processes

MDPI

Article

Development of a General PAT Strategy for Online Monitoring of Complex Mixtures—On the Example of Natural Product Extracts from Bearberry Leaf (*Arctostaphylos uva-ursi*)

Christoph Jensch, Larissa Knierim, Martin Tegtmeier and Jochen Strube *

Institute for Separation and Process Technology, Clausthal University of Technology, Leibnizstr. 15, 38678 Clausthal-Zellerfeld, Germany; jensch@itv.tu-clausthal.de (C.J.); larissa.knierim@tu-clausthal.de (L.K.); tegtmeier@itv.tu-clausthal.de (M.T.)

* Correspondence: strube@itv.tu-clausthal.de; Tel.: +49-5323-72-2872

Abstract: For the first time, a universally applicable and methodical approach from characterization to a PAT concept for complex mixtures is conducted—exemplified on natural products extraction processes. Bearberry leaf (*Arctostaphylos uva-ursi*) extract is chosen as an example of a typical complex mixture of natural plant origin and generalizable in its composition. Within the quality by design (QbD) based process development the development and implementation of a concept for process analytical technology (PAT), a key enabling technology, is the next necessary step in risk and quality-based process development and operation. To obtain and provide an overview of the broad field of PAT, the development process is shown on the example of a complex multi-component plant extract. This study researches the potential of different process analytical technologies for online monitoring of different component groups and classifies their possible applications within the framework of a QbD-based process. Offline and online analytics are established on the basis of two extraction runs. Based on this data set, PLS models are created for the spectral data, and correlations are conducted for univariate data. In a third run, the prediction potential is researched. Conclusively, the results of this study are arranged in the concept of a holistic quality and risk-based process design and operation concept.

Keywords: quality by design (QbD); process analytical technology (PAT); chemometrics; multivariate data analysis; solid-liquid extraction (SLE); pressurized hot water extraction (PHWE); phytoextraction; complex mixtures; Raman spectroscopy; attenuated total reflection Fourier transformed infrared spectroscopy (ATR-FTIR); fluorescence; diode array detector

Citation: Jensch, C.; Knierim, L.; Tegtmeier, M.; Strube, J. Development of a General PAT Strategy for Online Monitoring of Complex Mixtures—On the Example of Natural Product Extracts from Bearberry Leaf (*Arctostaphylos uva-ursi*). *Processes* **2021**, *9*, 2129. https://doi.org/10.3390/pr9122129

Academic Editor: Ibrahim M. Abu-Reidah

Received: 2 November 2021
Accepted: 19 November 2021
Published: 25 November 2021

1. Introduction

The quality by design (QbD) approach on process development has its background in the pharmaceutical industry. However, the core aspects of this process design method, process robustness, and process understanding paired with cleverly implemented digital twins and process analytical technologies (PAT) leading towards advanced process control (APC) based on model predictive control and process status evaluation, can be beneficial for every field of process industry.

Products based on renewable resources, such as plants, represent a growing market and the associated industry is an important supplier of versatile products. The applications range from pharmaceutical products, the food, health, and nutrition sector as well as plant protection for ecological farming or construction material, basic chemicals, and energy resources [1–7]

In the plant processing industry, especially in regulated industries, traditional extraction processes are widely used. The regulatory environment sets the framework, as traditional approvals are tied to manufacturing processes, which limits the possibilities for further development of the processes or optimization of operations. The associated dependence on traditional manufacturing processes can lead to the fact that raw materials are

not optimally utilized through the selection of suboptimal process parameters. Traditional extraction processes can furthermore lead to high process variabilities and bad process stability [8]. Innovative methods such as Quality by Design including Process Analytical Technology offer opportunities for manufacturers to meet regulatory requirements such as low product variability between production batches [9,10].

In a quality by design-based process development, a design space is defined in which the variation of critical quality attributes (CQAs) of the product are linked to process parameters [11]. The ICH guidelines for quality by design offer a logical approach in implementing the process design concept into process development. Based on these guidelines the flowsheet presented in Figure 1 illustrates the different steps [11–13]. In previous studies, the variability of different batches of plant material as well as QbD-based process design for plant-based processes are shown and researched following the flowsheet in Figure 1 dealing with risk assessment and the creation of design spaces via experimental or in silico DoEs using validated physico-chemical models which are further developed into digital twins [9,10,14–18].

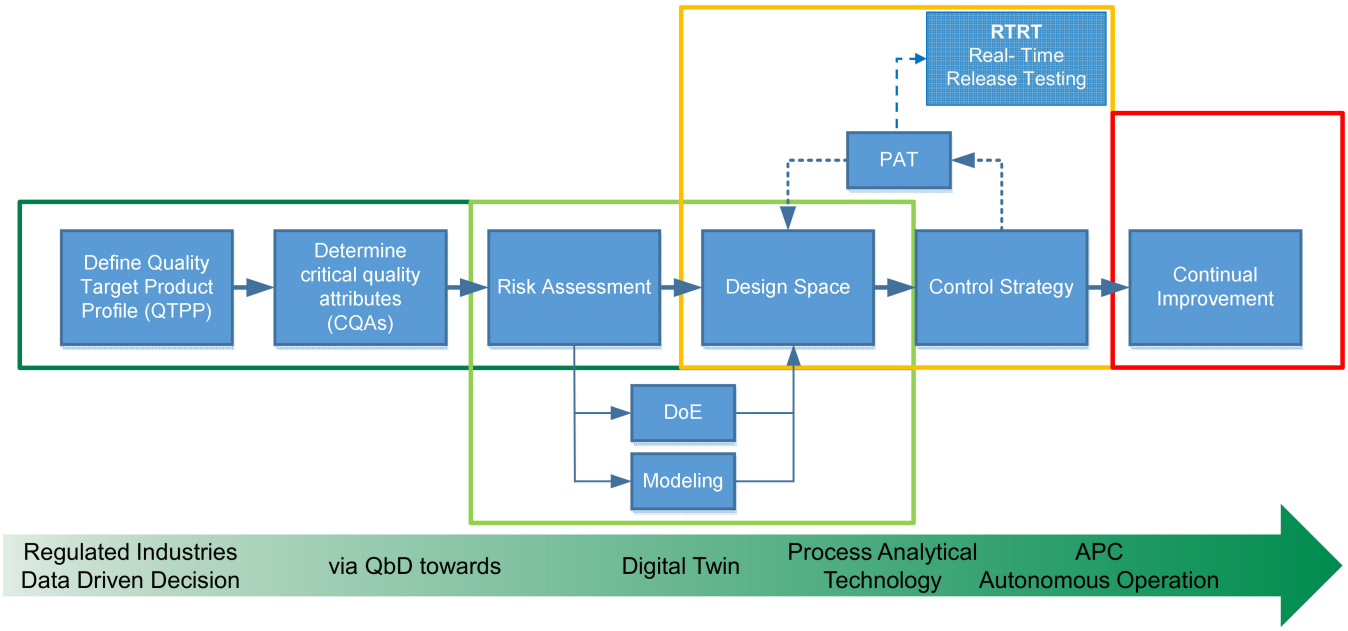


Figure 1. Overview of QbD and PAT concept and implementation, extended from [18].

The development of a PAT concept is the next logical step in researching process design and development applicating the QbD approach [18,19]. PAT can be used to measure the CQAs and the current value of the process parameters [20]. Paired with substance parameters and based on the signals of the PAT the extraction profile and process status can be evaluated and predicted by a digital twin and compared with the required process status defined by the design space [9,18]. Based on this comparison, the process parameters can be adjusted if the CQAs would otherwise drop out of the design space. This QbD-based approach for solid-liquid extraction process operation leads to greater flexibility and more possibilities to achieve constant high product qualities compared to stiff operation systems defined by a recipe. PAT is therefore a key enabling technology for a variety of benefits in process and product quality control. When using PAT for quality assurance purposes as well as for advanced process control (APC), it is demanded to differentiate between two PAT circuits. One of each for the respective field of action. For example, a combination of FTIR, DAD, and pH could be used for QA-measurement leading towards real-time release testing (RTRT) and Raman with Conductivity could be used for APC, by monitoring the CQAs in real time.

For the development of the PAT concept, a systematic approach is chosen to guaranty a broad field of view. In the first step, the complex mixture is broken down into seven substance groups with subgroups each. The substance groups are chosen based on chemical equality, such as in the case of organic acids, regulatory aspects for plant extracts, such as ash and secondary metabolites and established component groups such as proteins, matrix components/structural carbohydrates, and fat [21–23]. Dry residue is the sum of all non-volatile components in the extract and consists of the described component groups. Afterwards, offline analytics has to be available to detect each of these component groups in the extracts to generate a reference data set for the training of PLS models. The respective methods for the offline analytic in described in depth in Section 2. Based on the literature and the possible interaction of each component group with the respective spectroscopic or sensoric measurement method, the chosen measuring unit is assigned to the respective component group [24–33]. The overview is given in Figure 2.

This method is conducted on a pressurized hot water extraction (PHWE) of bearberry leaf (*Arctostaphylos uva-ursi*). The PHWE is a solid-liquid extraction process, in which water under subcritical conditions (100 °C, 0.1 MPa to 374 °C, 22.1 MPa) is utilized as a solvent. Due to the variable solubility properties of water in these ranges of temperature and pressure, water is able to extract components, which are usually only soluble in organic solvents such as ethanol or acetone [34,35]. A priori studies have proven bearberry leaf extract as an example is a typical complex mixture of natural plant origin and generalizable in its composition [15,36–38].

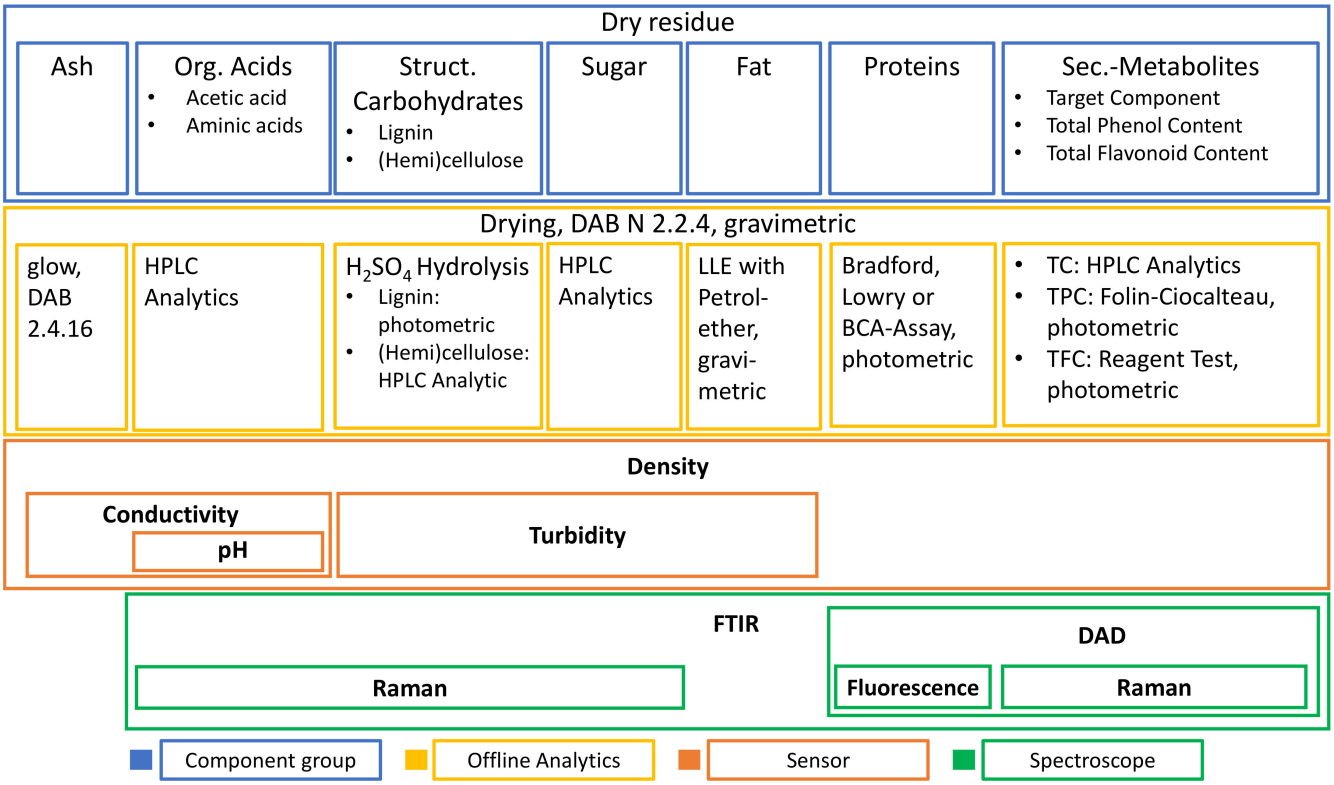


Figure 2. Overview of systematic approach on development of a PAT concept for complex mixtures with postulated sensor and spectroscopy concept.

2. Materials and Methods

2.1. Extraction Setup

The extraction was carried out as a pressurized hot water/liquid hot water/subcritical water extraction (PHWE). The extraction plant consists of an extraction column with a volume of 0.1 L, solvent vessel, heating unit with a heat exchanger, cooling unit with a heat exchanger, and an extract vessel. Extraction conditions were 140 °C at 1 L/h. The extraction column was filled with 30 g of ground plant material and 12 fractions with 40 mL each were collected. The third extraction run was carried out at 1.3 L/h. The extraction solvent was deionized water from an in-house deionization plant.

2.2. Dry Residue Determination

Dry residues of the extracts were determined after the method described in the European pharmacopoeia (2.8.16 dry residue of extracts). An amount of 2 g per sample was dried in glass vials at 105 °C for 2h, cooled down under a dry atmosphere and the residual mass was determined gravimetrically.

2.3. Target Component Analytic for Arbutin

RP-HPLC is used for the analysis of arbutin. The buffers used are water and methanol in a gradient mode, the column is a PharmPrep P 100 RP 18e (10 µm) by Merck KGaA, Darmstadt, Germany. For detection, a DAD is used, which detects arbutin at 280 nm. Methanol in HPLC grade is bought from VWR International GmbH, 30,163 Hannover, Germany. For calibration, arbutin standards of the concentrations 5 g/L, 1.6 g/L, 0.6 g/L, and 0.2 g/L from Sigma-Aldrich, St. Louis, MO, USA are used.

2.4. Hydrolysis of Structural Carbohydrates

Structural carbohydrates are the sum of (hemi-) celluloses and lignin. Lignin consists of acid-soluble lignin and acid-insoluble lignin. The determination of structural carbohydrates was conducted according to the Laboratory Analytical Procedure by the National Renewable Energy Laboratory [23] and consists of a two-step hydrolysis with sulfuric acid. The 72% sulfuric acid was ordered at VWR International GmbH, 30,163 Hannover, Germany. In the first step, 1.5 mL of 72% sulfuric acid is added to the dry extract in the centrifuge vial and heated to 30 °C for one hour in a water bath with stirring. The n, 42 mL of HPLC water is added, diluting the sulfuric acid to 4%. The second stage of hydrolysis takes place in the autoclave. The subsequent autoclaving process was chosen to maintain the temperature of 120 °C and the pressure of 2 bar for one hour. Following hydrolysis, the samples are filtered to separate the acid-insoluble lignin (retentate) from the acid-soluble lignin and structural carbohydrates (permeate). The filter paper (pore size 4–12 µm) is previously dried for one hour at 80 °C and then the weight is determined. The filter cake is rinsed with 50 mL of HPLC water after the sample has been completely transferred, and the filter paper with the filter cake is dried again at 80 °C for one hour, after which the weight is determined. The permeate is used for further analysis for acid-soluble lignin and structural carbohydrates. Acid soluble lignin is determined by UV/Vis spectroscopy by determining the adsorption of the sample in a glass cuvette at 215 nm and 280 nm. The structural carbohydrates are determined using the same HPLC method as the sugars in Section 2.5 since these are also present as free sugars in the solution after hydrolysis. The free sugars before hydrolysis are subsequently subtracted from the free sugars from the structural carbohydrate determination so that these are not evaluated twice.

2.5. Sugar and Acetic Acid Analytics

RP-HPLC is used to determine the various sugars and acetic acid [23]. Here, a 0.005 M sulfuric acid solution is used as buffer solution in an isocratic method and a refractive index detector is used for detection. The column is a Biorad Aminex HPX-87P. Glucose, arabinose, xylose, mannose, galactose, and cellobiose at concentrations of 0.1 g/L, 1 g/L,

2 g/L, and 4 g/L, and acetic acid at concentrations of 0.15 g/L, 0.2 g/L, 0.3 g/L, and 0.5 g/L are used as calibration standards.

2.6. TPC and TFC Determination

The phenol content of the samples is determined by UV/Vis spectroscopy using the Folin-Ciocalteu reagent. Gallic acid solutions of concentrations 0.05 g/L, 0.1 g/L, 0.2 g/L, and 0.25 g/L are used as calibration solutions. To prepare the calibration lines, 0.5 mL of gallic acid solution is mixed with 1.5 mL of Folin-Ciocalteu reagent, which is diluted to 10% of the original concentration beforehand and incubated for 5 min at room temperature. The n, 1.5 mL of 7% sodium carbonate solution is added, everything is made up to 10 mL with HPLC water and incubated for 1.5 h. This is followed by a triplicate measurement at 750 nm. HPLC water is used as a blank sample instead of the gallic acid standards. The Folin-Ciocalteu reagent was supplied by VWR International, Hannover, Germany.

The aluminum chloride method is used to determine the flavonoid content. This is UV/Vis spectroscopy. Quercetin in the concentrations 0.25 g/L, 1 g/L, 2 g/L, 4 g/L, 6 g/L, and 10 g/L is used as the equivalent for calibration. For the determination, 1 mL of the respective calibration solution is added to a centrifuge vial (15 mL) with 4 mL of HPLC water and 0.3 mL of 5% sodium nitrite solution is added. After standing at room temperature for 5 min, 0.3 mL of 10% aluminum chloride solution is added. After another 6 min, 1 mL of a 1 M sodium hydroxide solution is added and made up to 10 mL with HPLC water. The absorbance of the sample is determined at 510 nm as a triplicate measurement. A sample in which HPLC water is added instead of the calibration solution serves as a blank. When determining the flavonoid content of the extracts, they are diluted to 0.25 vol.% with HPLC water beforehand. Reagents are supplied by VWR International, Hannover, Germany.

2.7. Aminic Acids

RP-HPLC is used to analyze the free amino acids. The following buffers are required for this purpose: Buffer A, consisting of sodium hydrogen phosphate anhydrate at a concentration of 1.4 g/L (10 mM) and sodium tetraborate at a concentration of 3.8 g/L (10 mM), whose pH is adjusted to 8.2 with hydrochloric acid and then filtered with a filter of pore size 0.45 µm. Buffer B is composed of a mixture of acetonitrile and water in a 70/30 ratio. Double the amount of buffer A should be prepared, as this is consumed more quickly than buffer B.

For the derivatization reaction, a sodium borate buffer with a concentration of 8.24 g/L boric acid is required, which is adjusted to a pH of 10.2 with sodium hydroxide solution, as well as the OPA reagent, which is prepared directly in the HPLC vial. The latter consists of 10 mg o-phthalic dialdehyde, 6.5 µL mercaptopropionic acid, 500 µL methanol, and 500 µL of the borate buffer. An injection diluent is also required, which is composed of 100 mL of buffer A and 0.4 mL of phosphoric acid.

For calibration, a ready-made standard from Sigma-Aldrich, St. Louis, MO, USAis used, but it does not contain the amino acids glutamine and asparagine, so these are prepared separately. For this purpose, 50 mL each are prepared with a concentration of 50 mM in 0.1 M hydrochloric acid. The standards for calibration are prepared directly in 300 µL HPLC vials to save material. First, a standard with a concentration of 1.8 mM is prepared by mixing 31.2 µL of injection diluent, 108 µL of Sigma standard, 5.4 µL of glutamine standard, and 5.4 µL of asparagine standard. The concentrations are 0.9 mM, 0.45 mM, 0.225 mM, 0.1125 mM, 0.0563 mM, 0.0281 mM, and 0.01405 mM. An Agilent Poroshell HPH C18 3.0 × 100, 2.7 µm column is used. A DAD is used to detect the different amino acids.

2.8. Protein Determination

The protein content of the extracts was determined by UV/Vis spectroscopy using a Bradford assay. A BSA standard is used for calibration. Calibration solutions are analyzed

by adding 1.5 mL of Coomassie Plus reagent from Thermo Scientific to 0.05 mL of the calibration solution in a UV/Vis cuvette and then incubating the solution for 10 min at room temperature. During the subsequent measurement, make sure that each sample stands for exactly 10 min, otherwise, the staining will have a different intensity. The measurement is then performed at a wavelength of 595 nm and repeated three times. The assay is from Thermo Fischer Scientific, Osterode am Harz, Germany.

2.9. Fat Determination

The fat determination is conducted by extracting the fats with petrol ether from the liquid extract. After centrifugation for phase separation, the organic phase is transferred into a glass vessel for drying. The fat content is determined gravimetrically.

2.10. Ash Determination

For the analysis of the inorganic components, the method of the National Renewable Energy Laboratory for the determination of these components in biomass is modified [23].

To determine the inorganic components of the liquid extract, the required amount of porcelain crucibles is first annealed in a muffle furnace at 575 °C for 4 h, then cooled in a desiccator and the empty weight is determined. The n, 2 g of liquid extract per crucible are weighed out and this is dried at 40 °C under vacuum for 48 h in the drying oven. The subsequent ashing in the muffle furnace is carried out in three steps. First, it is heated from room temperature to 105 °C and the temperature is maintained for 12 min. The n, within 15 min, the temperature is raised to 250 °C (approx. 10 °C/min) and held for 30 min. In the final step, the temperature is heated to 575 °C (approx. 20 °C/min) within 16 min and held at this temperature for 3 h. After the time has elapsed, the muffle furnace is left closed for one hour to cool down. The crucibles are then removed and placed in the desiccator, where they remain until they have reached room temperature. The crucibles are then balanced and the amount of ash is determined together with the empty weight of the crucibles.

2.11. Spectroscopic and Sensoric Data Acquisition

The basis of data used for the PLS regressions are consolidated off of two extraction runs with the same extraction parameters. Each extraction process is collected as 12 fractions from which the spectroscopic and sensoric data are measured. Additionally, to broaden the data set, three mixed fractions between fractions 1 and 2 were prepared with 25%, 50%, and 75% of fraction 1 and the other part with fraction 2.

The sample solution is pumped with an HPLC pump at a flow rate of 1 mL/min. The dead volume of the system is 5 mL. Before a sample series is measured, the complete measuring section is rinsed with HPLC water and the zero signal of the DAD, the FTIR as well as the fluorescence spectrometer is determined by means of the water. For the Raman spectrometer, no background is subtracted from the measurement signal. The DAD (Smartline DAD 2600, Knauer Wissenschaftliche Geräte GmbH, Berlin, Germany) acquires a spectrum from 190–520 nm at a sampling rate of 0.2 s^{-1} and an integration time of 32 ms each. Since this instrument records a continuous signal over the entire measurement period, the time at which a constant signal has been established is noted and, during evaluation, the measured value is averaged over a time range of one minute from the noted time. The spectrum recorded by the FTIR spectrometer (Alpha II, Bruker, Billerica, MA, USA) spans a range of 400–4000 1/cm. This is determined by averaging 100 scans (approximately 2 min), the number of scans for background measurement is also 100. For the fluorescence spectrometer (Jasco FP-2020 Fluorescence detector), excitation is at 340 nm ([39–41]) as this is a common excitation intensity for sugars. The spectrum is then recorded in a range of 0–900 nm, the measurement is repeated three times per sample, and the spectra are averaged before further processing. The Raman spectroscopy spectrum (785 nm laser, 1.5 mW, Ocean Insight, Ostfildern, Germany, with a Raman probe from InPhotonics) is obtained by automatically averaging a total of three spectra with an integration time of

10 s each. For each sample, at least three spectra are measured and averaged before further processing.

The density of the extracts is determined with the DMA 500 density meter from Anton Paar (Anton Paar Group AG, Graz, Austria). Conductivity is determined using the GLMU 200 MP-TR meter from Greisinger (GHM Messtechnik GmbH, Regenstauf, Germany), turbidity using the Communication Interface ECI-03 in combination with the EXcell 231 turbidity probe from Exner (Exner Process Equipment GmbH, Ettlingen, Germany) and pH using the inoLab® pH 7110 pH-meter together with the SenTix™ 81 probe from WTW (Xylem Analytics Germany Sales GmbH & Co. KG., Weilheim, Germany).

2.12. Data Analysis

After data acquisition, the obtained spectra were processed and analyzed using Unscrambler® χ (Camo Analytics AS, Oslo, Norway). The raw spectra were analyzed using descriptive statistics to select an appropriate preprocessing strategy. An important tool was the graphical representation of the spectra compared to the mean spectrum, also known as the scatter effects plot, which shows the nature of the distortion effects in the spectra. Figure 3 shows example scattering effects plots for (a) additive effects, (b) scattering effects, (c) additive and scattering effects, and (d) complex effects. Additive effects were eliminated by applying a baseline correction and/or derivative. Unless otherwise stated, the first derivatives were taken. Multiplicative scatter effects were removed by applying the standard normal variant or multiplicative scatter correction (MSC). Combinations of additive and multiplicative effects were removed by a combination of baseline correction and scatter correction. Complex effects were not observed, although they could be removed using extended multiplicative scatter correction (EMSC). The preprocessed spectra were then used to correlate changes in component concentration with changes in spectral intensity in specific spectral regions using PLSR. A detailed description of the basics of PLSR was published by Esbensen et al. [42]. Hereby, a maximum of four principal components were allowed in the regression to avoid overfitting the data. The quality of the model was assessed by the plot of spectral line loadings, the score plot, and the plot of explained variance against the number of principal components.

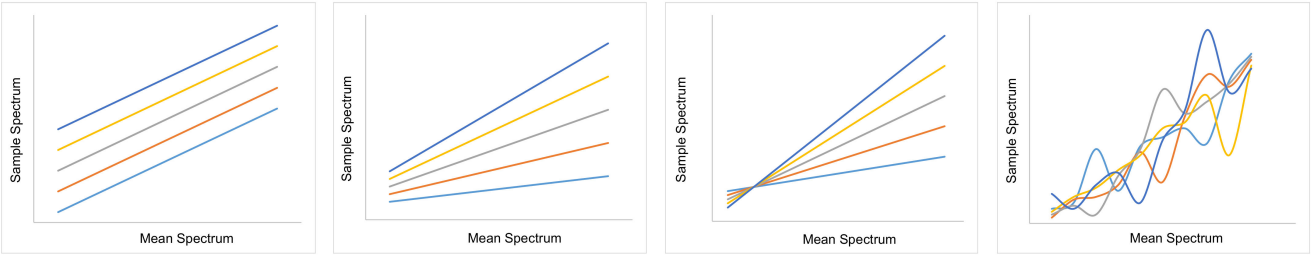


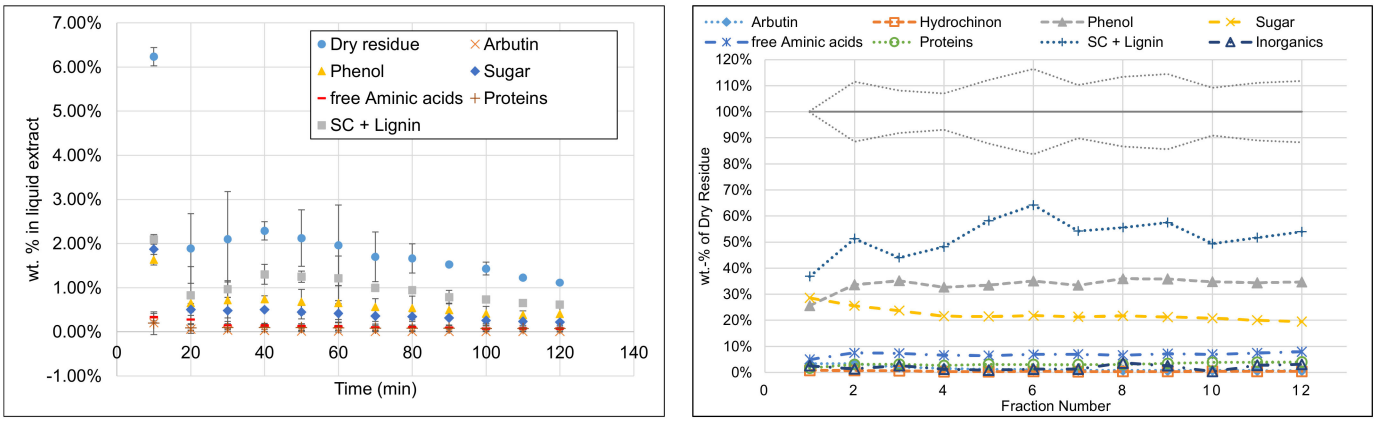
Figure 3. Exemplary scatter effects plots of (**a**) additive effects, (**b**) scatter effects, (**c**) additive and scatter effects, and (**d**) complex effects. Colors indicate spectra from different samples.

3. Results

3.1. Extract Composition

In this section, the results of the offline analytics for the two extraction runs are presented. In Figure 4a, a typical extraction profile for percolation processes can be seen, in which the different component groups are in the highest concentration in fraction 1 and the extraction profile flattens out over the extraction processes. However, there are differences between the extraction profiles of the individual extracted component group. For the component groups sugar and arbutin, for example, a constant decrease can be observed over the extraction process. For phenols and structural carbohydrates and lignin, it can be observed, that there is an increase in extracted mass between 20 and 40 min in the process.

This can be explained by the hydrolysis of the plant matrix during the extraction process. This impacts the composition of the dry mass, which varies over the extraction process. The variation of the composition is shown in Figure 4b with an indicator of how well the mass balance is closed in the respective fraction. The composition shows an equivalent trend as the extraction profile.



(**a**) (**b**)

Figure 4. (**a**) Extraction profile for components and component groups and (**b**) composition of dry residue with fidelity of mass balance shown as dotted line around 100%.

After conducting the analyses of the individual groups of substances, it turns out that the largest part of the extract consists of structural carbohydrates and lignin with 43% of the extracted dry mass. The second group is formed by phenolic components with 34% and the third largest group is the sum of free sugars with a partition of 22% of extracted dry mass. The rest is made up of free aminic acids (7%), proteins (3%), inorganic components (2%), fats (0.5%), and acetic acid (0.3%). The overview of the composition of the overall extracted dry mass is presented in Figure 5.

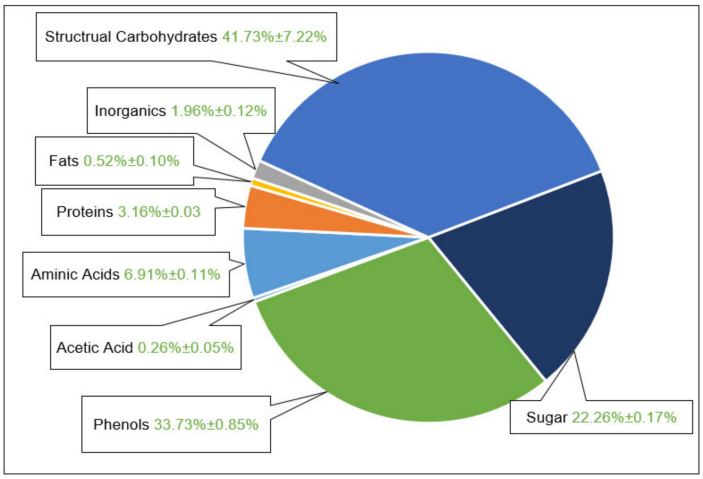


Figure 5. Overview of extracted dry mass composition.

3.2. Spectral Data and Preprocessing

The determined spectra are processed as described in Section 3.4 based on the ranges used for correlation and the scattering effects. When processing on the basis of the scattering effects, care must be taken not to remove too many differences in the spectra, but only those that are not related to the concentration change within the extraction. When processing the Raman spectra, these are, for example, effects caused by fluorescence, since there are many

chromophoric compounds in plant extracts. If all differences are filtered out of the spectra, one and the same spectrum is obtained, which means that there are no more differences between the individual spectra, which are, however, required for the correlation of the results from the offline analysis. In order to correlate the data generated by spectroscopy with the data from offline analysis, the individual spectra must be narrowed down and processed. First, the raw spectrum is examined and an assessment is made as to which areas can already be excluded. The se can be areas in which there is no deflection at all, or areas in which there are many deflections with high amplitude and high frequency. The next step is to correlate the reduced raw spectrum with the data from the offline analysis, in this case, the dry extract, and to assess in which areas of the spectrum data for correlation are available. When the correlation is made using PLS, one-half of the data set is used for calibration and the other half for validation of the system.

For FTIR, the scatter effects—as well as the processed part of the spectra—are shown in Figure 6. At first, spectra were correlated to the offline data to identify the part of the spectrum, which contains the required information. For FTIR the spectrum is cut down to 930–1800 1/cm. This represents the majority of the fingerprint region (500–1500 1/cm) and includes the acidic O-H bend, C-O stretch, and C-N stretch within the fingerprint region extended by the strong C=O stretch of acids, aldehydes, carbonyls, and esters at approximately 1650–1800 1/cm. Accordingly, the most relevant strong interactions with chemical bonds of the expected chemical compounds in the extract show a good correlation with the offline data, as expected. The narrowed spectra are further preprocessed with a 1st order derivative to eliminate additive effects as well as impacts of the measurement equipment itself. The derived spectra are used for the creation of the PLS models.

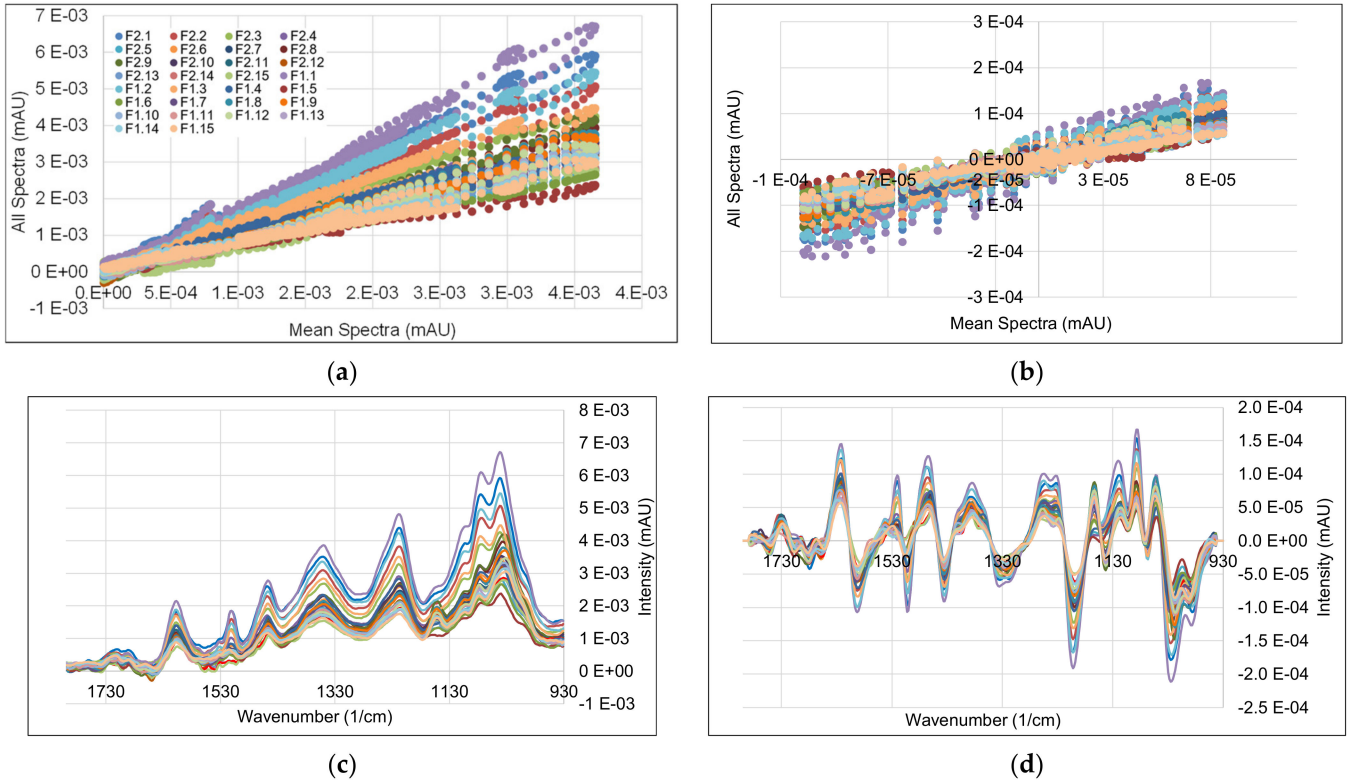


Figure 6. (**a**) Scatter effects of FTIR spectra before and (**b**) after preprocessing; (**c**) data of FTIR spectra before and (**d**) after preprocessing, legend from (**a**) applies for all figures.

For Raman, the scatter effects—as well as the processed part of the spectra—are shown in Figure 7. At first, spectra were correlated to the offline data to identify the part of the spectrum, which contains the required information. For Raman, the spectrum is cut

down to 340–2000 1/cm. The narrowed spectra are further preprocessed with a 1st order derivative to eliminate additive effects as well as impacts of the measurement equipment itself. The derived spectra are used for the creation of the PLS models. In Figure 7b, it can already be seen that the spectra of Raman are a lot less differentiated than the spectra of FTIR represented in Figure 6b.

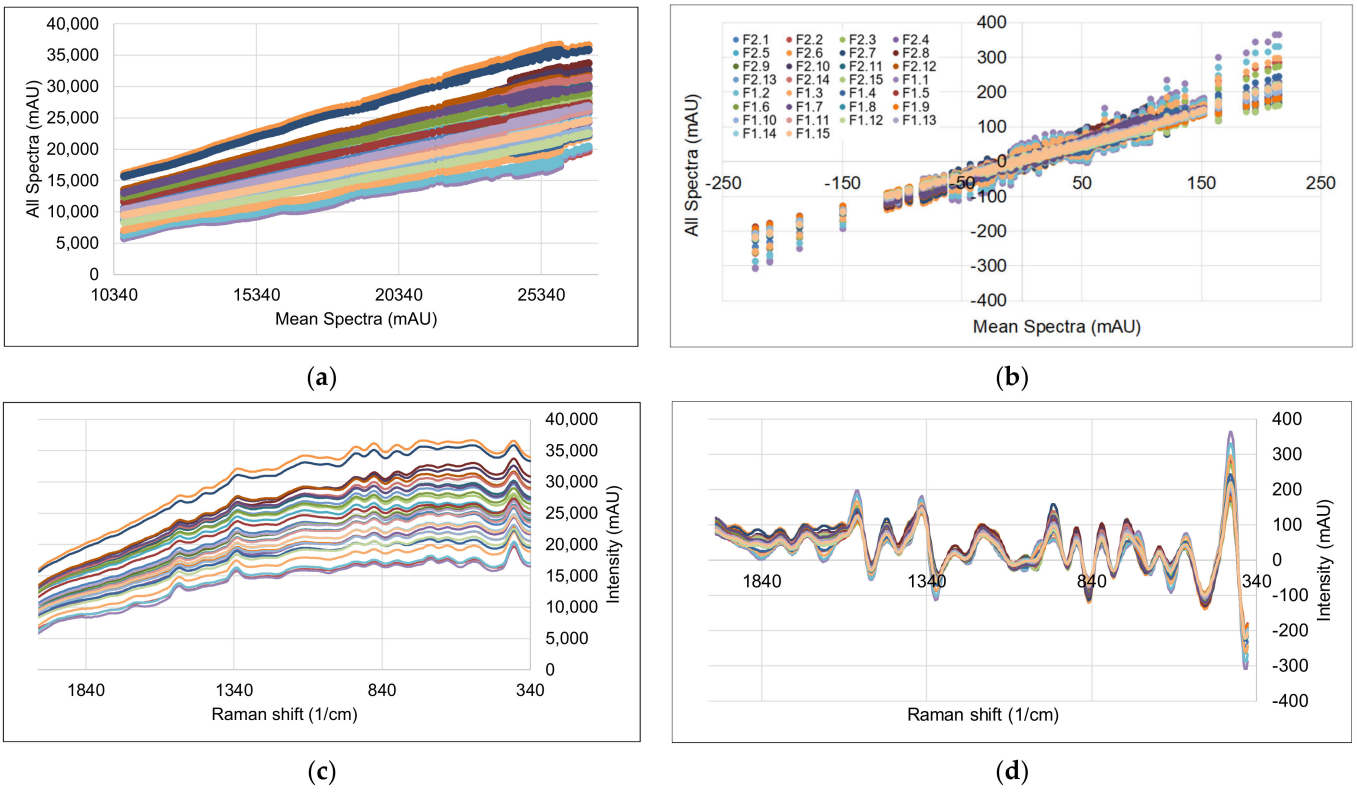


Figure 7. (**a**) Scatter effects of Raman spectra before and (**b**) after preprocessing; (**c**) data of Raman spectra before and (**d**) after preprocessing, legend from (**b**) applies for all figures.

Since the spectroscopy methods DAD and fluorescence are not suitable for the measurement of turbid samples, the samples are filtered before measurement and the dry residue is redetermined. The data of run 1 had to be discarded because they were measured without prefiltration. After recognizing, that the data cannot be processed, there was not enough sample volume left to re-measure the samples. DAD spectra and scatter effects are shown in Figure 8. A PCA was performed and according to the results, no correlation was found for wavelengths up to 350 nm. A significant variance in the spectrum between 250 and 300 nm was expected, since many phenols have their specific absorption maxima in this wavelength range, for Arbutin the absorption maximum is at 280 nm [43]. DAD spectra were then shortened to 350–450 nm and for correlation, the 1st order derivative is used. For DAD, a shift in the absorption spectrum can be observed which is not explainable by the observed extraction profiles (cf. Figure 4a). However, the occurring effect correlates to the extraction process.

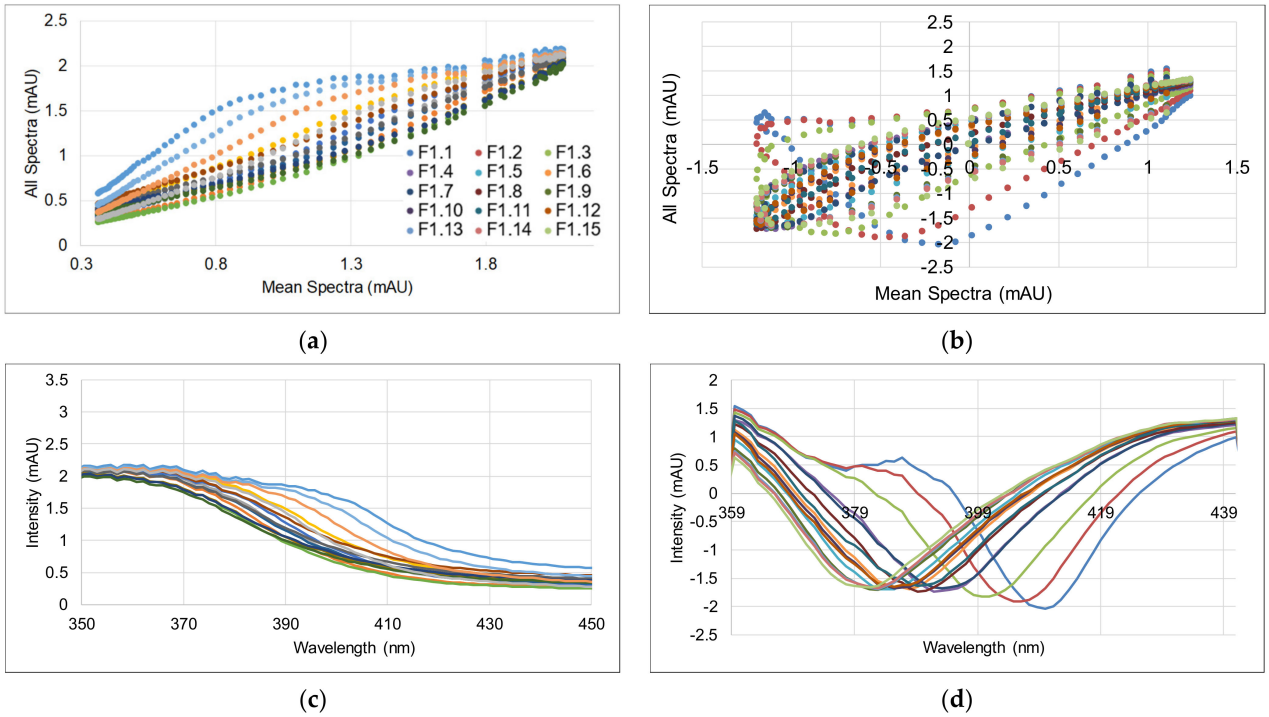


Figure 8. (**a**) Scatter effects of DAD spectra before and (**b**) after preprocessing; (**c**) data of DAD spectra before and (**d**) after preprocessing, legend from (**a**) applies for all figures.

Spectra and scatter effects for fluorescence are given in Figure 9. The spectra were cut to 380–630 nm. In the spectra as well as in the scatter effects, it can be observed that before (Figure 9a) and after (Figure 9b) 1st order derivation there is little information in the spectral data. Effects in the extraction process, however, can be explained to a certain extent.

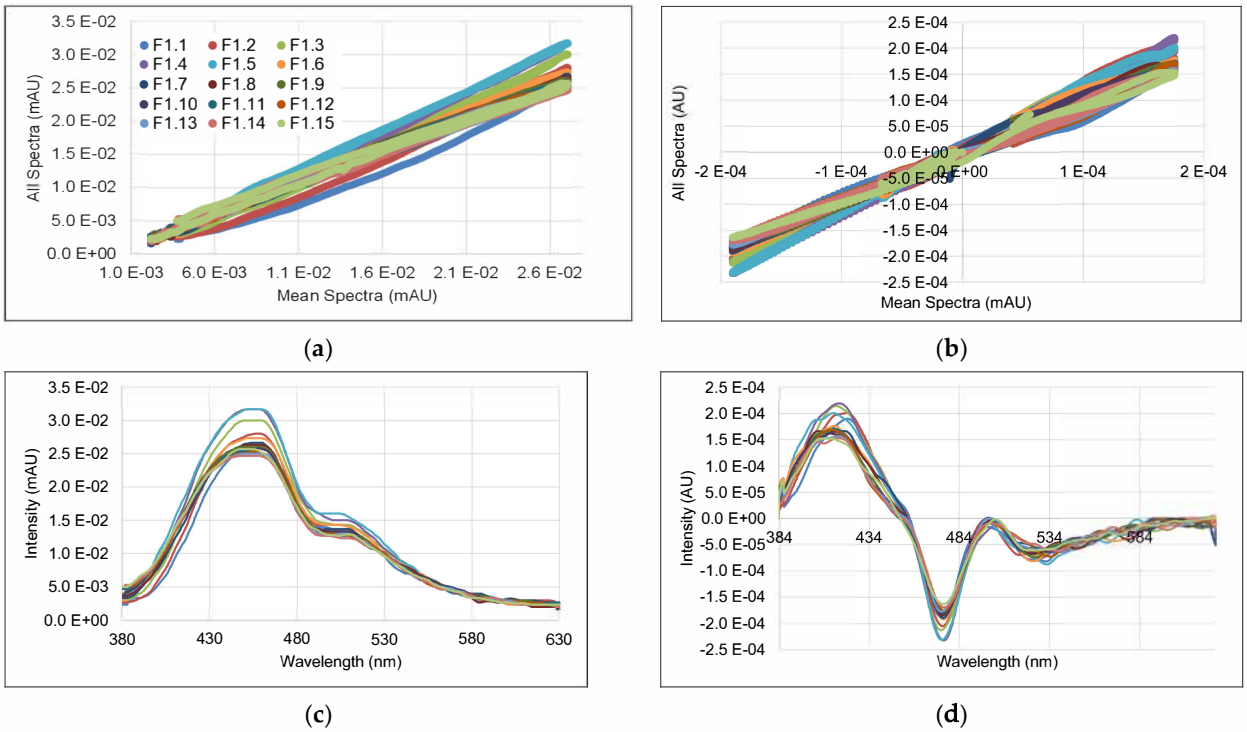


Figure 9. (**a**) Scatter effects of fluorescence spectra before and (**b**) after preprocessing; (**c**) data of fluorescence spectra before and (**d**) after preprocessing, legend from (**a**) applies for all figures.

3.3. PLS Models

The orthogonal factors used for correlation are used to evaluate the individual PLS systems. The se are considered here up to the fourth factor. The PLS with the highest correlation in the validation for the respective component group is selected and analyzed in more detail. Due to the great number of plots resulting from four spectroscopy methods plus sensoric measurements for eight components/component groups, the PLS models are shown and discussed for two representatives. The PLS models and their application on extraction curves as well as the half-normal plots are shown for all components in Appendix A in Figures A1–A4. In the following, the component group dry residue, which is the sum of all extracted components, and the component arbutin, which is the marker substance of bearberry, are discussed. Both have different extraction profiles, so they can show the independence of the PAT system from co-linearity. In Figure 10a,b, the offline data of both extraction runs are plotted against the predicted values of the PLS model for FTIR. The correlation for dry residue and arbutin works well with an R^2 of 0.99 for calibration as well as for validation for each component. It can be seen that the PLS is able to represent components with a second local extraction maximum at 40–60 min as well as components with a constant decrease over the extraction process. The model is able to distinguish between these two groups. FTIR has even the sensitivity to differ between very similar components like in this example arbutin and hydroquinone.

In Figure 10c,d, the PLS predicted results—as well as the offline data of both extraction runs—are plotted for Raman spectroscopy. The R^2 are for calibration 0.93 of dry residue and 0.93 for arbutin and for validation 0.82 for dry residue and 0.56 for arbutin. It can be seen, that the prediction accuracy is only partially satisfactory. Especially for run 2, some data points from the reference analytics are not matched by the predicted value. Furthermore, the prediction error is rather large compared to the prediction error of the PLS with FTIR data. In principle, the Raman spectrum contains sufficient variance to be able to represent the different characteristics of the two components under consideration. However, due to the low signal strengths, fine concentration differences can no longer be described, or only with great uncertainty.

In Figure 10e,f, PLS predicted data and reference analytics are shown for the DAD. As already mentioned, for DAD there was only one of the two extraction runs evaluable, because turbid samples could not be correlated. The DAD spectrum in Figure 8 already shows that there is more or less only one single peak, which can be used for correlation. This leads to a problem, which shows itself in the predicted data for arbutin. For the dry residue, the PLS model reaches an R^2 of 0.99 for calibration and 0.96 for validation and is able to predict the extraction run sufficiently. However, due to the low variance in the spectra, the divergence of extraction profile for arbutin cannot be represented by the PLS model with R^2 of 0.93 in correlation but only 0.83 in validation.

The same behavior can be seen for fluorescence, for which the reference and predicted extraction profiles are shown in Figure 10g,h. With an R^2 of 0.98 for calibration of dry residue and 0.97 for validation, the fluorescence spectra are able to represent the extraction profile of dry residue well. However, the PLS model cannot predict the extraction profile of arbutin suitably well, as it overestimates the deciding value at 40 min, which indicates, whether the PLS can represent the different extraction profiles or not. The R^2 for arbutin is still acceptable with 0.93 for calibration and 0.81 for validation, but the missing variance in the spectrum (cf. Figure 9) disqualifies this spectral method for process analytical technology of complex multi-component mixtures, as well as the DAD.

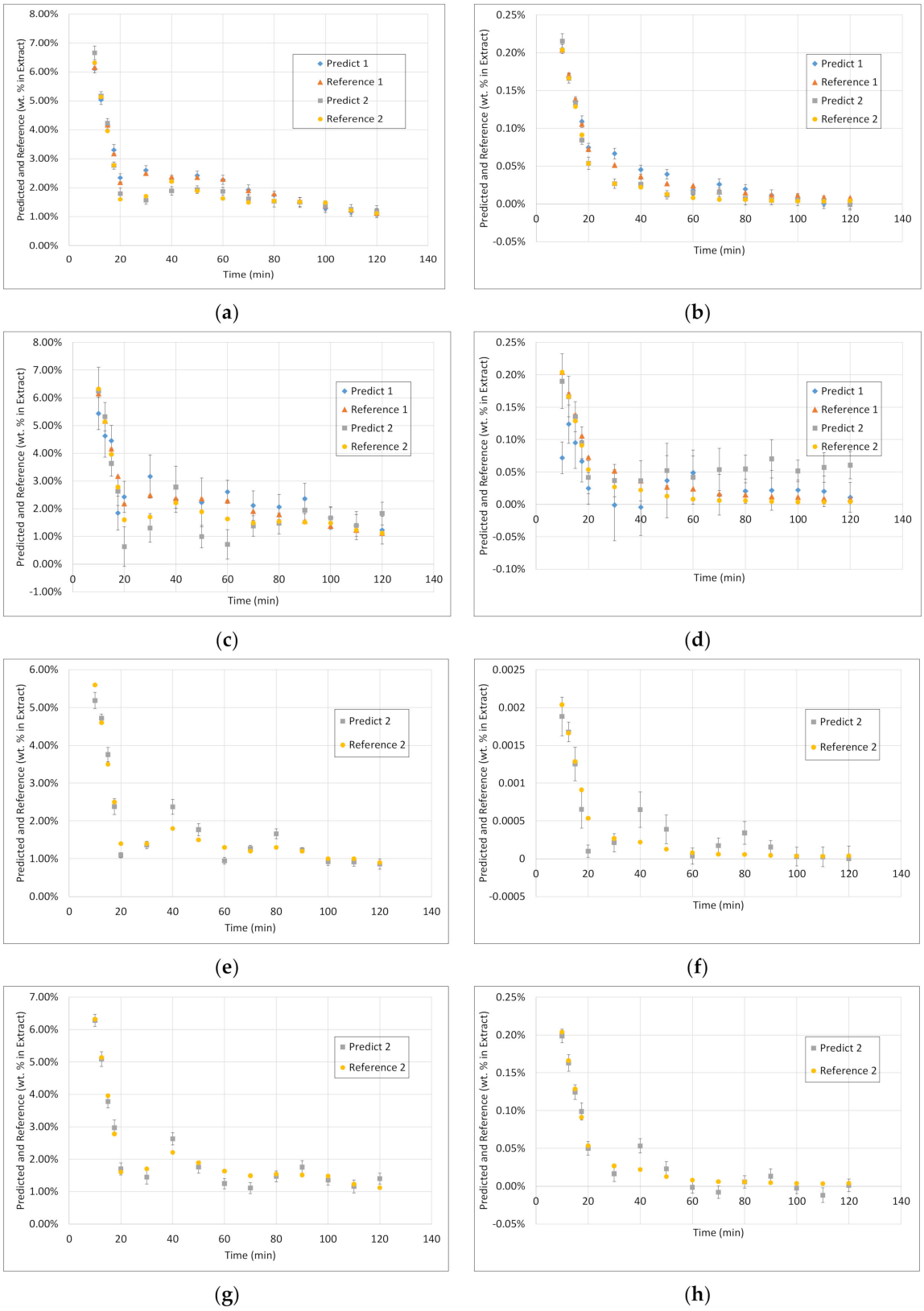
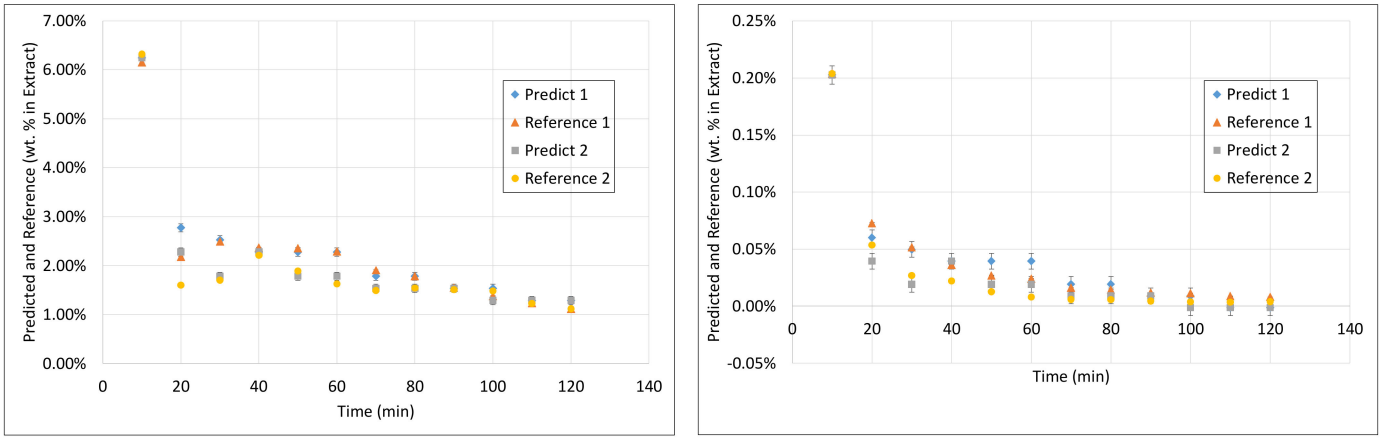


Figure 10. Extraction profiles with reference offline analytics and predicted with PLS based on spectra from extraction run 1 and 2. (**a**) Dry residue (FTIR), (**b**) arbutin (FTIR), (**c**) dry residue (Raman), (**d**) arbutin (Raman), (**e**) dry residue (DAD), (**f**) arbutin (DAD), (**g**) dry residue (fluorescence), (**h**) arbutin (fluorescence).

Besides the spectroscopic data, data for sensoric measurement methods such as conductivity, pH, density, and turbidity were collected. For pH, density, and turbidity no correlation for the respective component groups was found. Explanatory approaches are to be made. Due to the extraction of components such as acetic acid, naturally occurring buffer systems are present in the extracts, which have the consequence that pH is only able to describe the extraction behavior up to that point, where the buffer is formed. In a buffered solution, the concentration of single components cannot be correlated with the pH value. The dry residue would be the only component group that could be correlated on density as the sum of all extracted components. In the case of density, the impact of the overall extracted dry mass is too small to have a significant influence on the density. The refore, no significant variance in the density signal is detected to which the dry residue could be correlated. For the turbidity, postulation was to correlate to the concentration of structural carbohydrates, which are stably dispersed after cooling of the extract. This postulation turned out to be wrong. No correlation for turbidity for any of the researched component groups could be found.

However, for conductivity, the correlation with dry residue was exceptionally good. With an R^2 of 0.99 in correlation and 0.97 for validation. Due to the fact, that conductivity is only a univariate signal, it is only able to represent one particular system behavior. Even there are also good R^2 for calibration (0.98) and validation (0.97) for arbutin, in the prediction vs. reference extraction profiles in Figure 11b can be seen, that the conductivity signal is not able to predict the arbutin extraction properly. At 40 and 60 min, the PLS model predicts a little jump according to the extraction profile of dry residue. Where a multivariate signal such as the FTIR is able to differentiate between these two, conductivity as a univariate signal is not.



(**a**) (**b**)

Figure 11. Extraction profiles with reference offline analytics and predicted with correlation of conductivity measurement for (**a**) dry residue and (**b**) arbutin.

The Calibration and Validation R^2 for all spectroscopic methods are shown in Table 1. For all component groups, FTIR delivers good results.

Merely the improvable validation of proteins leads to a higher inaccuracy of the prediction. However, the predicted value fits well, with only a larger RMSE (cf. Figure A1 (n)). The best prediction for proteins was possible with fluorescence with an R^2 of 1 in calibration and 0.95 in validation. An improvement in prediction accuracy could therefore be reached by implementing a fluorescence spectroscope for the monitoring of proteins. However, due to the low variance in the spectrum, with only one significant peak, this method is far more vulnerable for variations in the extraction profile, which applies for each component group for the correlation with DAD and Fluorescence.

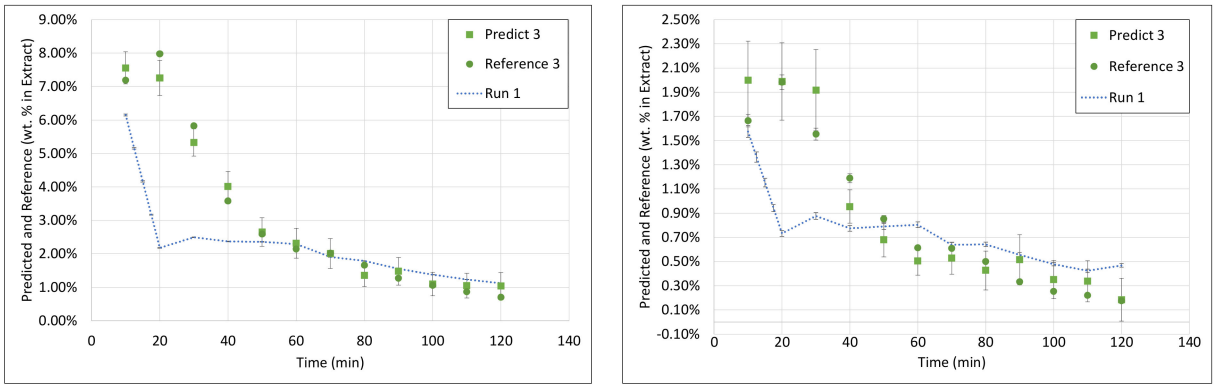
Table 1. Overview of Calibration and Validation R^2 and RMSE for each component group and spectroscopic method.

Component	FTIR				Raman				DAD				Fluorescence			
	R^2 Calibration	R^2 Validation	RMSE Calibration	RMSE Validation	R^2 Calibration	R^2 Validation	RMSE Calibration	RMSE Validation	R^2 Calibration	R^2 Validation	RMSE Calibration	RMSE Validation	R^2 Calibration	R^2 Validation	RMSE Calibration	RMSE Validation
Dry Residue	1	0.99	0.0007	0.0018	0.92	0.78	0.0042	0.007	1	0.95	0.0007	0.0027	1	0.94	0.0011	0.003
Arbutin	0.99	0.99	0.00005	0.00007	0.89	0.56	0.0002	0.0004	0.93	0.83	0.0002	0.0002	0.99	0.91	0.00006	0.0001
Hydroquinone	0.97	0.81	0.00002	0.00004	0.78	0.6	0.00006	0.00007	0.9	0.76	0.00003	0.00003	0.98	-	0.00001	0.00006
TPC	1	0.98	0.0002	0.0005	0.92	0.73	0.001	0.002	0.99	0.96	0.0005	0.0007	1	0.92	0.0002	0.001
Free sugars	1	0.99	0.0008	0.0005	0.92	0.77	0.0014	0.0024	0.98	0.91	0.0008	0.001	1	0.95	0.0004	0.001
Acetic acid	0.96	0.74	0.00002	0.00006	0.89	0.58	0.00004	0.00008	0.82	0.8	0.00005	0.00006	0.99	0.85	0.00001	0.00006
Proteins	0.93	0.4	0.0001	0.0005	0.87	0.81	0.0002	0.0003	0.99	0.91	0.00005	0.0001	1	0.95	0.00004	0.0001
Ash	0.96	0.94	0.00009	0.0001	0.88	0.71	0.0002	0.0003	0.99	0.81	0.00007	0.0002	0.99	0.8	0.00006	0.0002
Aminic acids	0.99	0.98	0.00007	0.0001	0.88	0.72	0.0002	0.0005	0.99	0.95	0.00008	0.0001	1	0.91	0.00001	0.0001
Structural carbohydrates	0.97	0.79	0.0008	0.0019	0.76	0.62	0.0022	0.0026	0.84	0.74	0.0017	0.0016	0.98	0.59	0.0006	0.002

In general, Raman delivered the least accurate results in calibration and validation. Due to fluorescence effects in the mixture, the Raman spectrum is superimposed and the Raman specific peaks are only small with low variance in the extraction range.

3.4. Application of PAT

Definitively, the created PLS models are tested on a third extraction run with slightly changed extraction conditions, with a higher flow of the extraction solvent which shifts the extraction profile. For the third extraction run, only the easy and fast obtainable components dry residue and total phenolic content were determined. Afterwards, the collected FTIR spectra were used to predict the offline analytics with the created PLS. Hereby, the PLS model consists only of spectra from the first two extraction runs. The reference analytics as well as the PLS-predicted extraction profiles are shown in Figure 12. As a reference, the respective extraction profile of extraction run 1 is added. For dry residue, it can be seen, that the PLS model is able to predict the extraction profile sufficiently, with exception of the value at 20 min. However, considering that the values of 10 and 20 min are out of the calibrated area, the model still works sufficiently with an R^2 of 0.95 and an accuracy of mass balance of only −3%. Similar behavior is observed for the total phenolic content in Figure 12b. Within the calibrated area, the model predicts the reference analytics sufficiently well. Even the values outside of the calibrated area lie within the confidence area of the model, though it is relatively broad. However, the model closes the mass balance with a deviation of −4% deviation and an R^2 of 0.93.



(**a**) (**b**)

Figure 12. Extraction profiles with reference offline analytics and predicted with FTIR-PLS prediction for (**a**) dry residue and (**b**) TPC.

As there has to be differentiated between a quality-focused PAT circuit and a control-focused PAT circuit, the good correlation of the dry residue with conductivity qualifies this technique as a potent candidate for the control-focused PAT.

Moreover, the calibration of conductivity is tested for the third extraction run. The extraction profile of the reference analytic and the predicted values as well as the reference to extraction run 1 are shown in Figure 13. Especially the first predicted fractions deviate from the reference analytics. From 40 min on, the prediction fits the reference. However, R^2 for correlation is still 0.93, however, the mass balance is missing 9.99%.

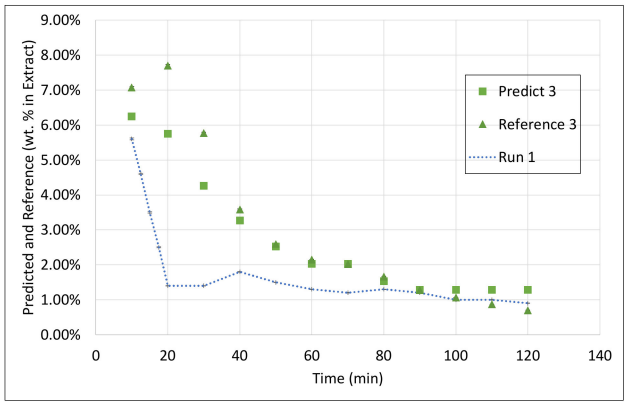


Figure 13. Extraction profiles with reference offline analytics and predicted with conductivity correlation for dry residue.

In Figure 14 the composition of the extract is shown. The composition, which was determined with offline analytics is already known from Section 3.1. The trained PLS model from FTIR data predicts the extract composition very well. PAT values of all component groups and the corresponding error range is definitely within the sophisticated standard offline analytics.

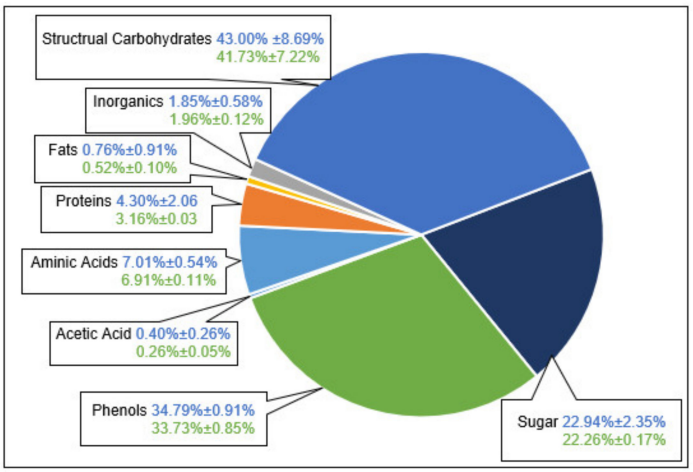


Figure 14. Composition of bearberry leaf extract determined by offline analytics (green) and presented FTIR spectra (blue).

4. Discussion

The aim of this research is the establishment and proof of concept of a method for the integration of a process analytical technology for complex component mixtures—exemplified on a natural product. This is one of the current challenges in a holistic QbD-based process design. This field is currently researched intensively in the biopharmaceutical field [44–49]. However, the QbD approach is mostly unknown to other working fields such as natural product extraction. Nevertheless, the are some research approaches in online analytics using spectroscopic data to measure concentration in different kinds of extracts or the plant material itself [24–27,29,33,50,51]. However, these works mainly focus on offline profiling of single components or multiple secondary metabolites, contrary to this work which focuses on the development of a method for process analytical technology and profiling of the whole extract and closing the mass balance.

For the elaboration of this method, three general working hypotheses were stated:

1. The first is that all the main component groups should be defined and quantified which cover the largest part of the entirety of ingredients.

- This hypothesis could be verified as the mass balance is closed sufficiently.
  - In the first fraction with the highest reproducibility of offline analytics, the mass balance was closed with only 0.21% deviation.
  - In the later fractions, the uncertainty regarding the mass balance increased, inter alia, because of less accurate results of offline analytics to about +/−5%.
  - In particular, the accuracy of structural carbohydrates in later fractions had a relatively high uncertainty about +/−20% due to the analytical methods.
- With such a reliable offline basis and the data sets of two extraction runs, the next step was the elaboration of a process analytical technology concept for these component groups for online measurement.
  - Based on the literature, in which different component groups or typical representatives of the regarding group are detected with one of the used spectroscopic methods, a concept was developed which links the components to a measurement method.
  - In addition, the chemical structures and properties and detection mechanisms of the spectroscopic and sensoric PATs were considered here.

2. As the second working hypothesis, it could be proven that main component groups could be defined and quantified. After conducting the analyses of the individual groups of substances, it turns out:
   a. That the largest part of the extract consists of structural carbohydrates and lignin with 43% of the extracted dry mass.
   b. The second group is formed by phenolic components with 34%
   c. and the third largest group is the sum of free sugars with a partition of 22% of extracted dry mass.
   d. The rest is made up of free aminic acids (7%), proteins (4%), inorganic components (2%), fats (0.5%), and acetic acid (0.3%).
3. Resulting of this concept, the third working hypothesis has been stated, that the combination of multiple spectroscopic methods such as FTIR and Raman with an additional insert of sensoric methods such as conductivity and turbidity would be necessary to distinguish between the proposed component groups.
   - Surprisingly, the FTIR alone shows enough variance in its spectra to be able to represent the slightly different extraction profiles of the different component groups (see 2. above) by itself. With R^2 between 0.96 and 0.99 for calibration and validation of FTIR for each component group, FTIR reached the overall best results.
     - During the establishment of PLS models, a combination of different spectra as well as spectra and sensoric data were researched. However, there was no benefit in predictability observed by combining these different measurement techniques.
     - Accordingly, FTIR is the best candidate for a standalone solution for the quality-focused PAT, in which product quality and composition is determined.
   - Within the field of turbidity, pH value, conductivity, and density, conductivity was the most promising technique.
     - Here, one has to distinguish between different possibilities of implementation.
       - The conductivity in a control-focused PAT circuit can be either used for model-predicted process control or as a process abort criterion.
       - In both cases, a well-predicted dry residue would be sufficient because the extraction process can be well described by the dry residue and

there is no possibility to shift the extract composition significantly by small adjustments of the process parameters.

- However, for a model-predicted control, the data set has to be increased to be able to describe the higher concentrated fractions with sufficiently higher accuracy.
  - However, nevertheless, as a process abort criterion, accurate prediction of medium to low concentrated fractions would be sufficient, which is already gained successfully at this point.

5. Conclusions

As already mentioned in the introduction, the creation of digital twins, digital twin-based evaluation of a process status, and the implementation possibilities of design spaces were already conducted in previous studies [9,10,16,17,52].

This results in the possibility of generating design spaces both in vitro and in silico via DoEs, of displaying and accordingly also predicting the extraction processes via digital twins, and of classifying these in the context of the design spaces.

Furthermore, mixtures from different raw material batches can be predictively tailored to the extraction result: Recipe-controlled batch operation is on hand.

The missing key enabling technology for linking all single parts of the overall concept together was a reliable method for the development and implementation of PAT. In Figure 15, the implementation concept of such a QbD-based process is shown.

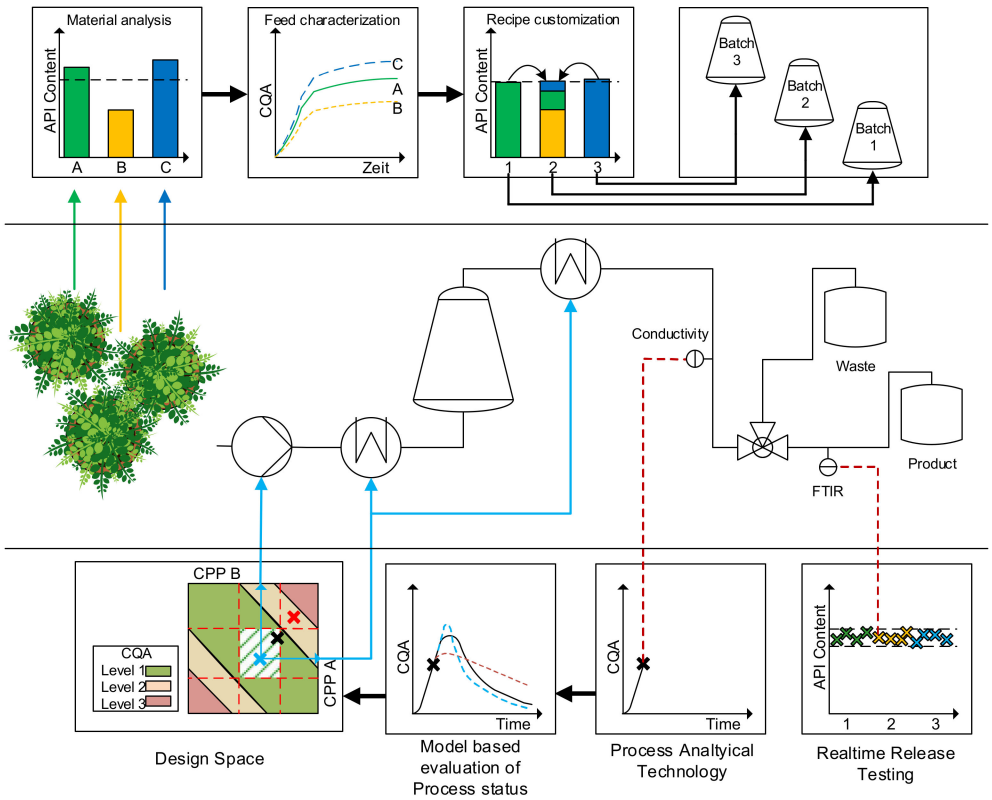


Figure 15. Overview of QbD and PAT concept and implementation, extended from [18].

This missing link between the process level and the digital twin with the design space level, the establishment of a method for PAT development, was conducted in this study.

Therefore, the complex mixture was divided into different component groups based on chemical equality.

Offline analytic methods were collected from different research fields and conceptualized and established in one coherent characterization concept for natural product extracts.

In addition, spectroscopic and sensoric measurement techniques were established, screened, and evaluated for each of the component groups.

- Based on this study, conductivity will be used for the monitoring of dry residue as the deciding component group for the process control PAT.
    - Based on the conductivity of the extract, the digital twin will evaluate the current process status and check, whether the extract lies within the permitted area of the design space.
    - Based on this evaluation the decision, whether and how process parameters have to be adjusted or not, will be made based on the digital twins prediction.
- The FTIR spectra will be used to determine the composition of the extract regarding the component groups and can therefore be used as a method for quality assurance and post-process analytics.
- Once validated, such robust methods can then be extended to real-time release testing, which will be followed up by further studies.

Figure 16 sums up the PAT methods evaluation results giving an updated overview of component groups and their corresponding offline and online analytical methods as a recommendation based on this study.

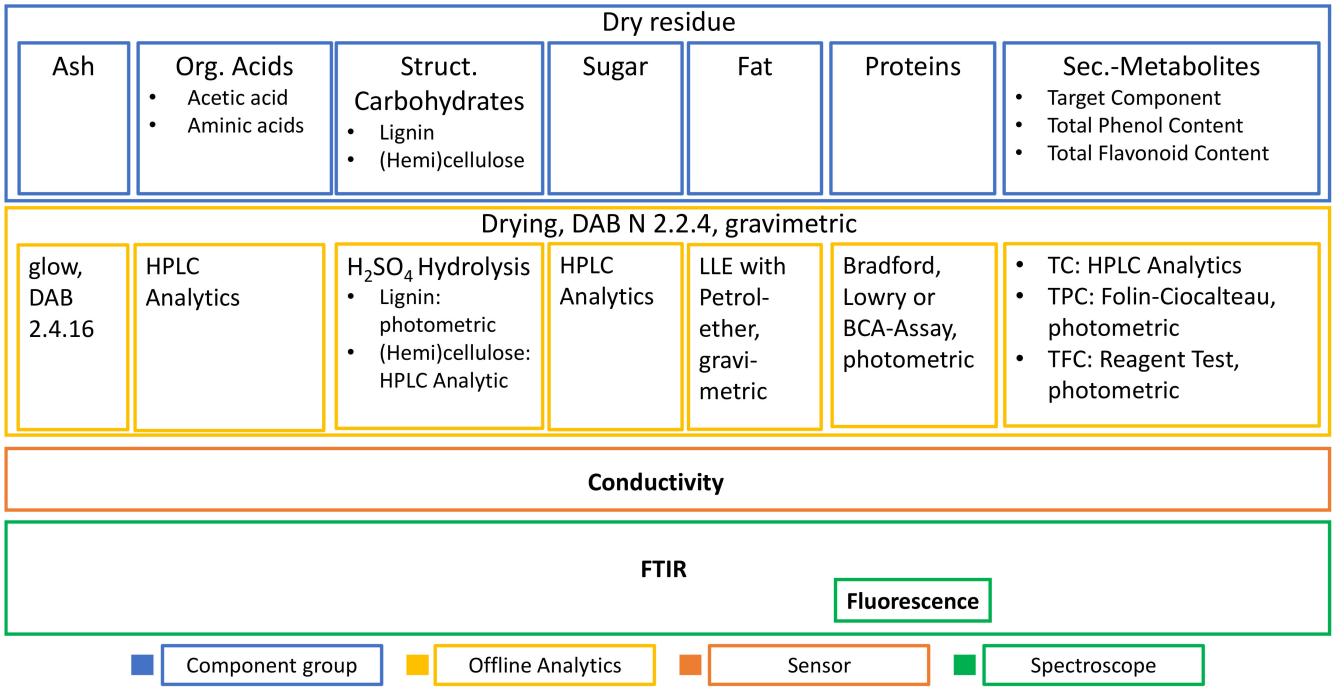


Figure 16. Updated overview of component groups and corresponding offline and online analytics, based on results in this study.

Author Contributions: Conceptualization, C.J. and J.S.; methodology, experimental design, and evaluation, C.J. and L.K.; writing and editing, C.J. and L.K.; reviewing and discussion, M.T. and J.S.; supervision, J.S. All authors have read and agreed to the published version of the manuscript.

Funding: The authors want to gratefully acknowledge the Bundesministerium für Wirtschaft und Energie (BMWi), especially Dr. Michael Gahr (Projektträger FZ Jülich), for funding the scientific work. We also kindly acknowledge the support of the Open Access Publishing Fund of Clausthal University of Technology.

Institutional Review Board Statement: Not applicable.

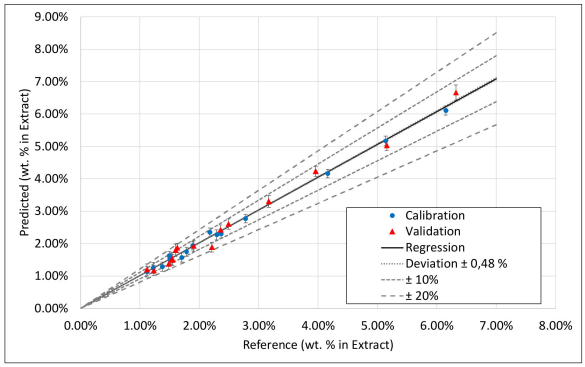
Informed Consent Statement: Not applicable.

Data Availability Statement: Data cannot be made publicly available.

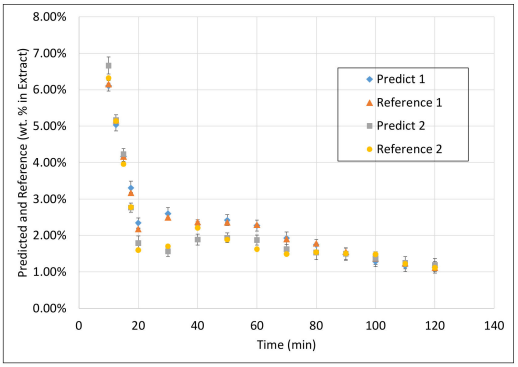
Acknowledgments: The authors would like to thank the ITVP lab-team, especially Frank Steinhäuser, Colin Herzberger, Volker Strohmeyer, and Thomas Knebel for their efforts and support. Furthermore, the authors would like to thank Axel Schmidt and Florian Lukas Vetter for the fruitful discussion.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analysis or interpretation of data; in the writing of the manuscript or in the decision to publish the results.

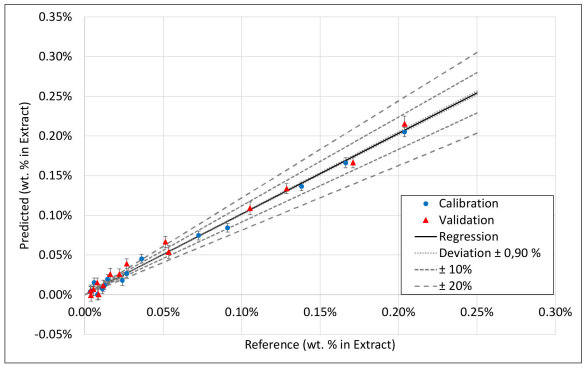
Appendix A



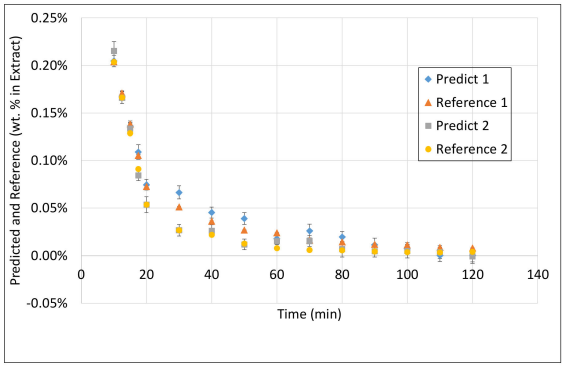
(**a**)



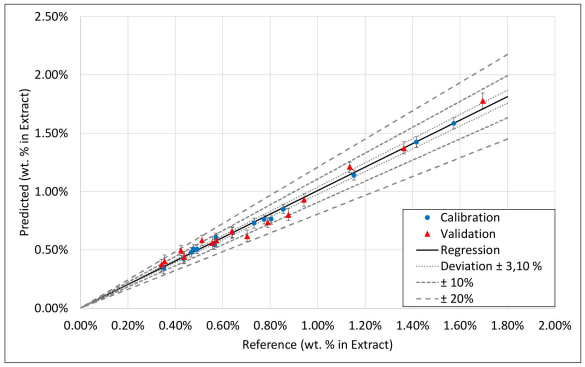
(**b**)



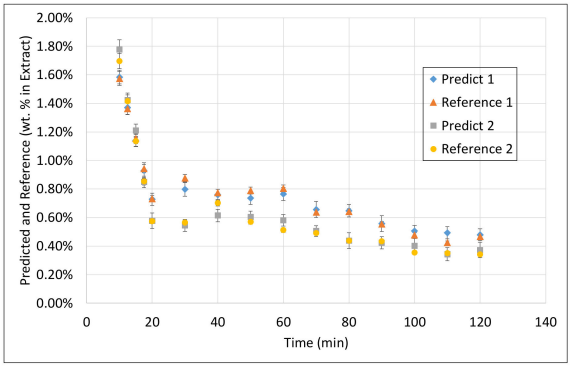
(**c**)



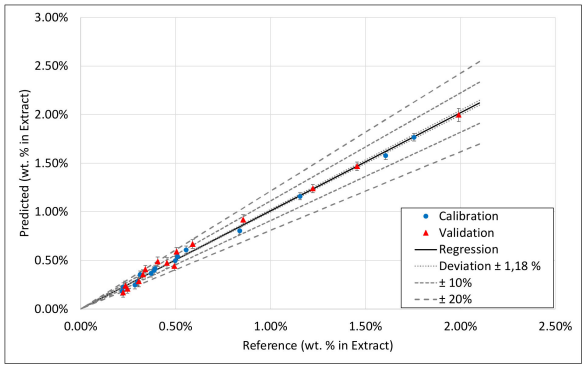
(**d**)



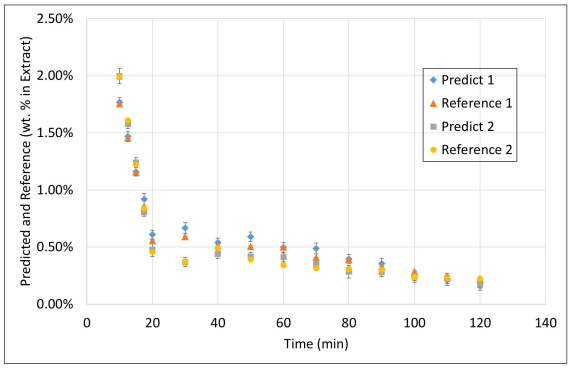
(**e**)



(**f**)



(**g**)



(**h**)

Figure A1. *Cont.*

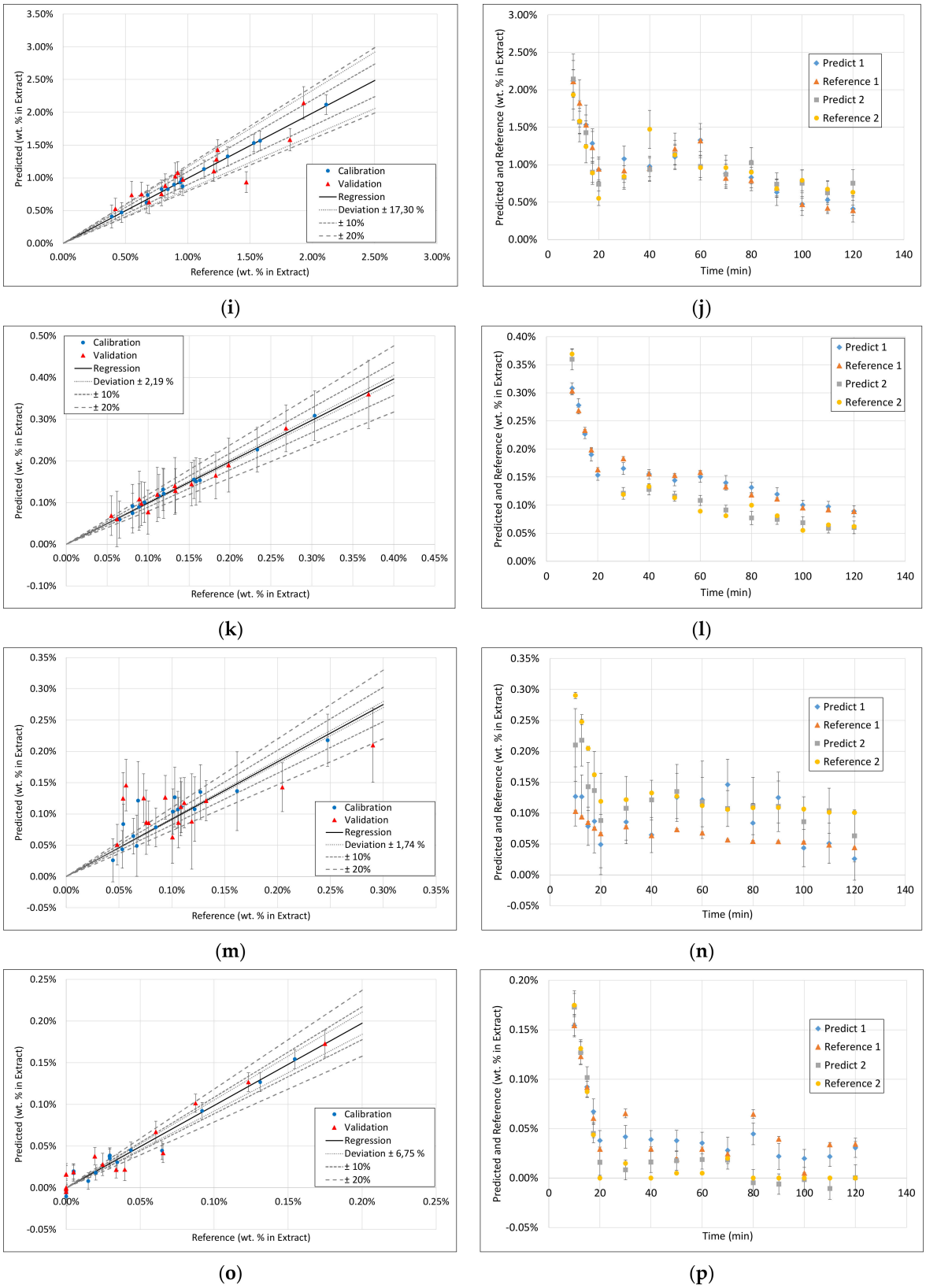
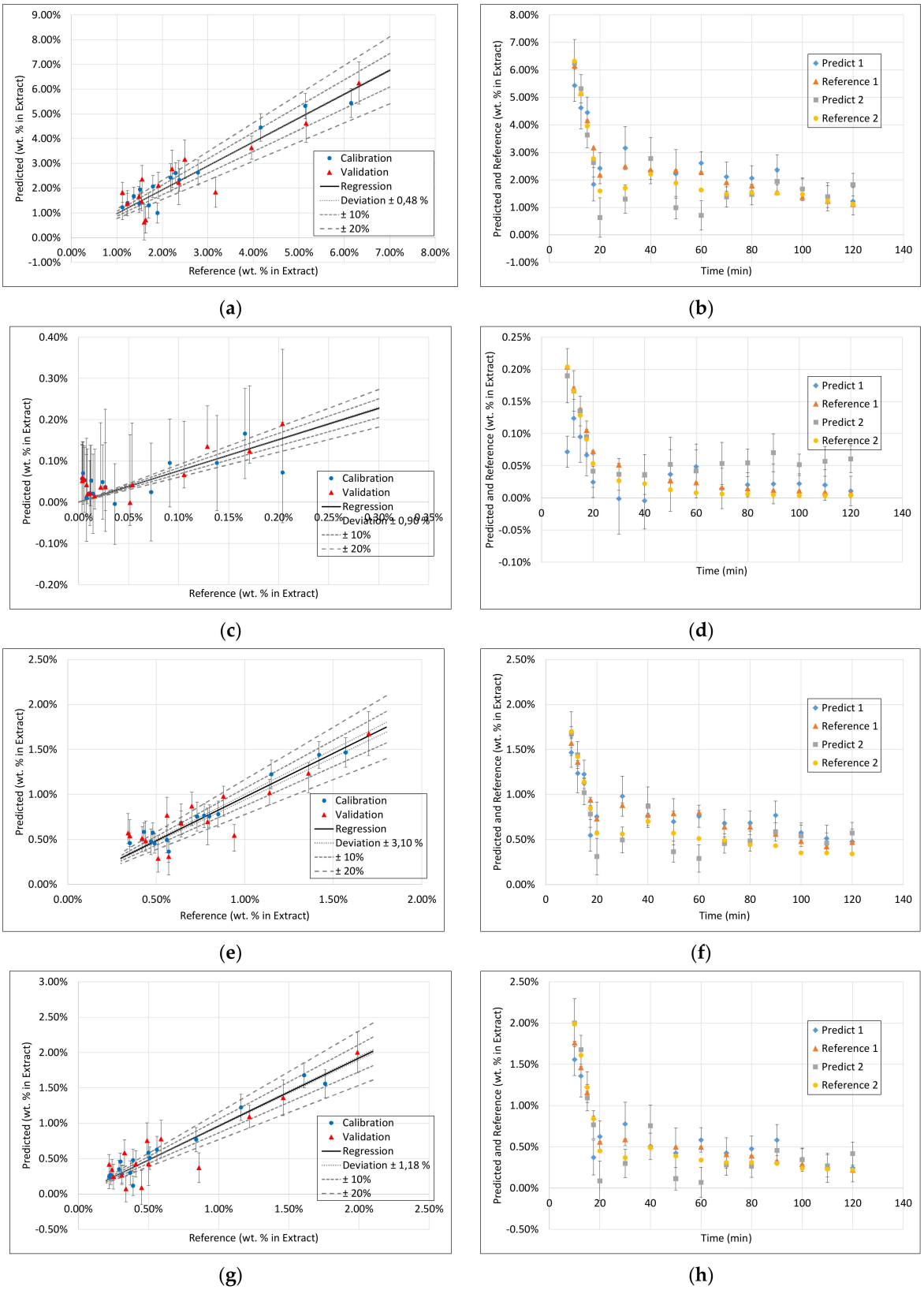


Figure A1. FTIR half-normal plots and measured vs. predicted extraction profiles for (**a**,**b**) dry residue, (**c**,**d**) arbutin, (**e**,**f**) total phenolic content, (**g**,**h**) sugar monomers, (**i**,**j**) structural carbohydrates, (**k**,**l**) aminic acids, (**m**,**n**) proteins, and (**o**,**p**) inorganics/ash.



(**a**) (**b**)

(**c**) (**d**)

(**e**) (**f**)

(**g**) (**h**)

Figure A2. *Cont.*

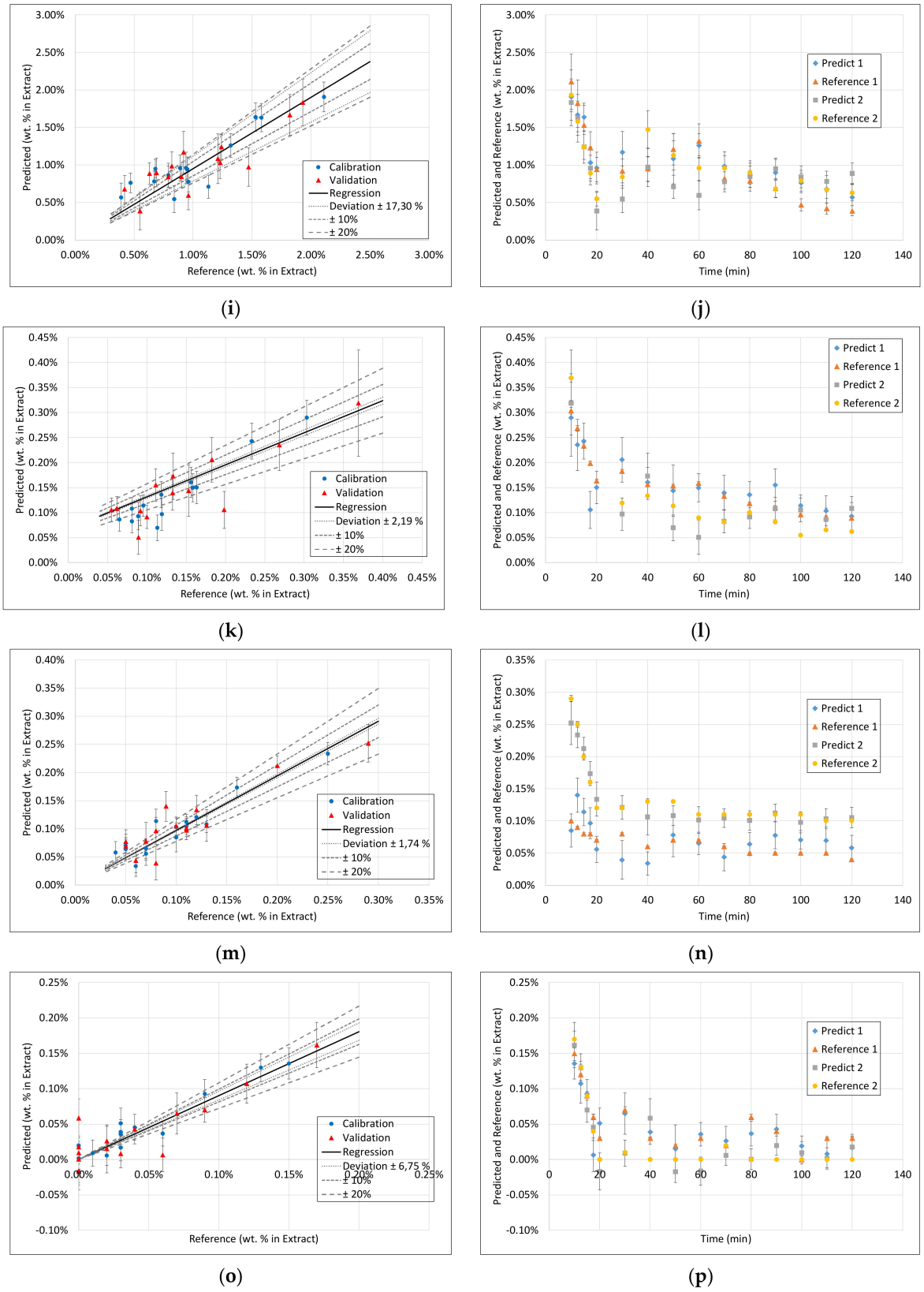
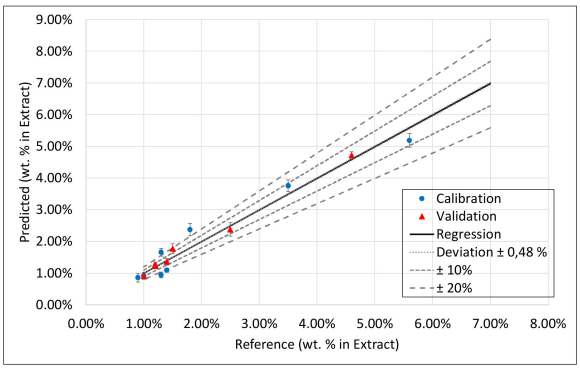
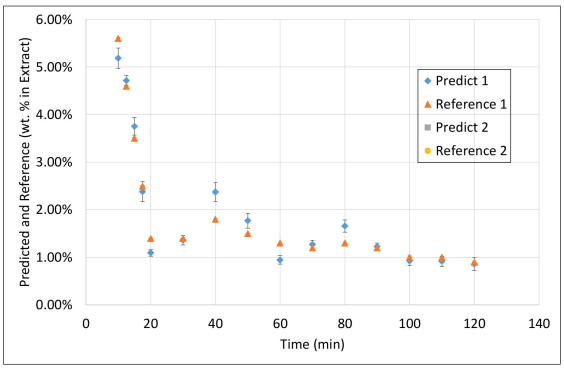


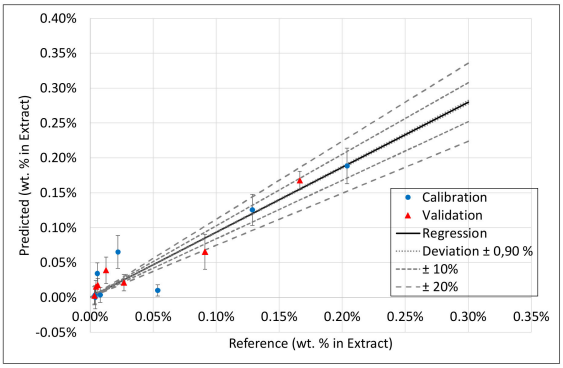
Figure A2. Raman half-normal plots and measured vs. predicted extraction profiles for (**a**,**b**) dry residue, (**c**,**d**) arbutin, (**e**,**f**) total phenolic content, (**g**,**h**) sugar monomers, (**i**,**j**) structural carbohydrates, (**k**,**l**) aminic acids, (**m**,**n**) proteins, and (**o**,**p**) inorganics/ash.



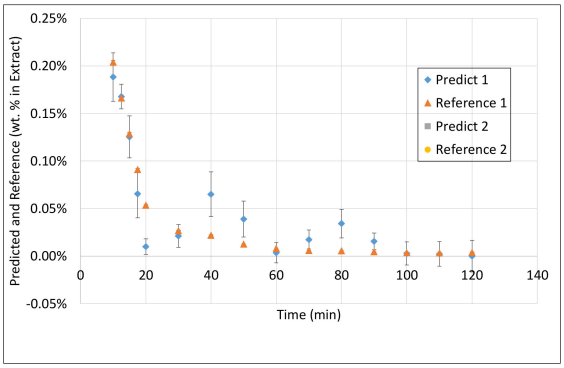
(**a**)



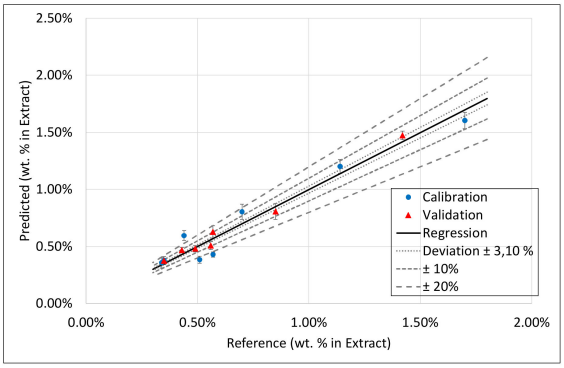
(**b**)



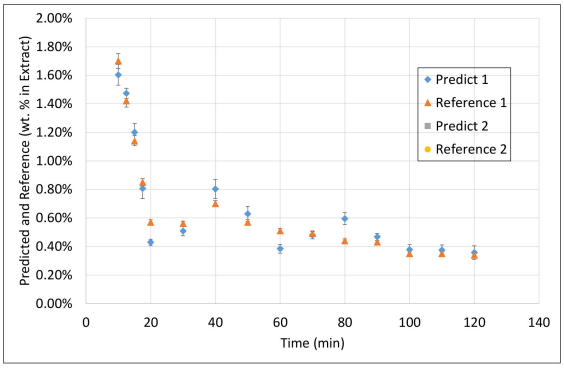
(**c**)



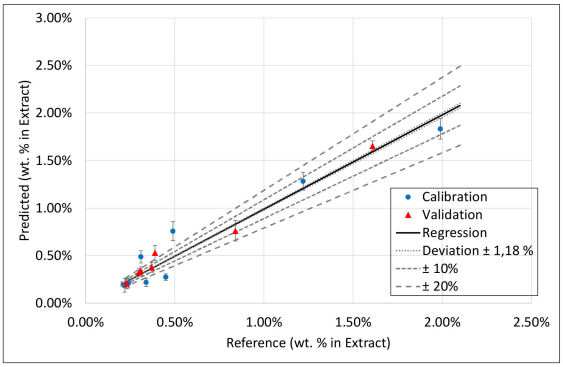
(**d**)



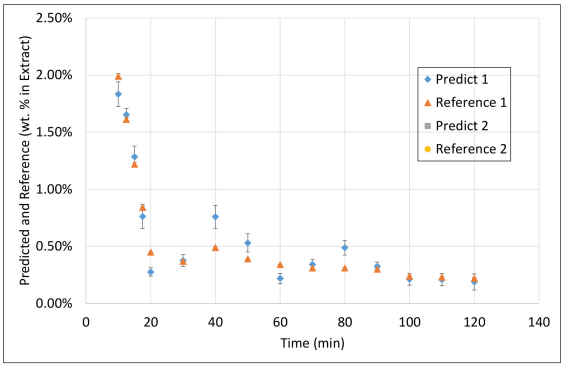
(**e**)



(**f**)



(**g**)



(**h**)

Figure A3. *Cont.*

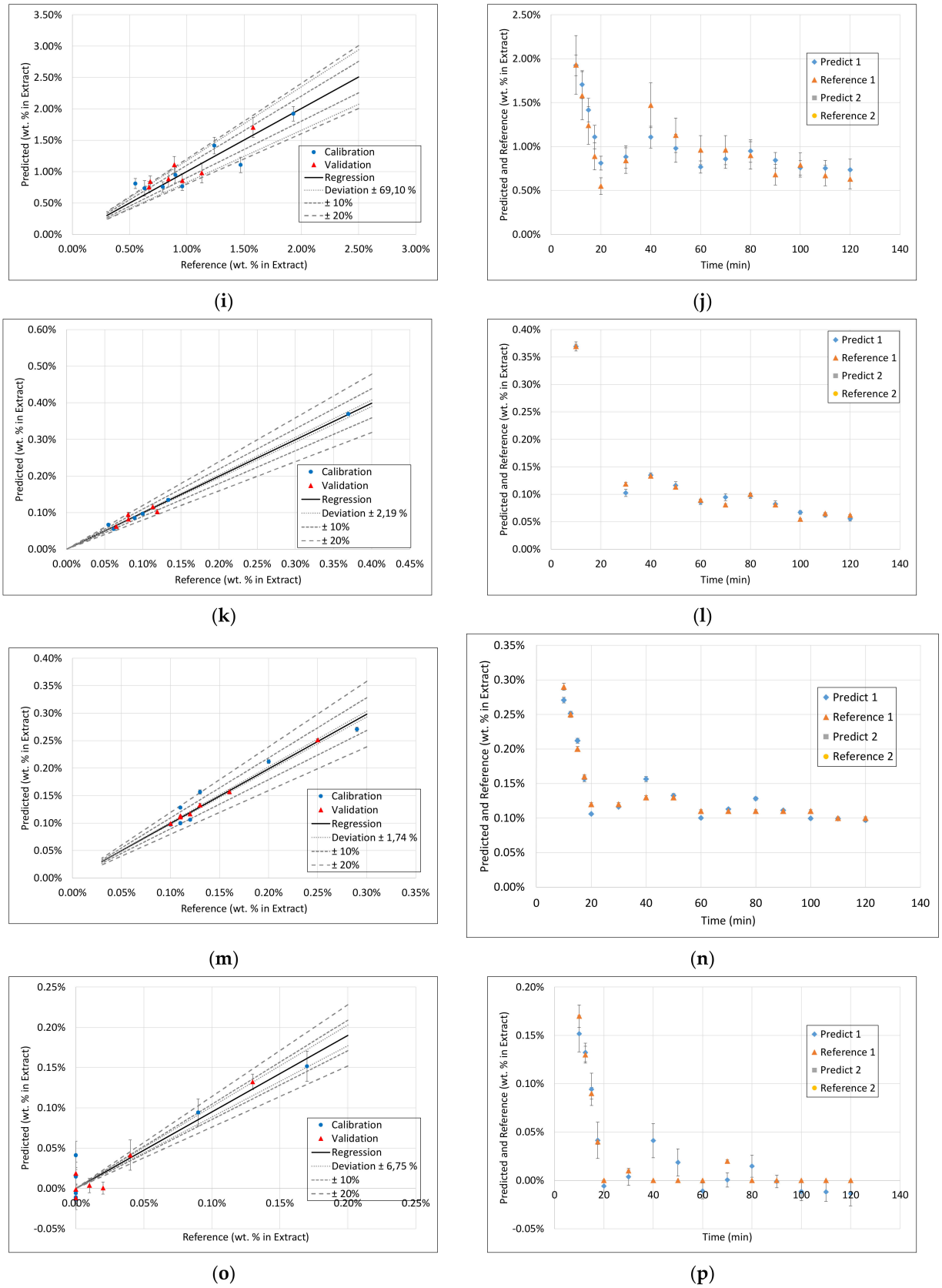


Figure A3. DAD half-normal plots and measured vs. predicted extraction profiles for (**a**,**b**) dry residue, (**c**,**d**) arbutin, (**e**,**f**) total phenolic content, (**g**,**h**) sugar monomers, (**i**,**j**) structural carbohydrates, (**k**,**l**) aminic acids, (**m**,**n**) proteins, and (**o**,**p**) inorganics/ash.

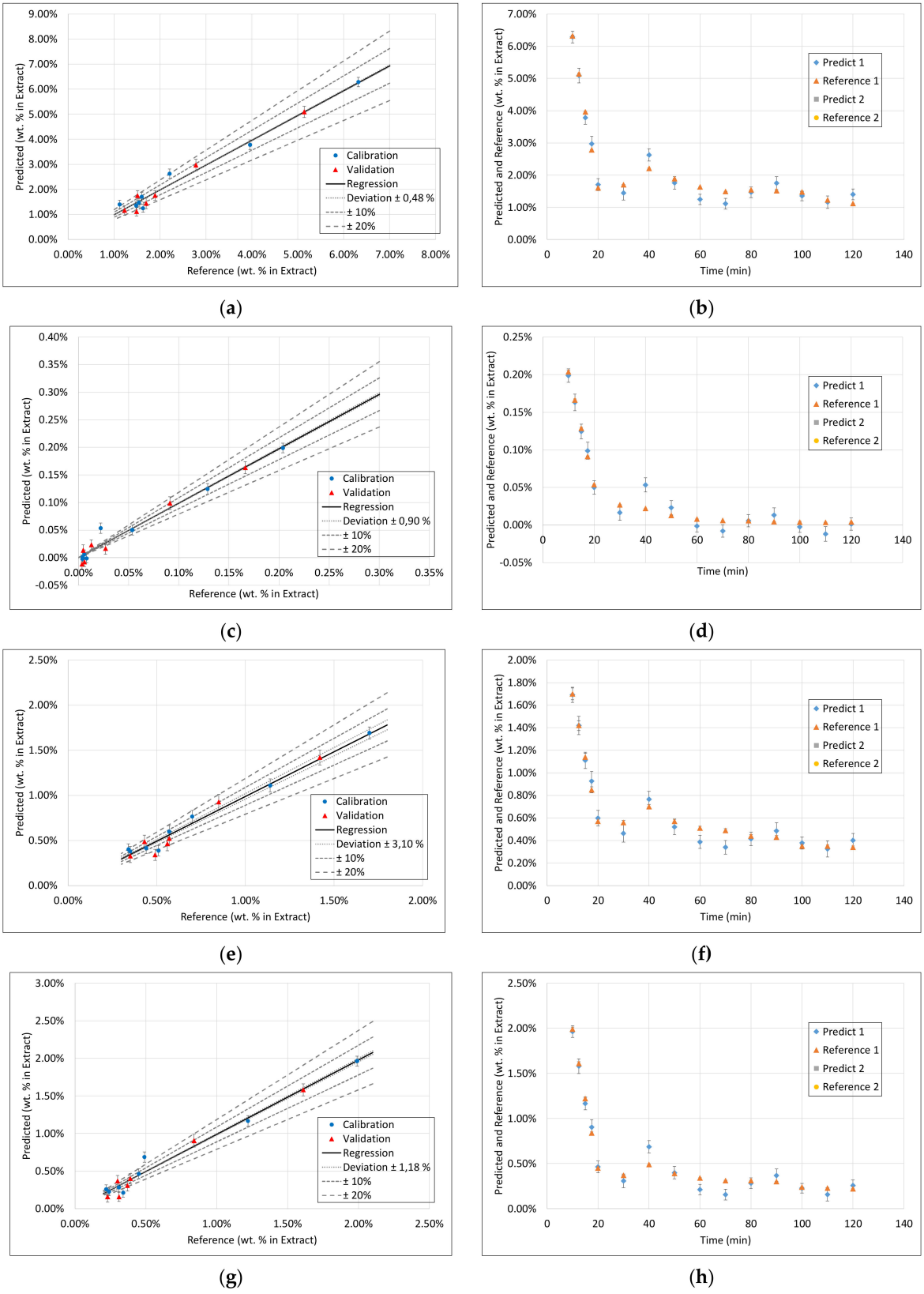


Figure A4. *Cont.*

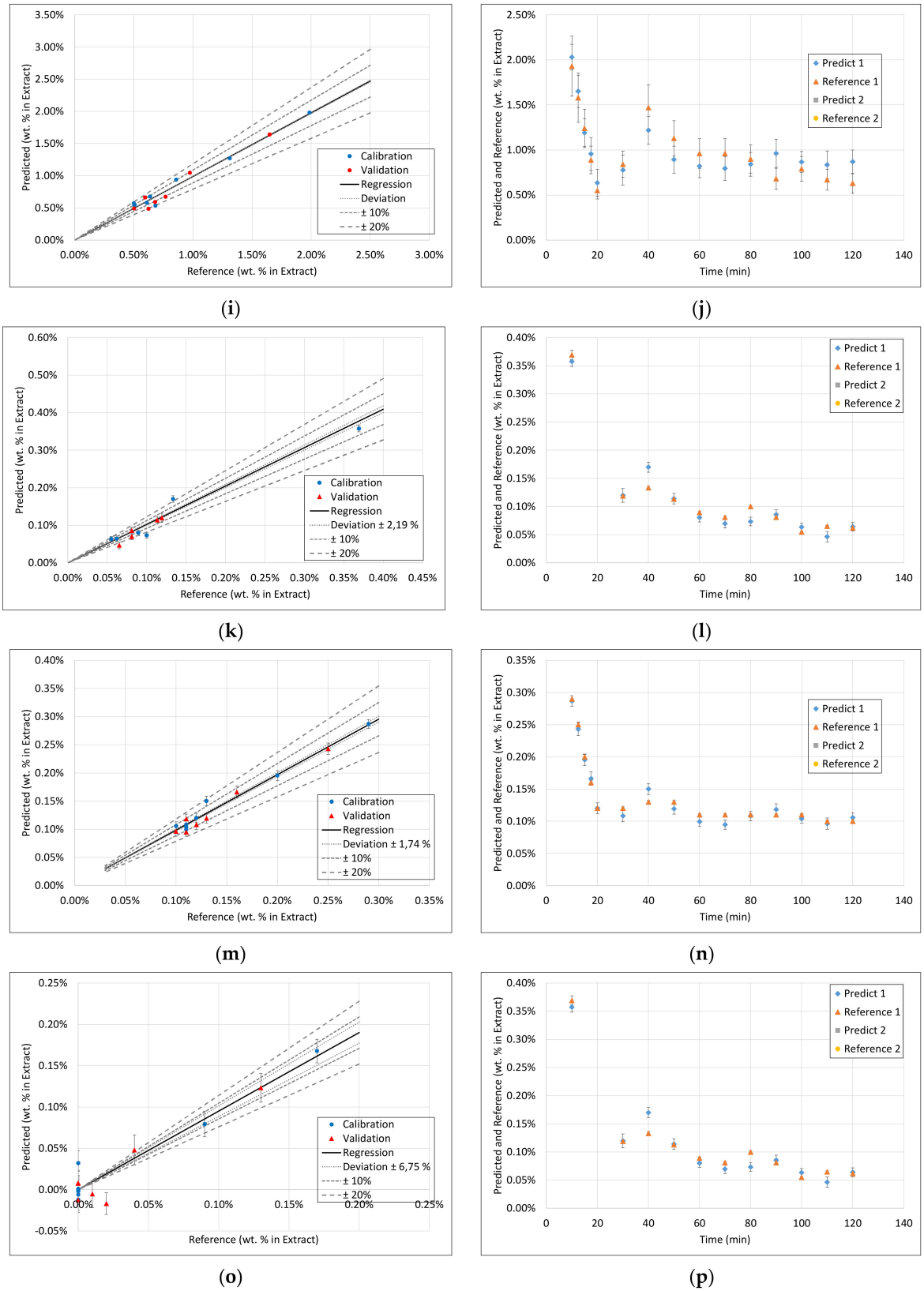


Figure A4. Fluorescence half-normal plots and measured vs. predicted extraction profiles for (**a**,**b**) dry residue, (**c**,**d**) arbutin, (**e**,**f**) total phenolic content, (**g**,**h**) sugar monomers, (**i**,**j**) structural carbohydrates, (**k**,**l**) aminic acids, (**m**,**n**) proteins, and (**o**,**p**) inorganics/ash.

References

1. Borges, A.; Abreu, A.; Dias, C.; Saavedra, M.; Borges, F.; Simões, M. New Perspectives on the Use of Phytochemicals as an Emergent Strategy to Control Bacterial Infections Including Biofilms. *Molecules* **2016**, *21*, 877. [CrossRef] [PubMed]
2. Leonel, M.; Sarmento, S.B.S.; Cereda, M.P. New starches for the food industry: Curcuma longa and Curcuma zedoaria. *Carbohydr. Polym.* **2003**, *54*, 385–388. [CrossRef]
3. D'Antuono, M.; Carola, A.; Sena, L.S.; Linsalata, V.; Cardinali, A.; Logrieco, A.F.; Colucci, M.G.; Apone, F. Artichoke polyphenols produce skin anti-age effects by improving endothelial cell integrity and functionality. *Molecules* **2018**, *23*, 2729.
4. Kleeberg, H. Method for the Production of Storage Stable Azadirachtin from Seed Kernels of the Neem Tree. U.S. Patent 5,695,763, 9 December 1997.
5. Fraunhofer-Projektgruppe für Wertstoffkreisläufe und Ressourcenstrategie. *Positionspapier zu Bioplastik*. 2018. Available online: https://www.iwks.fraunhofer.de/de/presse-und-medien/pressemeldungen-2018/positionspapier-zu-bioplastik.html (accessed on 3 December 2018).
6. DECHEMA. *Fortschrittliche Alternative Flüssig Brenn- und Kraftstoffe: Für Klimaschutz im Globalen Rohstoffwandel*; DECHEMA: Frankfurt, Germany, 2017.
7. Bundesministerium für Bildung und Forschung. *Nationale Forschungsstrategie BioÖkonomie 2030*; Herausgeber BMBF Referat Bioökonomie: Bonn/Berlin, Germany, 2010.
8. Ditz, R.; Gerard, D.; Hagels, H.; Igl, N.; Schäffler, M.; Schulz, H.; Stürtz, M.; Tegtmeier, M.; Treutwein, J.; Strube, J.; et al. *Position Paper of European Working Group on „Phytoextracts—Products and Processes": Proposal towards a New Comprehensive Research Focus*; DECHEMA: Frankfurt am Main, Germany, 2016.
9. Uhlenbrock, L.; Jensch, C.; Tegtmeier, M.; Strube, J. Digital Twin for Extraction Process Design and Operation. *Processes* **2020**, *8*, 866. [CrossRef]
10. Uhlenbrock, L.; Sixt, M.; Strube, J. Quality-by-Design (QbD) process evaluation for phytopharmaceuticals on the example of 10-deacetylbaccatin III from yew. *Resour. Effic. Technol.* **2017**, *3*, 137–143. [CrossRef]
11. ICH Expert Working Group. *Pharmaceutical Developement Q8(R2): ICH Harmonised Tripartite Guideline*; ICH: Geneva, Switzerland, 2009.
12. ICH Expert Working Group. *Riskmanagement (Q9): ICH Harmonised Tripartite Guideline*; ICH: Geneva, Switzerland, 2005.
13. ICH Expert Working Group. *Pharmaceutical Quality System (Q10): ICH Harmonised Tripartite Guideline*; ICH: Geneva, Switzerland, 2008.
14. Sixt, M.; Strube, J. Systematic and Model-Assisted Evaluation of Solvent Based- or Pressurized Hot Water Extraction for the Extraction of Artemisinin from Artemisia annua L. *Processes* **2017**, *5*, 86. [CrossRef]
15. Sixt, M.; Strube, J. Systematic Design and Evaluation of an Extraction Process for Traditionally Used Herbal Medicine on the Example of Hawthorn (Crataegus monogyna JACQ.). *Processes* **2018**, *6*, 73. [CrossRef]
16. Sixt, M.; Uhlenbrock, L.; Strube, J. Toward a Distinct and Quantitative Validation Method for Predictive Process Modelling—On the Example of Solid-Liquid Extraction Processes of Complex Plant Extracts. *Processes* **2018**, *6*, 66. [CrossRef]
17. Roth, T.; Uhlenbrock, L.; Strube, J. Distinct and Quantitative Validation for Predictive Process Modelling in Steam Distillation of Caraway Fruits and Lavender Flower Following a Quality-By-Design (QbD) Approach. *Processes* **2020**, *8*, 594. [CrossRef]
18. Uhlenbrock, L. Quality-by-Design zur Systematischen Entwicklung von Wertschöpfungsprozessen Pflanzlicher Rohstoffe. Ph.D. Thesis, TU Clausthal, Clausthal, Germany, 2021.
19. Zobel-Roos, S.; Schmidt, A.; Uhlenbrock, L.; Ditz, R.; Köster, D.; Strube, J. Digital Twins in Biomanufacturing. *Adv. Biochem. Eng. Biotechnol.* **2021**, *176*, 181–262. [CrossRef]
20. Sixt, M.; Gudi, G.; Schulz, H.; Strube, J. In-line Raman spectroscopy and advanced process control for the extraction of anethole and fenchone from fennel (Foeniculum vulgare L. MILL.). *C. R. Chim.* **2018**, *21*, 97–103. [CrossRef]
21. DAB. *Deutsches Arzneibuch*; Deutscher Apotheker Verlag: Stuttgart, Germany, 2017.
22. Williams, S. *Official Methods of Analysis of the Association of Official Analytical Chemists*, 14th ed.; Association of Official Analytical Chemists: Arlington, VA, USA, 1984.
23. Sluiter, A.; Hames, B.; Ruiz, R.; Scarlata, C.; Sluiter, J.; Templeton, D.; Crocker, D. *Determination of Structural Carbohydrates and Lignin in Biomass: Laboratory Analytical Procedure (LAP)*; Battelle: Columbus, OH, USA, 2005.
24. Baranska, M.; Roman, M.; Dobrowolski, J.C.; Schulz, H.; Baranski, R. Recent Advances in Raman Analysis of Plants: Alkaloids, Carotenoids, and Polyacetylenes. *CAC* **2012**, *9*, 108–127. [CrossRef]
25. Gierlinger, N.; Schwanninger, M. The potential of Raman microscopy and Raman imaging in plant research. *Spectroscopy* **2007**, *21*, 69–89. [CrossRef]
26. Holser, R.A. Principal Component Analysis of Phenolic Acid Spectra. *ISRN Spectrosc.* **2012**, *2012*, 1–5. [CrossRef]
27. Jović, O.; Habinovec, I.; Galić, N.; Andrašec, M. Maceration of Extra Virgin Olive Oil with Common Aromatic Plants Using Ultrasound-Assisted Extraction: An UV-Vis Spectroscopic Investigation. *J. Spectrosc.* **2018**, *2018*, 1–9. [CrossRef]
28. Mehrotra, R. Infrared Spectroscopy, Gas Chromatography/Infrared in Food Analysis. In *Encyclopedia of Analytical Chemistry*; Meyers, R.A., Ed.; John Wiley & Sons, Ltd: Chichester, UK, 2006; p. 177. ISBN 9780470027318.
29. Moldenhauer, M.; Sluchanko, N.N.; Buhrke, D.; Zlenko, D.V.; Tavraz, N.N.; Schmitt, F.-J.; Hildebrandt, P.; Maksimov, E.G.; Friedrich, T. Assembly of photoactive Orange Carotenoid Protein from its domains unravels a carotenoid shuttle mechanism. *Photosynth. Res.* **2017**, *133*, 327–341. [CrossRef] [PubMed]

30. Rys, M.; Szaleniec, M.; Skoczowski, A.; Stawoska, I.; Janeczko, A. FT-Raman spectroscopy as a tool in evaluation the response of plants to drought stress. *Open Chem.* **2015**, *13*, 21. [CrossRef]
31. Schulz, H.; Baranska, M. Identification and quantification of valuable plant substances by IR and Raman spectroscopy. *Vib. Spectrosc.* **2007**, *43*, 13–25. [CrossRef]
32. Trivedi, M.K.; Dahryn Trivedi, A.B.; Khemraj Bairwa, H.S. Fourier Transform Infrared and Ultraviolet-Visible Spectroscopic Characterization of Biofield Treated Salicylic Acid and Sparfloxacin. *Nat. Prod. Chem. Res.* **2015**, *3*, 1–6. [CrossRef]
33. Yuan, C.; Li, Y.; Li, Q.; Jin, R.; Ren, L. Purification of Tea Saponins and Evaluation of its Effect on Alcohol Dehydrogenase Activity. *Open Life Sci.* **2018**, *13*, 56–63. [CrossRef]
34. Sixt, M.; Strube, J. Pressurized hot water extraction of 10-deacetylbaccatin III from yew for industrial application. *Resour. Effic. Technol.* **2017**, *3*, 177–186. [CrossRef]
35. Plaza, M.; Turner, C. Pressurized hot water extraction of bioactives. *TrAC Trends Anal. Chem.* **2015**, *71*, 39–54. [CrossRef]
36. Schmidt, A.; Uhlenbrock, L.; Strube, J. Technical Potential for Energy and GWP Reduction in Chemical–Pharmaceutical Industry in Germany and EU—Focused on Biologics and Botanicals Manufacturing. *Processes* **2020**, *8*, 818. [CrossRef]
37. Sixt, M. *Methoden zur Systematischen Gesamtprozessentwicklung und Prozessintensivierung von Extraktions- und Trennprozessen zur Gewinnung Pflanzlicher Wertkomponenten*; Shaker Verlag: Aachen, Germany, 2018; ISBN 978-3-8440-6169-7.
38. Uhlenbrock, L.; Sixt, M.; Tegtmeier, M.; Schulz, H.; Hagels, H.; Ditz, R.; Strube, J. Natural Products Extraction of the Future—Sustainable Manufacturing Solutions for Societal Needs. *Processes* **2018**, *6*, 177. [CrossRef]
39. Bro, R. Exploratory study of sugar production using fluorescence spectroscopy and multi-way analysis. *Chemom. Intell. Lab. Syst.* **1999**, *46*, 133–147. [CrossRef]
40. Chakraborty, R.; Berglund, K.A. Steady state fluorescence spectroscopy of pyranine as a trace extrinsic probe to study structure in aqueous sugar solutions. *J. Cryst. Growth* **1992**, *125*, 81–96. [CrossRef]
41. Ghosh, N.; Verma, Y.; Majumder, S.K.; Gupta, P.K. A Fluorescence Spectroscopic Study of Honey and Cane Sugar Syrup. *FSTR* **2005**, *11*, 59–62. [CrossRef]
42. Esbensen, K.; Swarbrick, B.; Westad, F.; Anderson, M.; Whitcomb, P. *Multivariate Data Analysis: An introduction to Multivariate Analysis, Process Analytical Technology and Quality by Design*, 6th ed.; CAMO Software AS: Magnolia, TX, USA, 2018.
43. Dingermann, T.; Hiller, K.; Schneider, G.; Zündorf, I. *Schneider Arzneidrogen*, 5th ed.; Spektrum Akad. Verl.: Heidelberg, Germany, 2011; ISBN 978-3-8274-2765-6.
44. Lohmann, L.J.; Strube, J. Process Analytical Technology for Precipitation Process Integration into Biologics Manufacturing towards Autonomous Operation—mAb Case Study. *Processes* **2021**, *9*, 488. [CrossRef]
45. Vetter, F.L.; Zobel-Roos, S.; Strube, J. PAT for Continuous Chromatography Integrated into Continuous Manufacturing of Biologics towards Autonomous Operation. *Processes* **2021**, *9*, 472. [CrossRef]
46. Rolinger, L.; Rüdt, M.; Diehm, J.; Chow-Hubbertz, J.; Heitmann, M.; Schleper, S.; Hubbuch, J. Multi-attribute PAT for UF/DF of Proteins-Monitoring Concentration, particle sizes, and Buffer Exchange. *Anal. Bioanal. Chem.* **2020**, *412*, 2123–2136. [CrossRef]
47. Rolinger, L.; Rüdt, M.; Hubbuch, J. A critical review of recent trends, and a future perspective of optical spectroscopy as PAT in biopharmaceutical downstream processing. *Anal. Bioanal. Chem.* **2020**, *412*, 2047–2064. [CrossRef] [PubMed]
48. Thakur, G.; Thori, S.; Rathore, A.S. Implementing PAT for single-pass tangential flow ultrafiltration for continuous manufacturing of monoclonal antibodies. *J. Membr. Sci.* **2020**, *613*, 118492. [CrossRef]
49. Feidl, F.; Garbellini, S.; Vogg, S.; Sokolov, M.; Souquet, J.; Broly, H.; Butté, A.; Morbidelli, M. A new flow cell and chemometric protocol for implementing in-line Raman spectroscopy in chromatography. *Biotechnol. Prog.* **2019**, *35*, e2847. [CrossRef] [PubMed]
50. Grillo, G.; Boffa, L.; Binello, A.; Mantegna, S.; Cravotto, G.; Chemat, F.; Dizhbite, T.; Lauberte, L.; Telysheva, G. Analytical dataset of Ecuadorian cocoa shells and beans. *Data Brief* **2019**, *22*, 56–64. [CrossRef]
51. Zeise, I.; Heiner, Z.; Holz, S.; Joester, M.; Büttner, C.; Kneipp, J. Raman Imaging of Plant Cell Walls in Sections of Cucumis sativus. *Plants* **2018**, *7*, 7. [CrossRef] [PubMed]
52. Sixt, M.; Schmidt, A.; Mestmäcker, F.; Huter, M.; Uhlenbrock, L.; Strube, J. Systematic and Model-Assisted Process Design for the Extraction and Purification of Artemisinin from Artemisia annua L.—Part I: Conceptual Process Design and Cost Estimation. *Processes* **2018**, *6*, 161. [CrossRef]