Sensors, Systems, and Next-Generation Satellites VII, 2004; pp. 221-231.
38. Bro, R.; Smilde, A.K. Principal component analysis. *Analytical methods* **2014**, *6*, 2812-2831.
39. Calvini, R.; Ulrici, A.; Amigo, J. Sparse-Based Modeling of Hyperspectral Data. In *Data Handling in Science and Technology*; Elsevier: 2016; Volume 30, pp. 613-634.
40. DESY, D.E.-S. P66 Time-resolved luminescence spectroscopy. Available online: https://photon-science.desy.de/facilities/petra_iii/beamlines/p66_superlumi/index_eng.html (accessed on 06.09.2023).
41. Wiedemann, H.; Wiedemann, H. *Synchrotron radiation*; Springer: 2003.
42. Song, Z.; Li, J.; Davis, K.D.; Li, X.; Zhang, J.; Zhang, L.; Sun, X. Emerging Applications of Synchrotron Radiation X-Ray Techniques in Single Atomic Catalysts. *Small Methods* **2022**, *6*, 2201078, doi:10.1002/smtd.202201078.
43. Kim, K.J. Characteristics of synchrotron radiation. In Proceedings of the AIP conference proceedings, 1989; pp. 565-632.
44. Chawla, A.; Lobacz, A.; Tarapata, J.; Zulewska, J. UV light application as a mean for disinfection applied in the dairy industry. *Applied Sciences* **2021**, *11*, 7285, doi:10.3390/app11167285.
45. Tokode, O.; Prabhu, R.; Lawton, L.A.; Robertson, P.K. UV LED sources for heterogeneous photocatalysis. *Environmental Photochemistry Part III* **2015**, 159-179, doi:10.1007/698_2014_306.
46. Ditmire, T.; Gumbrell, E.; Smith, R.; Mountford, L.; Hutchinson, M. Supersonic ionization wave driven by radiation transport in a short-pulse laser-produced plasma. *Physical review letters* **1996**, *77*, 498, doi:10.1103/PhysRevLett.77.498.
47. Saunders, R.; Ott, W.; Bridges, J. Spectral irradiance standard for the ultraviolet: the deuterium lamp. *Applied Optics* **1978**, *17*, 593-600, doi:10.1364/AO.17.000593.
48. Zama, T.; Awazu, K.; Onuki, H. Improvement of the aging characteristics of deuterium lamp. *Journal of electron spectroscopy related phenomena* **1996**, *80*, 493-496, doi:10.1016/0368-2048(96)03024-1.
49. Lin, B.J. Deep UV lithography. *Journal of vacuum science technology* **1975**, *12*, 1317-1320, doi:10.1116/1.568527.
50. Davidson, M.W. Fundamentals of Mercury Arc Lamp. Available online: https://zeiss-campus.magnet.fsu.edu/articles/lightsources/mercuryarc.html (accessed on 15.05.2023).
51. Davidson, M.W. Fundamentals of Xenon Arc Lamp. Available online: https://zeiss-campus.magnet.fsu.edu/articles/lightsources/xenonarc.html (accessed on 15.05.2023).
52. Andreas Nolte, L.H., Michael W. Davidson. Fundamentals of Illumination Sources for Optical Microscopy. Available online: https://zeiss-campus.magnet.fsu.edu/articles/lightsources/lightsourcefundamentals.html (accessed on 15.05.2023).
53. Al Ktash, M.; Stefanakis, M.; Englert, T.; Drechsel, M.S.; Stiedl, J.; Green, S.; Jacob, T.; Boldrini, B.; Ostertag, E.; Rebner, K.; et al. UV hyperspectral imaging as process analytical tool for the characterization of oxide layers and copper states on direct bonded copper. *Sensors* **2021**, *21*, 7332, doi:10.3390/s21217332.
54. Al Ktash, M.; Stefanakis, M.; Wackenhut, F.; Jehle, V.; Ostertag, E.; Rebner, K.; Brecht, M. Prediction of Honeydew Contaminations on Cotton Samples by In-Line UV Hyperspectral Imaging. *Sensors* **2023**, *23*, 319, doi:10.3390/s23010319.

55. Knoblich, M.; Al Ktash, M.; Wackenhut, F.; Jehle, V.; Ostertag, E.; Brecht, M. Applying UV Hyperspectral Imaging for the Quantification of Honeydew Content on Raw Cotton via PCA and PLS-R Models. *Textiles* **2023**, *3*, 287-293, doi:10.3390/textiles3030019.
56. Trygg, J.; Holmes, E.; Lundstedt, T. Chemometrics in metabonomics. *Journal of proteome research* **2007**, *6*, 469-479, doi:10.1021/pr060594q.
57. Wold, S. Chemometrics; what do we mean with it, and what do we want from it? *Chemometrics Intelligent Laboratory Systems* **1995**, *30*, 109-115, doi:10.1016/0169-7439(95)00042-9.
58. Kumar, N.; Bansal, A.; Sarma, G.; Rawal, R.K. Chemometrics tools used in analytical chemistry: An overview. *Talanta* **2014**, *123*, 186-199, doi:10.1016/j.talanta.2014.02.003.
59. Rebner, K. Ortsaufgelöste Streulichtspektroskopie an mikrostrukturierten Systemen. Eberhard-Karls-Universität Tübingen, 2010.
60. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemometrics Intelligent Laboratory Systems* **1987**, *2*, 37-52, doi:10.1016/0169-7439(87)80084-9.
61. Geladi, P.; Isaksson, H.; Lindqvist, L.; Wold, S.; Esbensen, K. Principal component analysis of multivariate images. *Chemometrics Intelligent Laboratory Systems* **1989**, *5*, 209-220, doi:10.1016/0169-7439(89)80049-8.
62. Esbensen, K.H.; Guyot, D.; Westad, F.; Houmoller, L.P. *Multivariate data analysis: in practice: an introduction to multivariate data analysis and experimental design*, 6 ed.; Multivariate Data Analysis: 2002.
63. Kazakevich, Y.V.; Lobrutto, R. *HPLC for pharmaceutical scientists*; John Wiley & Sons: 2007.
64. Stefanakis, M.; Lorenz, A.; Bartsch, J.W.; Bassler, M.C.; Wagner, A.; Brecht, M.; Pagenstecher, A.; Schittenhelm, J.; Boldrini, B.; Hakelberg, S. Formalin fixation as tissue preprocessing for multimodal optical spectroscopy using the example of human brain tumour cross sections. *Journal of spectroscopy* **2021**, *2021*, 1-14, doi:10.1155/2021/5598309.
65. Mika, S.; Ratsch, G.; Weston, J.; Scholkopf, B.; Mullers, K.-R. Fisher discriminant analysis with kernels. In Proceedings of the Neural networks for signal processing IX: Proceedings of the 1999 IEEE signal processing society workshop (cat. no. 98th8468), 1999; pp. 41-48.
66. Bassler, M.C. Spectroscopic and multivariate approaches for tumor diagnostics and-therapy. Dissertation, Tübingen, Germany, 2022.
67. Stefanakis, M.; Lorenz, A.; Bartsch, J.W.; Bassler, M.C.; Wagner, A.; Brecht, M.; Pagenstecher, A.; Schittenhelm, J.; Boldrini, B.; Hakelberg, S. Formalin Fixation as Tissue Preprocessing for Multimodal Optical Spectroscopy Using the Example of Human Brain Tumour Cross Sections. *Journal of Spectroscopy* **2021**, *2021*, 14, doi:10.1155/2021/5598309.
68. Kessler, W. *Multivariate datenanalyse: für die pharma, bio-und Prozessanalytik*; John Wiley & Sons: 2011.
69. Esbensen, K.H.; Swarbrick, B. *Multivariate data analysis: an introduction to multivariate analysis, process analytical technology and quality by design*; Camo: 2018.
70. Ozer, D.J. Correlation and the coefficient of determination. *Psychological bulletin* **1985**, *97*, 307, doi:10.1037/0033-2909.97.2.307.
71. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE). *Geoscientific model development* **2014**, *7*, 1525-1534, doi:10.5194/gmdd-7-1525-2014.
72. Wikipedia. Root-mean-square deviation. Available online: https://en.wikipedia.org/wiki/Root-mean-square_deviation (accessed on 20.03.2023).
73. Lee, D.K.; In, J.; Lee, S. Standard deviation and standard error of the mean. *Korean journal of anesthesiology* **2015**, *68*, 220-223, doi:10.4097/kjae.2015.68.3.220.

74. Saffarian, S.; Elson, E.L. Statistical analysis of fluorescence correlation spectroscopy: the standard deviation and bias. *Biophysical journal* **2003**, *84*, 2030-2042, doi:10.1016/S0006-3495(03)75011-5.
75. Organization, W.H. March 2014 supplement to the 2013 consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. **2014**.
76. Abebe, W. Herbal medication: potential for adverse interactions with analgesic drugs. *Journal of clinical pharmacy*

therapeutics **2002**, *27*, 391-401, doi:10.1046/j.1365-2710.2002.00444.x.
77. Amani, M.B.; Kindenge, J.M.; Baruti, E.T.; Bakiantima, E.N.; Agasa, S.B.; Hubert, P.; Marini Djang'eing'a, R. Quality control of tramadol in kisangani: development, validation, and application of a uv-vis spectroscopic method. *American Journal of Analytical Chemistry* **2021**, *12*, 295-309, doi:10.4236/ajac.2021.128018
78. Sankar, R.; Snehalatha, K.S.; Firdose, S.T.; Babu, P.S. Applications in HPLC in pharmaceutical analysis. *International Journal of Pharmaceutical Sciences* **2019**, *59*, 117-124.
79. Saeed, M.; Ahmed, Q. Estimation of paracetamol, aspirin, ibuprofen, codeine and caffeine in some formulated commercial dosage using UV–spectroscopic method. *Eur. J. Pharm. Med. Res* **2017**, *4*, 33-38.
80. Akhtar, S.; Kareem, L.; Arif, A.; Siddiqui, M.; Hakeem, A. Development of a ceramic-based composite for direct bonded copper substrate. *Ceramics International* **2017**, *43*, 5236-5246, doi:10.1016/j.ceramint.2017.01.049.
81. Schulz-Harder, J. Advantages and new development of direct bonded copper substrates. *Microelectronics Reliability* **2003**, *43*, 359-365, doi:10.1016/S0026-2714(02)00343-8.
82. Gong, Y.; Lee, C.; Yang, C. Atomic force microscopy and Raman spectroscopy studies on the oxidation of Cu thin films. *Journal of Applied Physics* **1995**, *77*, 5422-5425, doi:10.1063/1.359234.
83. Lenglet, M.; Kartouni, K.; Delahaye, D. Characterization of copper oxidation by linear potential sweep voltammetry and UV-Visible-NIR diffuse reflectance spectroscopy. *Journal of applied electrochemistry* **1991**, *21*, 697-702, doi:10.1007/BF01034048.
84. Graham, M. Recent advances in oxide film characterization. *Pure applied chemistry* **1992**, *64*, 1641-1645, doi:10.1351/pac199264111641.
85. Liu, Y.; Gamble, G.; Thibodeaux, D. UV/visible/near-infrared reflectance models for the rapid and non-destructive prediction and classification of cotton color and physical indices. *Transactions of the ASABE* **2010**, *53*, 1341-1348, doi:10.13031/2013.32584.
86. Tschannerl, J.; Ren, J.; Jack, F.; Krause, J.; Zhao, H.; Huang, W.; Marshall, S. Potential of UV and SWIR hyperspectral imaging for determination of levels of phenolic flavour compounds in peated barley malt. *Food chemistry* **2019**, *270*, 105-112, doi:10.1016/j.foodchem.2018.07.089.
87. Stiedl, J.; Green, S.; Chassé, T.; Rebner, K. Characterization of oxide layers on technical copper material using ultraviolet visible (UV–Vis) spectroscopy as a rapid on-line analysis tool. *Applied Spectroscopy* **2019**, *73*, 59-66, doi:10.1364/AS.73.000059.
88. Wang, H.; Memon, H.J.P.S., Properties. *Cotton science and processing technology*; Springer: Singapore, 2020; Volume 5, pp. 79-98.
89. Tokel, D.; Dogan, I.; Hocaoglu-Ozyigit, A.; Ozyigit, I.I. Cotton agriculture in Turkey and worldwide economic impacts of Turkish cotton. *Journal of Natural Fibers* **2022**, *19*, 10648-10667, doi:10.1080/15440478.2021.2002759.
90. Karadag, R. Cotton Dyeing with Cochineal by Just in Time Extraction, Mordanting, Dyeing, and Fixing Method in the Textile Industry. *Journal of Natural Fibers* **2023**, *20*, 1-11, doi:10.1080/15440478.2022.2108184.

91. Fuhrer, L. Mapping of In-Field Cotton Fiber Quality Utilizing John Deere's Harvest Identification System (Hid). University of Georgia, 2022.
92. Ge, Y. *Mapping in-field cotton fiber quality and relating it to soil moisture*; Texas A&M University: 2007.
93. Campbell, B.T.; Hinze, L. Cotton production, processing and uses of cotton raw material. In *Industrial crops and uses*; CABI Wallingford UK: 2010; pp. 259-276.
94. BlouseRoumaine. What is slowfashion movement? Available online: https://www.blouseroumaine-shop.com/en/blog/what-is-slowfashion-movement (accessed on 23.02.2018).
95. Barotova, A.; Xurramov, A.; Raxmatullayev, S.; Ismoilova, A. Evaluation of fiber quality indexes in different varieties of cotton plants. *Journal of Agriculture Horticulture* **2023**, *3*, 41-46, doi:10.5281/zenodo.7655573.
96. Ravandi, S.H.; Valizadeh, M. Properties of fibers and fabrics that contribute to human comfort. In *Improving comfort in clothing*; Elsevier: 2011; pp. 61-78.
97. Chand, N.; Fahim, M. of NFPC, editors. Cotton reinforced polymer composites. **2008**, doi:10.1533/9781845695057.129.
98. Yu, C. Natural textile fibres: vegetable fibres. In *Textiles and fashion*; Elsevier: 2015; pp. 29-56.
99. Hequet, E.F.; Abidi, N.; Ethridge, D. Processing sticky cotton: Effect of stickiness on yarn quality. *Textile Research Journal* **2005**, *75*, 402-410, doi:10.1177/0040514505053953.
100. Chung, C.; Lee, M.; Choe, E.K. Characterization of cotton fabric scouring by FT-IR ATR spectroscopy. *Carbohydrate Polymers* **2004**, *58*, 417-420, doi:10.1016/j.carbpol.2004.08.005.
101. Behera, P.; Aravind, S.; Seetharaman, B. Honeydew contaminated cotton: a sticky problem needs a solution. *Research Journal of Textile and Apparel* **2022**, doi:10.1108/RJTA-05-2022-0053.
102. Hequet, E.; Abidi, N. Processing sticky cotton: implication of trehalulose in residue build-up. *Journal of Cotton Science* **2002**, *6*, 77-90.
103. Abidi, N.; Hequet, E. Fourier transform infrared analysis of cotton contamination. *Textile Research Journal* **2007**, *77*, 77-84, doi:10.1177/0040517507074624.
104. Amada44. Honeydew. Available online: https://www.thedailygarden.us/garden-word-of-the-day/honeydew (accessed on 02.09.2023).
105. Vazquez, J.R. Bemisia tabaci. Available online: https://www.biodiversidadvirtual.org/insectarium/Bemisia-tabaci-img925792.html (accessed on 02.09.2023).
106. Hequet, E.; Abidi, N. Effects of the origin of the honeydew contamination on cotton spinning performances. *Textile Research Journal* **2005**, *75*, 699-709, doi:10.1177/0040517505053909.
107. Perkins Jr, H.H. Identification and processing of honeydew-contaminated cottons. *Textile Research Journal* **1983**, *53*, 508-512.
108. Afzal, M.I.J.I.o.t.M.o.C.P.i.Z.a.b.S. Cotton stickiness–A marketing and processing problem. **2001**, 105-111.
109. Domelsmith, L.; Berni, R. Potassium: A new marker for washed cotton. *Textile Research Journal* **1984**, *54*, 210-214.
110. Fischer, J. Evaluation of Cleaning and Washing Processes for Cotton Fiber: Part VII. Microbiological Evaluation 1. *Textile Research Journal* **1980**, *50*, 93-95.
111. Tzanov, T.; Calafell, M.; Guebitz, G.M.; Cavaco-Paulo, A. Bio-preparation of cotton fabrics. *Enzyme Microbial Technology* **2001**, *29*, 357-362, doi:10.1016/S0141-0229(01)00388-X.
112. Shams, N.A.; Mohajerani, M. Evaluation of cotton fibers stickiness by colorimetric method. **2013**.

113. Luttrell, R.; Fitt, G.; Ramalho, F.; Sugonyaev, E. Cotton pest management: Part 1. A worldwide perspective. *Annual Review of Entomology* **1994**, *39*, 517-526.
114. Bradow, J.M.; Davidonis, G.H. Quantitation of fiber quality and the cotton production-processing interface: a physiologist's perspective. *Cotton Science* **2000**, *4*, 34-64.
115. Jiang, Y.; Li, C. Detection and discrimination of cotton foreign matter using push-broom based hyperspectral imaging: System design and capability. *PloS one* **2015**, *10*, e0121969, doi:10.1371/journal.pone.0121969.
116. Knowlton, J.L. USDA Cotton Classing and the Need for a Universal Cotton Quality Evaluation System. In Proceedings of the EFS® Systems Conference, Turkey, http://www. cottoninc. com/EFSConference, 2005.
117. Yang, S.; Gordon, S. Accurate prediction of cotton ring-spun yarn quality from high-volume instrument and mill processing data. *Textile Research Journal* **2017**, *87*, 1025-1039, doi:10.1177/0040517516646051.
118. Sayeed, M.A.; Schumann, M.; Wanjura, J.; Kelly, B.R.; Smith, W.; Hequet, E.F. Characterizing the total within-sample variation in cotton fiber length using the High Volume Instrument fibrogram. *Textile Research Journal* **2021**, *91*, 175-187, doi:10.1177/0040517520935212.
119. Steven Brown, T.S. How to Think About Fiber Quality in Cotton. Available online: https://www.aces.edu/blog/topics/crop-production/how-to-think-about-fiber-quality-in-cotton/ (accessed on 05.08.2022).
120. Barnhardt-Purified-cotton. Properties of Cotton. Available online: https://barnhardtcotton.net/technology/cotton-properties/ (accessed on 01.04.2023).
121. Zhou, J.; Xu, B. Reliability of cotton fiber length distributions measured by dual-beard fibrography and advanced fiber information system. *Cellulose* **2021**, *28*, 1753-1767, doi:10.1007/s10570-020-03611-x.
122. Abd El-Ghany, N.M.; Abd El-Aziz, S.E.; Marei, S.S. A review: application of remote sensing as a promising strategy for insect pests and diseases management. *Environmental Science and Pollution Research* **2020**, 1-13, doi:10.1007/s11356-020-09517-2.
123. Rego, C.H.Q.; França-Silva, F.; Gomes-Junior, F.G.; Moraes, M.H.D.d.; Medeiros, A.D.d.; Silva, C.B.d. Using Multispectral Imaging for Detecting Seed-Borne Fungi in Cowpea. *Agriculture* **2020**, *10*, 361, doi:10.3390/agriculture10080361.
124. Bannon, D. Hyperspectral imaging: Cubes and slices. *Nature photonics* **2009**, *3*, 627, doi:10.1038/nphoton.2009.205.
125. Tschannerl, J.; Ren, J.; Jack, F.; Krause, J.; Zhao, H.; Huang, W.; Marshall, S. Potential of UV and SWIR hyperspectral imaging for determination of levels of phenolic flavour compounds in peated barley malt. *Food chemistry* **2019**, *270*, 105-112, doi:10.1016/j.foodchem.2018.07.089.
126. Lodhi, V.; Chakravarty, D.; Mitra, P. Hyperspectral imaging system: Development aspects and recent trends. *Sensing Imaging* **2019**, *20*, 1-24, doi:10.1007/s11220-019-0257-8.
127. Lu, G.; Fei, B. Medical hyperspectral imaging: a review. *Biomedical optics* **2014**, *19*, 010901, doi:10.1117/1.JBO.19.1.010901.
128. Rebner, K. Hyperspectral Imaging for Quality Analysis and Control. In Proceedings of the Applied Industrial Optics: Spectroscopy, Imaging and Metrology, 2016; p. AITh2B. 1.
129. Biancolillo, A.; Marini, F. Chemometric methods for spectroscopy-based pharmaceutical analysis. *Frontiers in chemistry* **2018**, *6*, 576, doi:10.3389/fchem.2018.00576.
130. Tonnesen, H.H. *Photostability of drugs and drug formulations*, 2nd ed.; CRC Press: Boca Raton, 2004; p. 448.
131. Murtaza, G.; Hussain, I.; Khan, S.A.; Shabbir, A.; Mahmood, A.; Asad, M.H.H.B.; Farzanal, K.; Malik, N.S. Development of a UV-spectrophotometric method for the simultaneous determination of aspirin and paracetamol in tablets. *Scientific research and Essays* **2011**, *6*, 417-421, doi:10.5897/SRE10.925.

132. Saeed, M.; Ahmed, Q. Estimation of paracetamol, aspirin, ibuprofen, codeine and caffeine in some formulated commercial dosage using UV–spectroscopic method. *Eur J Pharm Med Res* **2017**, *4*, 33-38.
133. Rote, A.R.; Kumbhoje, P.A.; Bhambar, R.S. UV-visible spectrophotometric simultaneous estimation of paracetamol and nabumetone by AUC method in combined tablet dosage form. *Pharmaceutical methods* **2012**, *3*, 40-43, doi:10.4103/2229-4708.97722.
134. Jolliffe, I.T.; Cadima, J. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A* **2016**, *374*, 20150202, doi:10.1098/rsta.2015.0202.
135. Bianchi, F.; Riboni, N.; Trolla, V.; Furlan, G.; Avantaggiato, G.; Iacobellis, G.; Careri, M. Differentiation of aged fibers by Raman spectroscopy and multivariate data analysis. *Talanta* **2016**, *154*, 467-473, doi:10.1016/j.talanta.2016.04.013.
136. Lawson-Wood, K.; Robertson, I. Pharmaceutical Assay and Multicomponent Analysis using the LAMBDA 365 UV/Vis Spectrophotometer. **2016**.
137. Barnaby, J.Y.; Huggins, T.D.; Lee, H.; McClung, A.M.; Pinson, S.R.; Oh, M.; Bauchan, G.R.; Tarpley, L.; Lee, K.; Kim, M.S. Vis/NIR hyperspectral imaging distinguishes sub-population, production environment, and physicochemical grain properties in rice. *Scientific reports* **2020**, *10*, 9284, doi:10.1038/s41598-020-65999-7.
138. Atif, M.; Farooq, W.; Fatehmulla, A.; Aslam, M.; Ali, S.M. Photovoltaic and impedance spectroscopy study of screen-printed TiO2 based CdS quantum dot sensitized solar cells. *Materials* **2015**, *8*, 355-367, doi:doi.org/10.3390/ma8010355.
139. Esa, S.R.; Yahya, R.; Hassan, A.; Omar, G. Nano-scale copper oxidation on leadframe surface. *Ionics* **2017**, *23*, 319-329, doi:10.1007/s11581-016-1894-8.
140. Council, N.R. *Expanding the vision of sensor materials*; National Academies Press: 1995.
141. Al Ktash, M.; Hauler, O.; Ostertag, E.; Brecht, M. Ultraviolet-visible/near infrared spectroscopy and hyperspectral imaging to study the different types of raw cotton. *Journal of Spectral Imaging* **2020**, *9*, doi:10.1255/jsi.2020.a18.
142. Lodhi, V.; Chakravarty, D.; Mitra, P. Hyperspectral imaging system: Development aspects and recent trends. *Sensing and Imaging* **2019**, *20*, 1-24, doi:10.1007/s11220-019-0257-8.
143. Jin, S.; Hui, W.; Wang, Y.; Huang, K.; Shi, Q.; Ying, C.; Liu, D.; Ye, Q.; Zhou, W.; Tian, J. Hyperspectral imaging using the single-pixel Fourier transform technique. *Scientific reports* **2017**, *7*, 45209, doi:10.1038/srep45209.
144. Willoughby, C.T.; Folkman, M.A.; Figueroa, M.A. Application of hyperspectral-imaging spectrometer systems to industrial inspection. In Proceedings of the Three-Dimensional and Unconventional Imaging for Industrial Inspection and Metrology, 1996; pp. 264-272.
145. Lu, G.; Fei, B. Medical hyperspectral imaging: a review. *Journal of biomedical optics* **2014**, *19*, 010901, doi:10.1117/1.JBO.19.1.010901.
146. Stiedl, J.; Green, S.; Chassé, T.; Rebner, K. Auger electron spectroscopy and UV–Vis spectroscopy in combination with multivariate curve resolution analysis to determine the Cu2O/CuO ratios in oxide layers on technical copper surfaces. *Applied Surface Science* **2019**, *486*, 354-361, doi:10.1016/j.apsusc.2019.05.028.
147. Mazzeo, G.; Prestopino, G.; Conte, G.; Salvatori, S. Metal-diamond-metal planar structures for off-angle UV beam positioning with high lateral resolution. *Sensors Actuators A: Physical* **2005**, *123*, 199-203, doi:10.1016/j.sna.2005.02.016.
148. Ojeda, C.B.; Rojas, F.S. Process analytical chemistry: applications of ultraviolet/visible spectrometry in environmental analysis: an overview. *Applied Spectroscopy Reviews* **2009**, *44*, 245-265, doi:10.1080/05704920902717898.
149. Mazzeo, G.; Prestopino, G.; Conte, G.; Salvatori, S. Metal-diamond-metal planar structures for off-angle UV beam positioning with high lateral resolution. *Sensors and Actuators A: Physical* **2005**, *123*, 199-203, doi:10.1016/j.sna.2005.02.016.

150. Reyes, G.; Diaz, W.; Toro, C.; Balladares, E.; Torres, S.; Parra, R.; Vásquez, A. Copper Oxide Spectral Emission Detection in Chalcopyrite and Copper Concentrate Combustion. *Processes* **2021**, *9*, 188, doi:10.3390/pr9020188.
151. Obeidat, S.M.; Al-Ktash, M.M.; Al-Momani, I.F. Study of fuel assessment and adulteration using EEMF and multiway PCA. *Energy & Fuels* **2014**, *28*, 4889-4894, doi:10.1021/ef500718e.
152. Geladi, P.; Kowalski, B.R. Partial least-squares regression: a tutorial. *Analytica chimica acta* **1986**, *185*, 1-17, doi:10.1016/0003-2670(86)80028-9.
153. inno-spec GmbH. BlueEye UV Hyperspectral Imaging Camera (220 – 380 nm). Available online: https://inno-spec.de/wp-content/uploads/2021/10/210928_BlueEye.pdf (accessed on 29.10.2021).
154. PCO AG. pco.edge 4.2 bi cooled sCMOS camera. Available online: https://www.pco.de/fileadmin/user_upload/pco-product_sheets/DS_PCOEDGE42BI_V104.pdf (accessed on 29.10.2021).
155. Quantum Desgin Europe GmbH. Lamp Spectra and Irradiance. Available online: https://qd-europe.com/fileadmin/Mediapool/products/lightsources/en/LQ_Lamp_spectra_and_irradiance_en.pdf (accessed on 27.09.2021).
156. Zhang, H.; Li, D. Applications of computer vision techniques to cotton foreign matter inspection: A review. *Computers and Electronics in Agriculture* **2014**, *109*, 59-70, doi:10.1016/j.compag.2014.09.004.
157. Morais, J.P.S.; de Freitas Rosa, M.; Nascimento, L.D.; do Nascimento, D.M.; Cassales, A.R. Extraction and characterization of nanocellulose structures from raw cotton linter. *Carbohydrate polymers* **2013**, *91*, 229-235, doi:10.1016/j.carbpol.2012.08.010.
158. Dohlman, E.; Johnson, J.; MacDonald, S.; Meyer, L.; Soley, G. The world and United States cotton outlook. In Proceedings of the 2015-02-20)[2015-04-01]. http: www. usda. gov/oce/forum, Arlingron/Virginia, 21.-22.02.2019.
159. Pray, C.E.; Huang, J.; Hu, R.; Rozelle, S. Five years of Bt cotton in China–the benefits continue. *The Plant Journal* **2002**, *31*, 423-430, doi:10.1046/j.1365-313X.2002.01401.
160. Heuvels, S.; Molenaar, J.; Raap, N.; Petit, C.; Shanmugavel, A.; Li, W. *Sustainable cotton ranking 2017*; Assesing company performance: October 2017 2017; p. 28.
161. Mustafic, A.; Jiang, Y.; Li, C. Cotton contamination detection and classification using hyperspectral fluorescence imaging. *Textile Research Journal* **2016**, *86*, 1574-1584, doi:10.1177/0040517515590416.
162. Ghule, A.V.; Chen, R.K.; Tzing, S.H.; Lo, J.; Ling, Y.C. Simple and rapid method for evaluating stickiness of cotton using thermogravimetric analysis. *Analytica Chimica Acta* **2004**, *502*, 251-256, doi:10.1016/j.aca.2003.10.021.
163. Abidi, N.; Hequet, E. Fourier Transform Infrared Analysis of Cotton Contamination. *Textile Research Journal* **2016**, *77*, 77-84, doi:10.1177/0040517507074624.
164. Was-Gubala, J.; Starczak, R. Nondestructive identification of dye mixtures in polyester and cotton fibers using raman spectroscopy and ultraviolet-visible (UV-Vis) microspectrophotometry. *Appl. Spectrosc.* **2015**, *69*, 296-303, doi:10.1366/14-07567.
165. Goldfarb, A.R.; Saidel, L.J.; Mosovich, E. The ultraviolet absorption spectra of proteins. *Journal of Biological Chemistry* **1951**, *193*, 397-404.
166. Zhou, F.; Ding, T. Detection of Cotton Lint Trash within the Ultraviolet—Visible Spectral Range. *J Applied spectroscopy* **2010**, *64*, 936-941, doi:10.1366/000370210792081091.
167. Peets, P.; Leito, I.; Pelt, J.; Vahur, S. Identification and classification of textile fibres using ATR-FT-IR spectroscopy with chemometric methods. *Spectrochimica Acta* **2017**, *173*, 175-181, doi:10.1016/j.saa.2016.09.007.

168. Chung, C.; Lee, M.; Choe, E. Characterization of cotton fabric scouring by FT-IR ATR spectroscopy. *Carbohydrate Polymers* **2004**, *58*, 417-420, doi:10.1016/j.carbpol.2004.08.005.
169. Blanco, M.; Coello, J.; Iturriaga, H.; Maspoch, S.; PageÁs, J. Use of near-infrared spectrometry in control analyses of acrylic ®bre manufacturing processes. *Analytica Chimica Acta* **1999**, *383*, 291±298, doi:10.1016/S0003-2670(98)00804-6.
170. Cleve, E.; Bach, E.; Schollmeyer, E. Using chemometric methods and NIR spectrophotometry in the textile industry. *Analytica Chimica Acta* **2000**, *420*, 163–167, doi:10.1016/S0003-2670(00)00888-6.
171. Fortier, C.A.; Rodgers, J.E.; Cintrón, M.S.; Xiaoliang, C.; Foulk, J.A. Identification of cotton and cotton trash components by Fourier transform near-infrared spectroscopy. *Textile Research Journal* **2010**, *81*, 230-238, doi:10.1177/0040517510383620.
172. Ruckebusch, C.; Orhan, F.; Durand, A.; Boubellouta, T.; Huvenne, J.P. Quantitative Analysis of Cotton–Polyester Textile Blends from Near-Infrared Spectra.pdf>. *Applied Spectroscopy* **2006**, *60*, 539 - 544, doi:10.1366/000370206777412194.
173. Liu, Y. Recent Progress in Fourier Transform Infrared (FTIR) Spectroscopy Study of Compositional, Structural and Physical Attributes of Developmental Cotton Fibers. *Materials* **2013**, *6*, 299-313, doi:10.3390/ma6010299.
174. Ríos-Reina, R.; García-González, D.L.; Callejón, R.M.; Amigo, J.M. NIR spectroscopy and chemometrics for the typification of Spanish wine vinegars with a protected designation of origin. *Food Control* **2018**, *89*, 108-116, doi:10.1016/j.foodcont.2018.01.031.
175. Obeidat, S.M.; Khanfar, M.S.; Obeidat, W.M. Classification of edible oils and uncovering adulteration of virgin olive oil using FTIR with the aid of chemometrics. *Australian Journal of Basic and Applied Sciences* **2009**, *3*, 2048-2053.
176. Schneider, A.; Feussner, H. *Biomedical engineering in gastrointestinal surgery*; Academic Press: 2017.
177. Colarusso, P.; Kidder, L.H.; Levin, I.W.; Fraser, J.C.; Arens, J.F.; Lewis, E.N. Infrared spectroscopic imaging: from planetary to cellular systems. *Applied Spectroscopy* **1998**, *52(3)*, 106A–120A, doi:10.1366/0003702981943545.
178. Mirschel, G.; Daikos, O.; Scherzer, T.; Steckert, C. Near-infrared chemical imaging used for in-line analysis of functional finishes on textiles. *Talanta* **2018**, *188*, 91-98, doi:10.1016/j.talanta.2018.05.050.
179. Zhu, S.; Zhou, L.; Gao, P.; Bao, Y.; He, Y.; Feng, L. Near-infrared hyperspectral imaging combined with deep learning to identify cotton seed varieties. *Molecules* **2019**, *24*, 3268, doi:10.3390/molecules24183268.
180. Martens, H.; Høy, M.; Wise, B.M.; Bro, R.; Brockhoff, P.B. Pre-whitening of data by covariance-weighted pre-processing. *Journal of Chemometrics* **2003**, *17*, 153-165, doi:10.1002/cem.780.
181. Wise, B.M.; Martens, H.; Høy, M.; Bro, R.; Brockhoff, P.B. Calibration transfer by generalized least squares. In Proceedings of the Proceedings of the Seventh Scandinavian Symposium on Chemometrics (SSC7), Copenhagen, Denmark, 2001; pp. 19-23.
182. Jirata, M.T.; Chelule, J.C.; Odhiambo, R. Deriving some estimators of panel data regression models with individual effects. *International Journal of Science and Research* **2012**, *3*, 53-59.
183. Liu, Y.; Delhom, C.; Campbell, B.T.; Martin, V. Application of near infrared spectroscopy in cotton fiber micronaire measurement. *Information Processing in Agriculture* **2016**, *3*, 30-35, doi:10.1016/j.inpa.2016.01.001.
184. Femenias, A.; Marín, S. *Hyperspectral Imaging*; Vicente M. Gómez-López, R.B., Ed.; 2021; pp. 363-390.

185. Park, B.; Lu, R. *Hyperspectral imaging technology in food and agriculture*; Barbosa-Ca'novas, G.V., Ed.; Springer: Athens, GA,USA, 2015.
186. Chen, S.-Y.; Chang, C.-Y.; Ou, C.-S.; Lien, C.-T. Detection of insect damage in green coffee beans using VIS-NIR hyperspectral imaging. *Remote Sensing* **2020**, *12*, 2348, doi:10.3390/rs12152348.
187. Devassy, B.M.; George, S. Estimation of strawberry firmness using hyperspectral imaging: a comparison of regression models. *Journal of Spectral Imaging* **2021**, *10*, doi:10.1255/jsi.2021.a3.
188. Daikos, O.; Scherzer, T. In-line monitoring of the residual moisture in impregnated black textile fabrics by hyperspectral imaging. *Progress in Organic Coatings* **2021**, 106610, doi:10.1016/j.porgcoat.2021.106610.
189. Wang, C.; Xu, M.; Jiang, Y.; Zhang, G.; Cui, H.; Deng, G.; Lu, Z. Toward Real Hyperspectral Image Stripe Removal via Direction Constraint Hierarchical Feature Cascade Networks. *Remote Sensing* **2022**, *14*, 467, doi:doi.org/10.3390/rs14030467.
190. Blanch-Perez-del-Notario, C.; Saeys, W.; Lambrechts, A. Hyperspectral imaging for textile sorting in the visible–near infrared range. *Journal of Spectral Imaging* **2019**, *8*, doi:10.1255/jsi.2019.a17.
191. Mirschel, G.; Daikos, O.; Scherzer, T. In-line monitoring of the thickness distribution of adhesive layers in black textile laminates by hyperspectral imaging. *Computers & Chemical Engineering* **2019**, *124*, 317-325, doi:10.1016/j.compchemeng.2019.01.015.
192. Feng, X.; Cheng, H.; Zuo, D.; Zhang, Y.; Wang, Q.; Lv, L.; Li, S.; Yu, J.Z.; Song, G. Genome-wide identification and expression analysis of GL2-interacting-repressor (GIR) genes during cotton fiber and fuzz development. *Planta* **2022**, *255*, 1-18, doi:10.1007/s00425-021-03737-7.
193. Anthony, W.S. *Improvement of the Marketability of Cotton Produced in Zones Affected by Stickiness*, GOURLOT J.-P. ed.; Common Fund for Commodities: France, 2001; p. 99.
194. Rony, A.N.U. Technical Properties of Cotton Fiber, Textile Learner GmbH. Available online: https://textilelearner.net/technical-properties-of-cotton-fiber/ (accessed on 26 December 2020).
195. Abidi, N.; Hequet, E.; Cabrales, L. Changes in sugar composition and cellulose content during the secondary cell wall biogenesis in cotton fibers. *Cellulose* **2010**, *17*, 153-160, doi:doi.org/10.1007/s10570-009-9364-3.
196. Calvo-Agudo, M.; Tooker, J.F.; Dicke, M.; Tena, A. Insecticide-contaminated honeydew: risks for beneficial insects. *Biological Reviews* **2021**, doi:10.1111/brv.12817.
197. Balasubramanya, R.; Bhatawdekar, S.; Paralikar, K. A new method for reducing the stickiness of cotton. *Textile Research Journal* **1985**, *55*, 227-232, doi:10.1177/004051758505500405.
198. Jumaniyazov, K.; Egamberdiev, F.; Abbazov, I. The Effect of Crop Type on Cotton Quality Indicators. In Proceedings of the International Journal of Advanced Research in Science, Engineering and Technology, 2020; pp. 13510-13518.
199. Severino, L.; Leite, B.; Gambarra-Neto, F.; Araújo, J.; Medeiros, E. *Detection and Quantification of Stickiness on Cotton Samples Using Near Infrared Hyperspectral Images* Bremen Baumwollboerse; p. 8.
200. Gamble, G.R. Evaluation of cotton stickiness via the thermochemical production of volatile compounds. *Journal of Cotton Science* **2003**, 45-50.
201. Was-Gubala, J.; Starczak, R. Nondestructive identification of dye mixtures in polyester and cotton fibers using Raman spectroscopy and ultraviolet-visible (UV-Vis) microspectrophotometry. *Applied spectroscopy* **2015**, *69*, 296-303, doi:10.1366/14-07567.

202. Fortier, C.A.; Rodgers, J.E.; Cintron, M.S.; Cui, X.; Foulk, J.A. Identification of cotton and cotton trash components by Fourier transform near-infrared spectroscopy. *Textile Research Journal* **2011**, *81*, 230-238, doi:10.1177/0040517510383620.
203. Miller, W.B.; Peralta, E.; Ellis, D.R.; Perkins Jr, H.H. Stickiness potential of individual insect honeydew carbohydrates on cotton lint. *Textile research journal* **1994**, *64*, 344-350, doi:10.1177/004051759406400606.
204. Barton, F.; Bargeron III, J.; Gamble, G.; McAlister, D.; Hequet, E. Analysis of sticky cotton by near-infrared spectroscopy. *Applied spectroscopy* **2005**, *59*, 1388-1392, doi:10.1366/000370205774783214.
205. Stefanakis, M.; Lorenz, A.; Bartsch, J.W.; Bassler, M.C.; Wagner, A.; Brecht, M.; Pagenstecher, A.; Schittenhelm, J.; Boldrini, B.; Hakelberg, S. Formalin Fixation as Tissue Preprocessing for Multimodal Optical Spectroscopy Using the Example of Human Brain Tumour Cross Sections. *Journal of Spectroscopy* **2021**, *2021*, doi:10.1155/2021/5598309.
206. van Kollenburg, G.; Bouman, R.; Offermans, T.; Gerretzen, J.; Buydens, L.; van Manen, H.-J.; Jansen, J. Process PLS: Incorporating substantive knowledge into the predictive modelling of multiblock, multistep, multidimensional and multicollinear process data. *Computers & Chemical Engineering* **2021**, *154*, 107466, doi:10.1016/j.compchemeng.2021.107466.
207. van Kollenburg, G.; Bouman, R.; Offermans, T.; Gerretzen, J.; Buydens, L.; van Manen, H.-J.; Jansen, J. Process PLS: Incorporating substantive knowledge into the predictive modelling of multiblock, multistep, multidimensional and multicollinear process data. *Computers Chemical Engineering* **2021**, *154*, 107466, doi:10.1016/j.compchemeng.2021.107466.
208. Victorita, B.; Marghitas, L.A.; Stanciu, O.; Laslo, L.; Dezmirean, D.; Bobis, O. High-performance liquid chromatographic analysis of sugars in Transylvanian honeydew honey. *Entomologia Experimentalis et Applicata* **2008**, *65*, 229-232, doi:10.1111/j.1570-7458.2006.00505.x.
209. Lottspeich, F.; Zorbas, H. *Bioanalytik*, 4 ed.; Spektrum, Akad. Verlag: Springer Spektrum Berlin, Heidelberg, 2022.
210. Wang, H.; Memon, H. *Cotton science and processing technology*; Springer Nature Singapore: China, 2021; Volume 5, p. 565.
211. DIN Deutsches Institut für Normung e. V., D.G.I.f.S. Textiles - Determination of cotton fibre stickiness - Part 2: Method using an automatic thermodetection plate device. 2004.
212. k Mansuri, A.; Somani, S.; Pathak, R. Effect of Waste Control on Yarn Parameters and Yield Improvement in Spinning Mill. *Journal of critical reviews* **2020**, *07*.
213. DIN Deutsches Institut für Normung e. V., D.G.I.f.S. Textiles - Determination of cotton fibre stickiness - Part 2: Method using an automatic thermodetection plate device; German version EN 14278-2:2004. **2004**, 11.
214. Bi, J.; Ballmer, G.; Hendrix, D.; Henneberry, T.; Toscano, N. Effect of cotton nitrogen fertilization on Bemisia argentifolii populations and honeydew production. *Entomologia Experimentalis et Applicata* **2001**, *99*, 25-36, doi:10.1046/j.1570-7458.2001.00798.x.
215. Fischer, M.K.; Völkl, W.; Hoffmann, K.H. Honeydew production and honeydew sugar composition of polyphagous black bean aphid, Aphis fabae (Hemiptera: Aphididae) on various host plants and implications for ant-attendance. *European Journal of Entomology* **2005**, *102*, 155-160, doi:10.14411/eje.2005.025.
216. Hogervorst, P.A.; Wäckers, F.L.; Romeis, J. Effects of honeydew sugar composition on the longevity of Aphidius ervi. *Entomologia Experimentalis et Applicata* **2007**, *122*, 223-232, doi:10.1111/j.1570-7458.2006.00505.x.
217. Lottspeich, F.; Zorbas, H. *Bioanalytik*, 4th ed.; Spektrum, Akad. Verlag, Berlin/Heidelberg, Germany, 2022.

218. Janshekar, H.; Brown, C.; Fiechter, A. Determination of biodegraded lignin by ultraviolet spectrophotometry. *Analytica Chimica Acta* **1981**, *130*, 81-91, doi:10.1016/S0003-2670(01)84153-2.
219. Sadeghifar, H.; Venditti, R.; Jur, J.; Gorga, R.E.; Pawlak, J.J. Cellulose-lignin biodegradable and flexible UV protection film. *ACS Sustainable Chemistry Engineering* **2017**, *5*, 625-631, doi:10.1021/acssuschemeng.6b02003.
220. Mach, H.; Volkin, D.B.; Burke, C.J.; Russell Middaugh, C. Ultraviolet absorption spectroscopy. *Protein Stability Folding* **1995**, 91-114, doi:10.1385/0-89603-301-5:91.
221. Lottspeich, F.; Zorbas, H. *Bioanalytik*, 4th ed.; Spektrum, Akad. Verlag: Berlin/Heidelberg, Germany, 1998.
222. Al Ktash, M.; Haular, O.; Ostertag, E.; Brecht, M. UV-Vis/NIR spectroscopy and hyperspectral imaging to study the different types of raw cotton. *Journal spectral imaging* **2020**, 14.

10 Appendix

10.1 Supplementary Information for Paper II: UV Hyperspectral Imaging as Process Analytical Tool for the Characterization of Oxide Layers and Copper States on Direct Bonded Copper

Diffuse reflectance spectra of the copper powders were recorded in the range of 200 nm – 380 nm using a commercial UV spectrometer (Lambda 1050+, PerkinElmer, Inc., Waltham, MA, USA). The spectrometer was equipped with a 150 mm Spectralon® integrating sphere to acquire data in reflection mode with an R6872-Photomultiplier (PMT). A deuterium lamp was used as light source in the spectrometer. A 10 mm quartz SUPRASIL® cuvette (QS, 100-10-40, Hellma, Müllheim, Germany) was used for measuring the copper powder see Table 10.1. The filled cuvette was placed at the reflectance port of the integrating sphere. The port measuring area is approximately 4.9 cm².

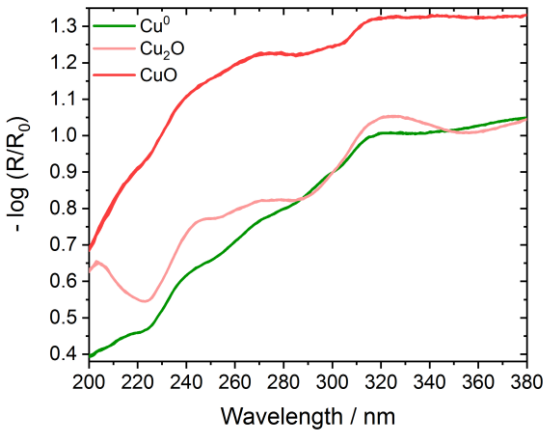


Figure 10.1: Reference spectra for the copper Cu^0, Cu_2O and CuO by using UV spectrometer).

Table 10.1: Description of the direct bonded copper substrates and their sample preparation.

Sample type	Description	Manufacturer	Article Number
Cu	Copper, powder, electrolytically produced	Merck KGaA, Darmstadt, Germany	2715
Cu2O	Copper (I) oxide powder, red	Riedel-de Haën AG, Seelze, Germany	12841
CuO	Copper (II) oxide powder, heavy, powder, technical	Riedel-de Haën AG, Seelze, Germany	12867

10.2 Supplementary Information for Paper IV: Prediction of Honeydew Contaminations on Cotton Samples by In-Line UV Hyperspectral Imaging

10.2.1 Cotton sample preparation

All details and measured values of the reference sample set are given in Table 10.2.

Table 10.2: Description of the direct bonded copper substrates and their sample preparation.

Sample type	Sugar concentration / wt	Cotton weight / g (± 0.0001 g)	Cotton weight / g after dried at 30 °C, 8 h (± 0.0001 g)	Humidity / % after dried at 30 °C, 8 h	Soaked cotton weight / g after dried at 30 °C, 44 h (± 0.0001 g)	Humidity / % after dried at 30 °C, 44 h	Amount of sugar on cotton / g	Ratio of sugar / g dried cotton / g
A1	2	0.3	0.2879		53.7	0.3734	50.0	0.0855
A2	2	0.3	0.2925		53.5	0.3643	50.0	0.0718
A3	2	0.3	0.2914		53.3	0.3600	50.0	0.0686
B1	1	0.3	0.2917		54.2	0.3295	50.0	0.0378
B2	1	0.3	0.2931		54.7	0.3291	50.0	0.0360
B3	1	0.3	0.2913		54.2	0.3341	50.0	0.0428
C1	0.5	0.3	0.2916		55.2	0.3142	50.0	0.0226
C2	0.5	0.3	0.2929		55.6	0.3137	50.0	0.0208
C3	0.5	0.3	0.2992		55.7	0.3215	50.0	0.0223
D1	0.25	0.3	0.2925		56.0	0.3050	51.0	0.0125
D2	0.25	0.3	0.2928		56.2	0.3030	51.0	0.0102
D3	0.25	0.3	0.2921		56.2	0.3033	51.0	0.0112
E1	0.125	0.3	0.2908		56.5	0.2979	50.5	0.0071
E2	0.125	0.3	0.2819		56.9	0.2915	50.0	0.0096
E3	0.125	0.3	0.2899		57.2	0.3013	49.9	0.0114
F1	0.0625	0.3	0.2903		57.4	0.2997	50.7	0.0094
F2	0.0625	0.3	0.2899		57.6	0.2990	49.9	0.0091
F3	0.0625	0.3	0.2823		57.7	0.2915	49.7	0.0093
CLN1	0	0.3	0.2984		57.7	-	-	-
CLN2	0	0.3	0.2991		57.9	-	-	-
CLN3	0	0.3	0.2894		58.1	-	-	-

10.2.2 Additional figures of the principal component analysis of the sugar cotton samples

For the PCA model of the cotton samples with different concentrations of sugar four PCs are necessary. The variance on PC3 is not necessary to distinguish between different sugar concentrations. The information on PC3 might be related to the morphology of the fiber itself. PC1 against PC2, PC3 and PC4 are shown respectively in Figure 10.2 to complement the PCA sugar model.

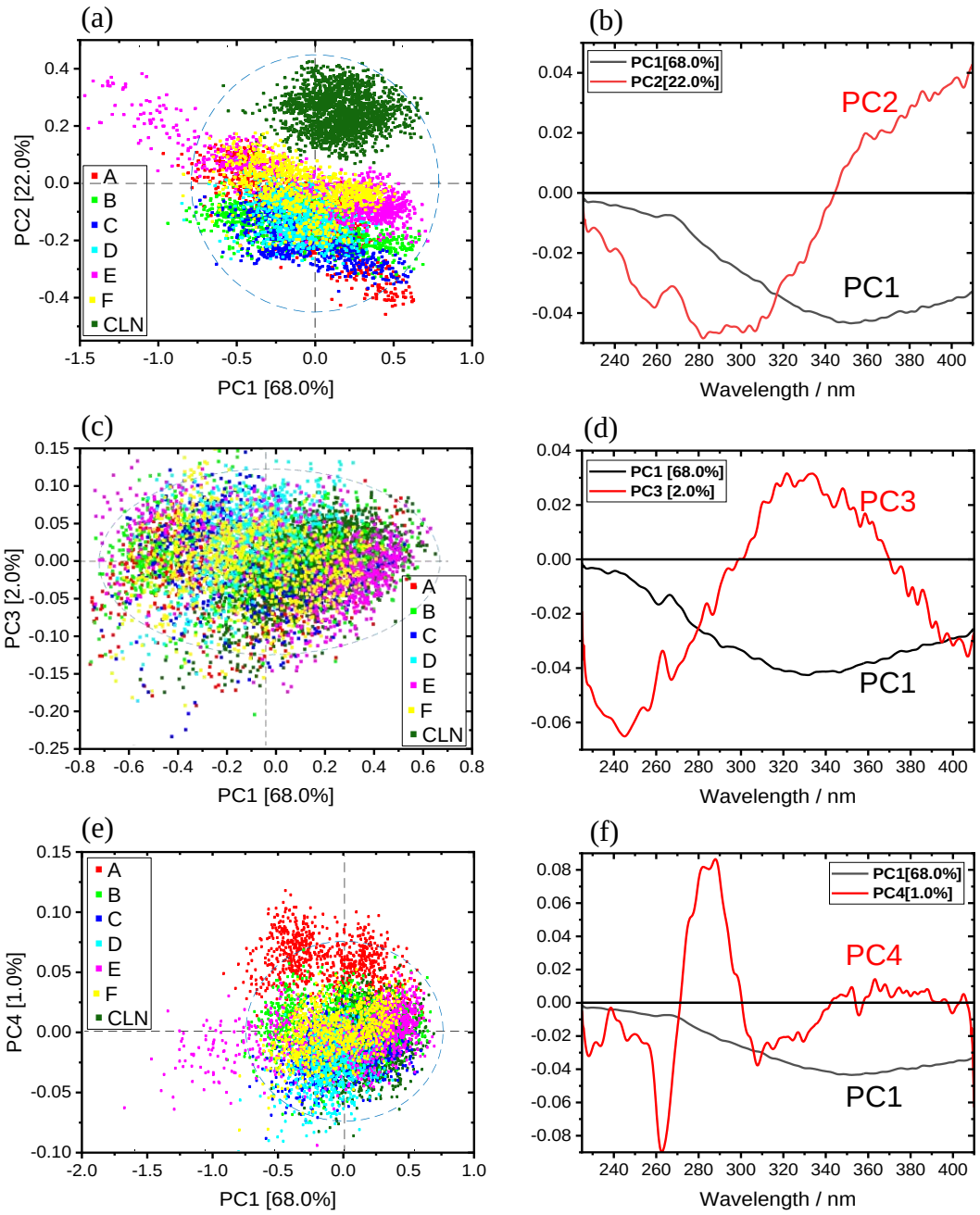


Figure 10.2: PCA sugar model for the cotton samples with (a,c,e) scores and (b,d,f) corresponding loadings (PC1 black, PC2, PC3 and PC4 red).

10.2.3 Pure dried protein spectrum

Protein spectra were acquired to identify the information in the range of 250 nm to 280 nm. The protein was solved in distilled water and the solution was dropped on a piece of PTFE. Afterwards, the sample was dried in a vacuum oven (see 6.3.2). Data was acquired with the hyperspectral imaging setup with the settings mentioned in 2.3 and 2.4. This experiment was necessary to

verify the spectral range between 250 nm and 280 nm contains true information and is not an artifact due to the efficiency of the detector and the weak intensity light source in the UV range [2].

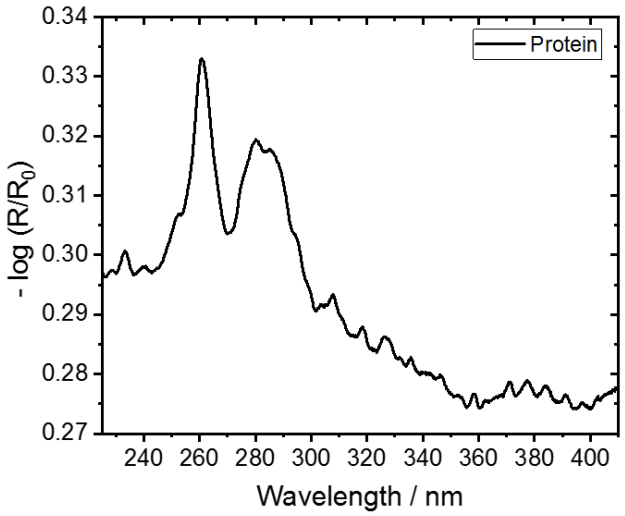
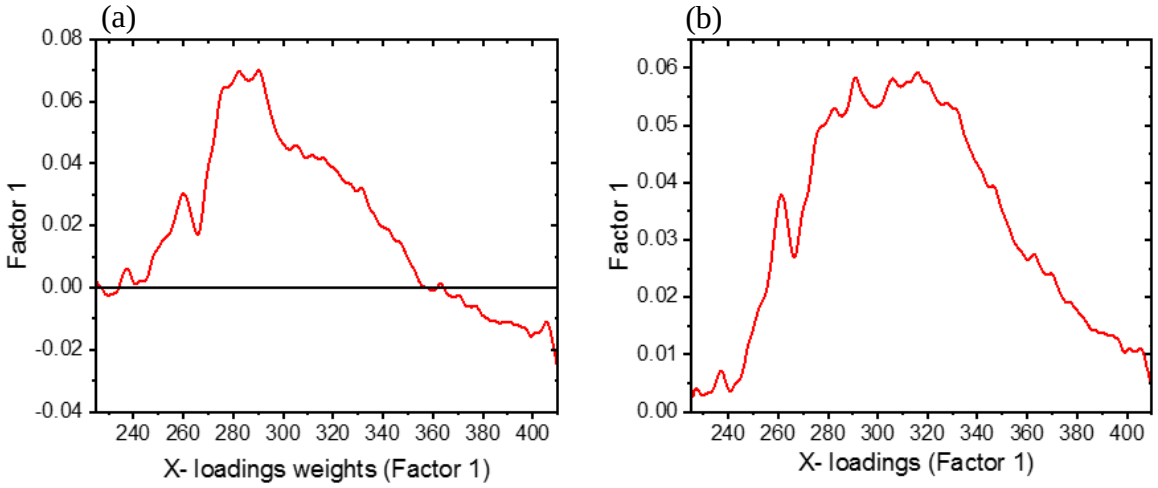
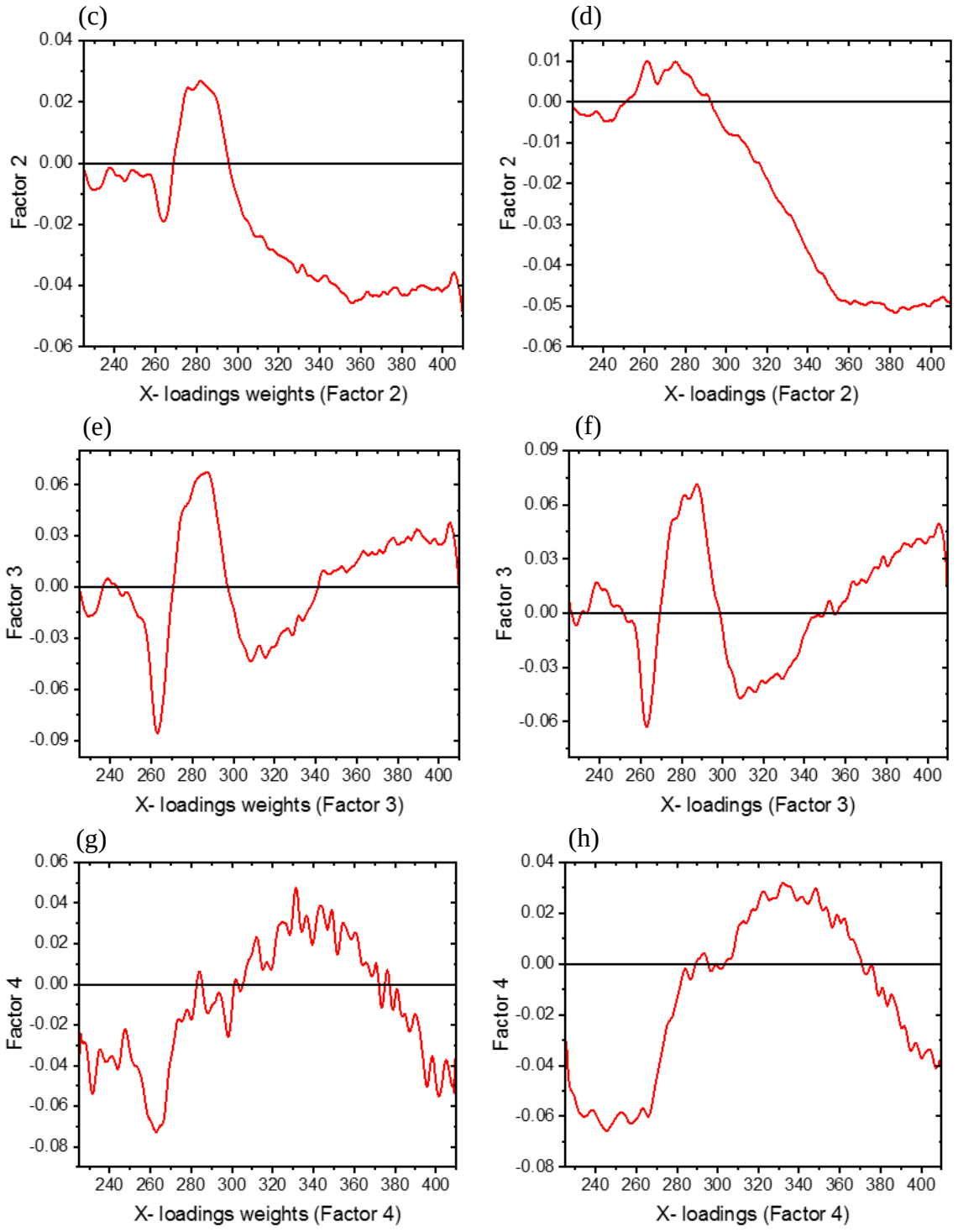


Figure 10.3: Mean spectrum of pure dried protein on PTFE.

10.2.4 X-loadings weights and x-loadings of the PLS-R model

For model building and understanding the PLS-R factor loadings and loading weights for all five factors are displayed in the Figure 10.4.




(c)
Factor 2
X- loadings weights (Factor 2)
(d)
Factor 2
X- loadings (Factor 2)
(e)
Factor 3
X- loadings weights (Factor 3)
(f)
Factor 3
X- loadings (Factor 3)
(g)
Factor 4
X- loadings weights (Factor 4)
(h)
Factor 4
X- loadings (Factor 4)

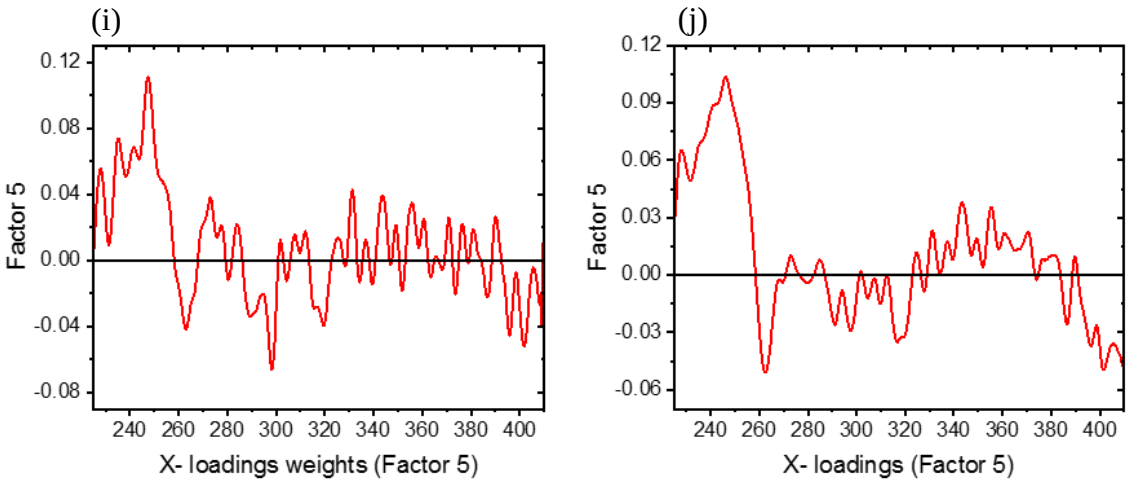
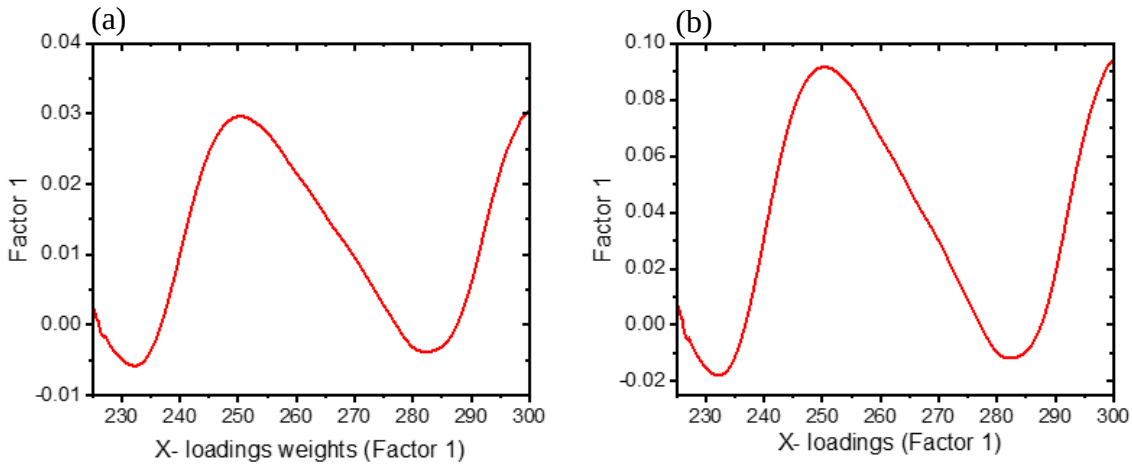


Figure 10.4: X-loadings weights and x-loadings for factor 1 (a, b), factor 2 (c, d), factor 3 (e, f), factor 4 (g, h) and factor 5 (i, j), respectively.

10.3 Supplementary Information for Paper V: Applying UV Hyperspectral Imaging for Quantification of Honeydew Content on Raw Cotton via PCA and PLS-R Models

For model building and understanding the PLS-R factor loadings and loading weights for all three factors are displayed in the Figure 10.5. X-loadings reflect the relationship between the predictor variables and the latent variables, while X-loadings weights represent the relationship between the predictor variables and the response variable. Both X-loadings and X-loadings weights play important roles in PLS-R for variable selection, model interpretation, and prediction.



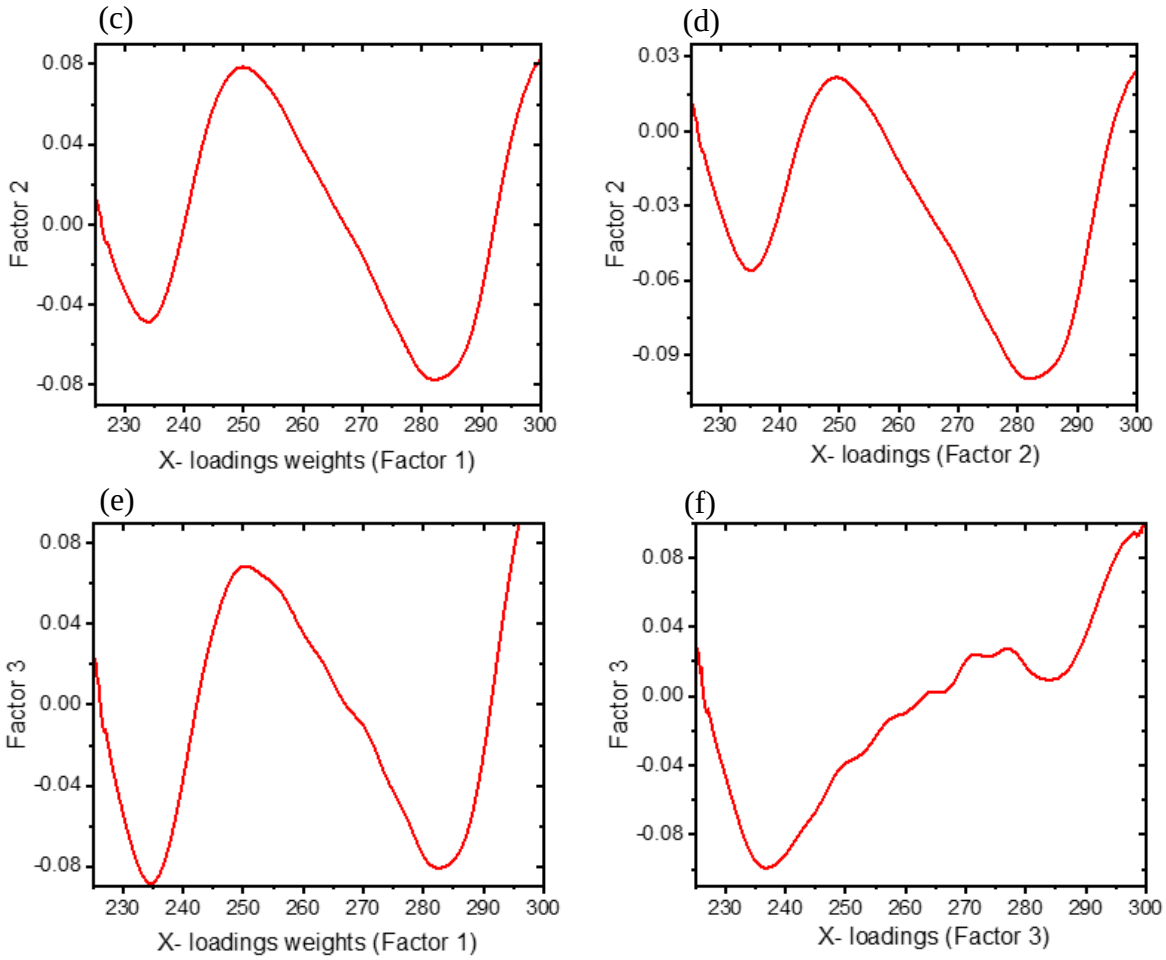


Figure 10.5: X-loadings weights and x-loadings for factor 1 (a, b), factor 2 (c, d) and factor 3 (e, f), respectively.

List of Abbreviations and Symbols

Abbreviation	Concept
EM	Electromagnetic
UV-Vis	Ultraviolet-Visible
NIR	Near Infrared
NMR	Nuclear Magnetic Resonance
MS	Mass Spectroscopy
HSI	Hyperspectral Imaging
MVA	Multivariate Data Analysis
PCA	Principal Component Analysis
PCs	Principal Components
PLS-R	Partial Least Square Regression
QDA	Quadratic Discriminant Analysis
R^2	High coefficient of determination
RMSEC	Root Mean Square Error of Calibration
RMSECV	Root Mean Square Error of Calibration and Cross-Validation
R^2_C	High Coefficient of Calibration
R^2_{CV}	High Coefficient of Cross-Validation
n	Sample number
PAT	Process Analysis and Technology
2D	Two Dimensional
3D	Three Dimensional
CCD	Charge Couple Device
CMOS	Complementary Metal Oxide Semiconductor
RGB	Red Green Blue
PTFE	Polytetrafluoroethylene
InGaAs	Indium Gallium Arsenide
HPLC	High-Performance Liquid Chromatography
MS	Mass Spectrometry
PMT	Photomultiplier
XBO	Xenon-arc lamp
HBO	Mercury-arc lamp
LEDs	Light Emitting Diodes
FDA	Food and Drug Administration
API	Active Pharmaceutical Ingredients
RoB	Organic raw cotton
HC	Hemp plant
RcO	Recycled cotton
RoSt	Standard raw cotton
RcBH	Recycled organic bright cotton
CLN	Cleaned Cotton
IBU_{pure}	Ibuprofen, >98%, API
ASA	Acetylsalicylic Acid, 99%, API
PAR_{pure}	Paracetamol, 99%, API
THO	Thomapyrin
IBU_{ratio}	Ibuprofen, Ratiopharm GmbH
IBU_{beTa}	Ibuprofen, Betapharm
ASP_{BAYER}	Acetylsalicylic Acid, Bayer vital GmbH

PAR_{ratio}	Paracetamol, Ratiopharm GmbH
DBC	Direct Bonded Copper
GLS	Generalized Least Squares
IC	Integrated Circuit
XPS	X-ray Photoelectron Spectroscopy
USDA	United States Department of Agriculture
HVI	High Volume Instrument
AFIS	Advanced Fiber Information System
FQI	Fibre Quality Index
SCI	Spinning Consistency Index
PDI	Premium-Discount Index

List of Figures

List of Tables

List of Publications

Published Manuscripts

1. **Al Ktash, M.;** Stefanakis, M. et al. Characterization of Pharmaceutical Tablets Using UV Hyperspectral Imaging as a Rapid In-Line Analysis Tool. *Sensors* **2021**, *21*,4436. https://doi.org/10.3390/s21134436.

2. **Al Ktash, M.;** Stefanakis, M. et al. UV Hyperspectral Imaging as Process Analytical Tool for the Characterization of Oxide Layers and Copper States on Direct Bonded Copper. *Sensors* **2021**, *21*, 7332. https://doi.org/10.3390/s21217332.

3. **Al Ktash, M.;** Hauler, O. et al. Ultraviolet-visible/near infrared spectroscopy and hyperspectral imaging to study the different types of raw cotton. *Journal of Spectral Imaging* **2020**; 9: a18. http://dx.doi.org/10.1255/jsi.2020.a18

4. **Al Ktash, M.;** Stefanakis, M.; et al. Prediction of Honeydew Contaminations on Cotton Samples by In-Line UV Hyperspectral Imaging. *Sensors* **2023**, 23, 319. https://doi.org/10.3390/s23010319**.**

5. Knoblich, M.; **Al Ktash, M**.; et al. Applying UV Hyperspectral Imaging for the Quantification of Honeydew Content on Raw Cotton via PCA and PLS-R Models. *Textiles* **2023**, 23, 287-293. https://doi.org/10.3390/textiles3030019.

Submitted Manuscript

6. Knoblich, M.; **Al Ktash, M.** et al. Rapid Detection of Cleanliness on Direct Bonded Copper Substrate Using UV Hyperspectral Imaging.

Declaration of Contribution

Description of the meaning of the personal contribution according to German law § 6 Abs. 2 sentences 3 of the doctoral regulations

No.	Published / submitted	Position of the candidate in the list of authors	Number of authors	Scientific ideas by the candidate (%)	Data generation by the candidate (%)	Interpretation and analysis by the candidate (%)	Paper writing by the candidate (%)
1	published	1	5	60	80	70	60
2*	published	1	11	45	80	45	45
3	published	1	4	40	80	70	60
4*	published	1	7	45	90	40	40
5*	published	1	6	45	50	30	40
6*	submitted	1	11	45	50	30	40

*Shared first author

Acknowledgments

There are a number of people without whom this thesis might not have been originated and finished, and to whom I am very grateful. There are also a number of people without whom the life of my family during these years in Germany had been very difficult, and to whom I am greatly indebted.

To:

Prof. Dr. Marc Brecht, you provided excellent supervision and assistance during my research work. Your gentlemanly conduct impressed me greatly, and your constructive criticism and advice were always delivered politely. You instilled confidence in me and transferred positive energy whenever I needed it. You have been a wonderful academic advisor, and I will be forever indebted to you. Please continue to be the same amazing person that you are.!

Porf. Dr. Alfred Meixner, thank you for your support during my time in Germany. Your expertise and ideas greatly contributed to the success of my study, and your impact on my scientific and personal life has been wonderful. Thank you very much!

Prof. Dr. Hermann Mayer and ***Dr. Erik Schäffer*** for agreeing to participate in this work and for attending as dissertation examiners.

Prof. Dr. Karsten Rebner, who push me in the best way and help me in all steps to get my scholarship for my doctoral studies. I am grateful for your support.

Dr. Edwin Ostertag, I would like to express my gratitude for my continuous support during my Ph.D. studies and related research. I am thankful for patience, motivation, and immense knowledge. His guidance has been invaluable to me throughout the research and writing of this thesis.

Prof. Dr. Volker Jehle, who provided invaluable assistance in supplying the necessary samples and engaging in a productive discussion that greatly aided my understanding.

My Friends Mona Knobblich, Tim Bäuerle, Tobias Drieschner, Frank Wackenhut, Barbara Blodrini, Miriam Bassler, Julia Steinbach, Alexandra Wagner and Ashatoush Mukherjee.

I would like to thank for sharing their time with me, for their support during all these years in Germany, and for always being glad to help or simply listen.

Reutlingen University, the university where I worked in Germany, allowed me to qualify and economically supported me all those years.

KAAD, the Katholischer Akademischer Ausländer-Dienst that gave me a scholarship for my doctoral studies, had an extraordinary fulfillment!

Lastly, I would like to express my gratitude to my family, including my parents ***Mahmoud and Nawla,*** and to my lovely wife ***Bayan*** and my son ***Osama*** my brother ***Alaa*** and sisters ***Madleen, Nivin, Nisrein, and Walaa*** for supporting me spiritually throughout this work and my life in general. God save and bless them.