Universidade de Lisboa

Faculdade de Farmácia / Instituto Superior Técnico

Optimizing Blending Processes: Establishing and Refining NIR and Transmission Raman methods

Jacinta Dias Ribeiro

Dissertation supervised by Professor João Almeida Lopes and co-supervised by Dr. Stefan Busche.

Master's in Pharmaceutical Engineering

2024

Universidade de Lisboa

Faculdade de Farmácia / Instituto Superior Técnico

Optimizing Blending Processes: Establishing and Refining NIR and Transmission Raman methods

Jacinta Dias Ribeiro

Dissertation supervised by Professor João Almeida Lopes and co-supervised by Dr. Stefan Busche.

Master's in Pharmaceutical Engineering

2024

Acknowledgments

I would like to express my deepest gratitude to Dr. Stefan Busche for his constant sup and patience during my time at Merck, as well as for the rigorous manner in which he revie[illegible] my thesis. I'm also extremely grateful to Professor João Almeida Lopes for the invaluable feedback and for providing me with this opportunity.

I extend my sincere thanks to my co-workers, who helped me through the different stages and taught me so much.

Lastly, I'd like to mention my family and friends, especially my parents and brother. Without their support and encouragement, I would not have been able to achieve this. They believed in me and pulled me through the toughest times. To my grandparents, Paulino and Adelina, who have always been there for me, and to my grandmother Clara, my guiding star, who I know was cheering for me every step of the way.

Statement

I declare that this document is an original work of my own and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Name: Jacinta Dias Ribeiro Student number: 71837

Lisboa, 30/09/2024

Abstract

Ensuring blend uniformity of pharmaceutical powders and granules is crucial in obtaining pharmaceutical products, especially solid pharmaceutical forms, with quality and in accordance with regulatory criteria. Near-Infrared Spectroscopy and Transmission Raman Spectroscopy have emerged as methods that enable the determination of when the mixture is homogeneous in a non-destructive, fast and effective way. The main objectives of this work were to use Near-Infrared Spectroscopy as a tool for monitoring the mixture in real time and to evaluate the sensitivity of this technique in detecting small variations in the formulation. Also to use Transmission Raman Spectroscopy to quantify API in samples and evaluate its potential as an alternative analytical reference technique for quality control.

In order to achieve these objectives investigation with two different formulations were carried out. First a Partial Least Squares model was developed and validated using caffeine as the model substance and then this same approach was applied to another formulation using an API from the development pipeline.

Near-Infrared Spectroscopy was able to identify small variations down to 0.5% in the caffeine formulations. The Partial Least Squares models developed were robust and allowed changes to be made to the batch size while maintaining a good prediction capability for blend uniformity. However, there are statistically significant differences between the predictions and the HPLC/UPLC results over wide concentration ranges.

As for Transmission Raman Spectroscopy, it made good predictions for API quantification in both formulations and was in line with expectations when compared with the HPLC/UPLC results. Transmission Raman Spectroscopy predictions showed no significant difference to HPLC/UPLC results, proving this method to give closer results to the latter reference analytics techniques and to be a more accurate and reliable method than Near-Infrared Spectroscopy.

Key words: Near-infrared spectroscopy, Transmission Raman spectroscopy, Blend Uniformity, Bulk powder

Resumo

Garantir a uniformidade da mistura de pós e grânulos farmacêuticos é crucial na obtenção de produtos farmacêuticos, especialmente, formas farmacêuticas sólidas, com qualidade e de acordo com os critérios regulatórios. A espetroscopia de infravermelho próximo e a espetroscopia de Raman por transmissão têm surgido como métodos que permitem alcançar e determinar quando a mistura está homogénea de forma não destrutiva, rápida e eficaz. Os principais objetivos deste trabalho foram utilizar a espetroscopia de infravermelho próximo como ferramenta de monitorização da mistura em tempo real e avaliar a sensibilidade desta técnica na deteção de pequenas variações na formulação. Assim como usar a espetroscopia de Raman por transmissão para quantificar a substância ativa em amostras e avaliar o seu potencial como uma técnica alternativa de referência analítica para controlo de qualidade.

De modo a alcançar estes objetivos foram realizados estudos com duas formulações diferentes. Primeiro desenvolveu-se e validou-se um modelo de Mínimos Quadrados parciais usando a cafeína como substância modelo e posteriormente aplicou-se esta mesma abordagem a outra formulação utilizando uma outra substância ativa, cuja identidade é confidencial.

A espetroscopia de infravermelho próximo mostrou conseguir identificar pequenas variações até 0.5% nas formulações de cafeína. Os modelos de Mínimos quadrados parciais desenvolvidos foram robustos, permitiram que mesmo com alterações ao tamanho do lote se mantivesse uma boa capacidade preditiva para determinar a uniformidade da mistura. Embora sejam de realçar, estatisticamente, diferenças significativas entre as previsões e os resultados de HPLC/UPLC em grandes intervalos de concentração.

Quanto à espetroscopia de Raman por transmissão, permitiu fazer boas previsões na quantificação da substância ativa em ambas as formulações estando de acordo com expetativa quando comparado com os resultados de HPLC/UPLC. De forma geral, as previsões obtidas através do modelo de Mínimos quadrados parciais para o Raman de transmissão mostraram não ser estatisticamente diferentes dos resultados de HPLC/UPLC, garantindo que este é o método que mais se aproxima destas técnicas de referência analíticas e a partir do qual de obtêm resultados mais exatos e fidedignos.

Palavras-chave: Espetroscopia de Infravermelho próximo, Espetroscopia Raman de Transmissão, Uniformidade da mistura, Pó farmacêutico

Index

List of tables and figures

List of figures

List of tables

Introduction

Although every unit operation is important in its own way, powder blending is critical in the manufacturing of solid dosage forms. This places a significant burden on manufacturers to monitor and control the uniformity of blends throughout the production process. To ensure blend uniformity is to ensure quality, safety and efficacy in the final product and compliance with the regulatory standards since deviations in blend composition can lead to variations in product performance and may jeopardize patient safety. Non-destructive analytical techniques, such as Near-Infrared Spectroscopy (NIR) and Transmission Raman Spectroscopy (TRS), have emerged as invaluable tools in the pursuit of reliable blend monitoring in detriment of more common off-line techniques. This master's thesis focuses on blend monitoring, with a particular emphasis on the sensitivity analysis of NIR in detecting subtle formulation changes and evaluates Transmission Raman as an alternative technique to be used as reference analytics for quantification of APIs.

In development and manufacturing of solid dosage forms, blend uniformity (BU) is among the most important critical quality attribute (CQA) and is a necessary prerequisite for delivering content uniformity [1]. Only a homogeneous mixture can be subdivided into individual doses to provide the correct proportions of the active compound [2]. Inhomogeneity and a skewed active pharmaceutical ingredient (API) to excipient ratio could affect not only drug bioavailability and efficacy, but potentially induce an adverse response in patients [1]. Therefore, blending uniformity ensures low dosage variability and high quality of the product [3]. For powder-blending processes the ultimate goal is to obtain a uniform blend with a quality close to optimal, in typical applications the relative standard deviation (RSD) should be below 5% [3].

The development of dry mixing processes is complex [4], so knowledge on blend uniformity and the related process design space is crucial. The process design includes several variables such as the characteristics of the blending vessel, like its design; the operating conditions of the blender, such as blending time, fill level, filling procedure, mixing speed, mixer type; the physical properties of the mixture components, such as particle-size distribution (PSD) or cohesivity. The latest shows a strong effect on flow patterns, segregation tendency and agglomeration. Finally, the environmental conditions, such as humidity also play a role [3, 5]. All these factors influence blend uniformity and therefore will have an impact on the quality of the final product.

The scale up of the blending operation can be difficult, leading to challenges in industrial applications. Problems related to content uniformity such as weight variability of the finished dosage form, particle segregation, particle agglomeration and poor equipment design and operation, can lead to non-uniform blends of the API and the excipients and may be hard to resolve [3]. Thus, methods to monitor the various processing steps are critical for maintaining high product quality and eliminating the risk of producing out-of-specification products [3].

Among various types of blending techniques there´s two that are more common: free fall and high shear mixers, each of these have their advantages and disadvantages [1]. In this work it was used a bin blender for the blending trials which is a special type of free fall mixer, so its highlights and downsides will be pointed out in more detail. Some of the positive points are enabling a slow rotating speed, ensuring uniform, gentle and low impact blending, helping to prevent unwanted destruction of granules. However, presents some drawbacks like that particle size distribution (PSD) and morphology of excipients and API need to be similar to avoid segregation – poorly flowable materials are harder to blend [1].

The goal of Quality by design (QbD) is to build quality into the product through the process, and to achieve that it is necessary to have knowledge on both the product and the process. The principle is that if critical sources of variability can be understood, then product performance can be controlled using the manufacturing process to mitigate variability in the material properties. [6, 7]. This approach is a bit different from the ones previously used where the process is developed, and the product quality is checked at the end. The ICH Q8[8] guideline, describing the principles of Quality by Design, emphasizes the application of appropriate risk management tools to guide development efforts.

Based on the literature reviewed [3, 7], the key parameters that affect the blending unit operation the most and therefore will have a bigger impact on final product quality were identified. An Ishikawa Diagram was designed to highlight the major problems as presented in *Figure 1*.

The risk analysis considers the different parameters that can affect the outcome, in this case, of the blending operation. It can be grouped roughly in four groups, namely raw materials characteristics, environment, equipment and operating conditions. For example, if the blending time is too long depending on the particle characteristics, may occur separation which then affects content uniformity of the mixture. If the blending speed it's too elevated may cause the particles to adhere to the wall due to centrifugal force which once again may affect content uniformity of the mixture [9].

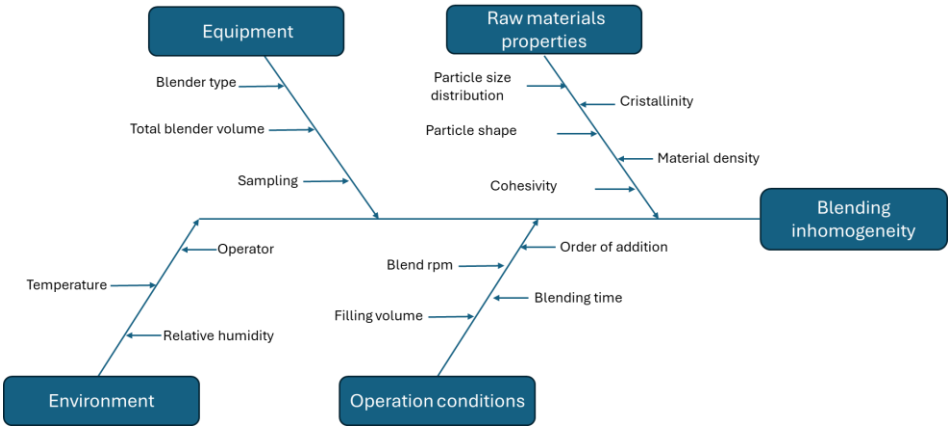


Figure 1: Ishikawa diagram for blending unit operation.

Currently the companies are slowly transitioning to continuous manufacturing and with that implementing new methods to monitor and control unit operations. These arise as an alternative to the previous analytical methods used to assess blend uniformity which are time-consuming [2, 10, 11], involve stopping the procedure taking a sample and analyzing it off-line. Typically, HPLC is the method used, which may lead to error such as loss of product in transfers from one container to other and consequently lead to mistakes in the results obtained, could potentially introduce contamination and segregation [10–12]. Moreover, the time difference between the end of the blending process and the results of the reference analytics not only prevent its release in real time but also can have a negative impact on quality. In addition, thief sampling can disturb the blend and be prone to biased sampling [11].

In parallel with continuous manufacturing, Process Analytical Technology (PAT) emerges as the main tool to achieve Quality by Design [13]. This control strategy can be defined as a mechanism to design, analyze, and control pharmaceutical manufacturing processes through the measurement of critical process parameters (CPPs) which influence CQAs [6].

Among several PAT tools to monitor processes, two of the most prominent ones are Near-Infrared Spectroscopy and Transmission Raman Spectroscopy. Near infrared spectroscopy and chemometrics provide a useful methodology to assess the blend homogeneity in real-time and to support the real time release of the batches[11].

The Near-Infrared region (780–2500 nm) is located between the red band of the visible light and the mid infrared region. The NIR signal is a consequence of the absorbance of light due to molecular vibrations (overtones and combinations of fundamental vibrations) of hydrogen bonds like C-H, N-H, O-H [14].

Near-Infrared spectroscopy has become a prominent non-destructive analytical technique, offering real-time insights into the chemical composition of substances [10, 14]. Some of the advantages include minimal sample preparation, broad applicability, no use of organic solvents and rapid analysis, making it an attractive choice for monitoring blend uniformity [10, 14, 15]. Thus, the low molar absorptivity of NIR bands permits operation in the reflectance mode and hence recording of spectra of solid samples with minimal or no pre-treatment, thereby substantially increasing the throughput. Also, the dual dependence of the analytical signal on the physical and chemical nature of the sample facilitates both its identification and the determination of physical and chemical parameters [16]. As with any analytical technique, NIR has its shortcomings namely its high detection limit.

The major pitfall of this technique is the complexity of the spectra due to their nature (overtones and combination bands of vibrational energy levels). In a complex matrix, many of these transformations can occur resulting in a broad range of possibly overlapping absorption bands. This makes it difficult to obtain relevant information about the feature of interest using measurements at only one wavelength from the raw spectra. Furthermore, physical conditions of sample and measuring environment also influence the spectra, making it even more complex to interpret the data[14].

NIR usually is not used as a direct analysis technique. It generally requires a correlation to be established, and for this, a reference method to assess the property of interest is required (often HPLC or UPLC) To extract the useful information that one wants to obtain (predict) from the spectra, e.g. the concentration of an active ingredient, chemometric techniques are applied [14].

NIR spectroscopy is frequently applied for the raw materials identification and has revealed itself very useful for on-line/in-line monitoring in several operations in the pharmaceutical industry

including determination of blend uniformity and specifically its endpoint by both qualitative and quantitative models of different pharmaceutical products [1, 10, 11]. Numerous qualitative and quantitative methods have been developed to help determine when a blending operation should be terminated. The endpoint is usually specified as when a particular statistic, qualitative or quantitative, fulfills a threshold-related condition or remains constant over a given number of consecutive blending observations [17].

Quantitative models [11] give a more precise prediction but the method's validation is time consuming and slow down the implementation of PAT tools, whereas qualitative models need less calibration burden and data sampling, offering a simpler way to build and validate the blend uniformity by NIR in early process development stages [4, 11, 13]. Ideally, a BU application for process development and up-scaling should require limited development work. In addition, for technology transfer and routine production, robustness, transferability and limited maintenance are prime considerations. The method should perform well for different batch sizes and blender types. Quantitative methods often require re-calibration in these cases. Selection of adequate calibration samples and measurement may also be challenging since it is very unattractive (from a time and cost perspective) to prepare calibration samples on the actual batch size scale(s). Use of static calibration samples typically fails to adequately incorporate the influence of scale, powder density and rheology, compromising robustness and reliability. Quantitative methods may also suffer from variation in ingredient properties (e.g., particle size or moisture content). While technically possible to incorporate these in a model, the extensive additional work is usually prohibitive. A practical and reliable qualitative method should be applicable at various stages of development and manufacturing without much modeling investment and should readily provide information regarding effects of variation in ingredient properties, process settings and environmental conditions on blend uniformity [4].

Several pharmaceutical applications of qualitative NIR spectroscopy include identity and quality testing of raw materials, detection of tablet and capsule tampering, detection of tablet degradation, analysis of parenteral products, and determination of ointment homogeneity. The capability of NIR to validate powder mixing has also been demonstrated [15].

Regarding qualitative methods, principal component analysis (PCA) and moving block standard deviation (MBSD) are the most commonly used. Both these approaches compress NIR data and monitor spectral variance over time during the blending process [11–13]. Other examples

of qualitative analysis include dissimilarity index, soft independent modeling of class analogies, Hotelling's T-square statistic and bootstrap error-adjusted single-sample techniques. The development of these qualitative methods is efficient as little to no calibration runs are required, although suitable thresholds for stopping the blend must be determined and justified [12].

Some methods like MBSD and moving block F-test compare spectral variance between adjacent block of spectra throughout the blending process [11, 17]. For example, MBSD is calculated by monitoring the pooled standard deviation across wavelengths between time points, then blend endpoint is determined when standard deviation reaches a minimal value that is predefined [12, 13]. The data are organized into a matrix of time versus wavelength. A new matrix is then created by calculating the standard deviation of intensity values over a set time-window or block for each wavelength. After that, the average of these standard deviation values across all wavelengths is calculated. This average standard deviation is plotted over time, and the blending endpoint is identified as the point where this average reaches its minimum value [12, 13]. However, MBSD does not ensure that a possible absence of the API or some excipients will be noticeable.

A qualitative method for assessing blend uniformity that overcomes the drawbacks of other qualitative techniques while ensuring both ease of use and reliability: the moving F-test. The method evaluates the spectral variance (SV) ratio (F-value) of two adjunct blocks of NIR spectra via the standard F-test. SV within a block typically declines over time and stabilizes when 'uniformity' is achieved. To confirm the 'equal variance hypothesis' with a small chance for false rejects, the F-value should be lower than a critical value *Fcrit.* An advantage is that the threshold value is defined statistically, not empirically and thus does not suffer from threshold ambiguities as with the MBSD. The F-test method has important practical advantages such as sensitivity is easily determined and optimized without extensive calibration, information on excipient uniformity is easily obtained, the method has relatively short development time, is essentially maintenance free and easily transferable to different blenders [4].

The additional information provided by quantitative NIRS methods represents an advantage over qualitative approaches. Quantitative predictions from regression models such as partial least squares (PLS) are based on monitoring the concentration of each component over the course of the blend process and then relating that to NIR powder spectra [12, 13]. The blend process is considered complete when the deviations fall within a pre-specified limit for a certain amount of time.

The changes in homogeneity over a period of time and the deviation of the prediction from target concentration are simultaneously assessed. However, end-point criteria based solely on the evolution of a single compound, usually the active ingredient, may not be sufficient when considering the homogeneity of a global pharmaceutical formulation. When the excipients' distributions are also of interest to obtain a particular drug release profile, it may be more relevant to implement a quantitative approach predicting all the compounds of interest, or a qualitative method assessing the global homogeneity of the powder blend [17].

A calibration design, spanning relevant variables such as component concentration and scale, is required. The main variable under consideration is the concentration of the components, particularly the active pharmaceutical ingredient. Additionally, interfering variables like blend scale are included in the calibration design to improve model robustness. Conducting multiple calibration blends using industrial-scale blenders is inefficient and may not be practical.

The continuous sampling provided by NIRS allows the investigation of two blending process characteristics: the process trajectory and the process endpoint.

Process trajectory refers to the trends observed in the evolution of active pharmaceutical ingredient and excipients homogeneity as a function of time during mixing. Because of variable powder handling and dispensing, blender loading pattern differences, and raw material variability (i.e., lot-to-lot, relative humidity), no two blends will present the exact same mixing kinetics. The consequences will be that until homogeneity is reached, different blends of the same formulation will not necessarily have the same composition at a given period of time. Knowledge of this variability will be of great value to understand better the natural process variations that can be expected in normal operating conditions. In addition, abnormal process behaviors may be identified, and the related trajectories may inform about the root causes of the process anomaly.

Process endpoint, on the other side, refers to the time point at which a satisfactory homogeneity level has been reached for a particular product. Due to the various sources of process variability potentially impacting the process trajectory, no two blends will reach homogeneity at precisely the same time. In addition, the conditions for use of a certain algorithm employed to check for homogeneity may lead to different answers related to the end-point detection (even for a same blend) [17]. Indeed, an end-point will be affected by the number of observations/predictions taken into account (period of time considered to estimate homogeneity), the number of components considered (API alone or global formulation), the parameter used for end-point detection (a

particular wavelength intensity, API prediction, global formulation predictions, etc.), and the criterion used to stop the unit operation (related to the closeness to a nominal concentration value, to the relative variation among consecutive predictions, predictions within confidence limits, predictions following a particular statistical distribution, etc.)[17].

There are two key calibration issues that must be addressed if a practical NIR analysis is to be developed. First, the instrumental configuration and sampling interface must be designed to provide stable spectral measurements with minimal variation. Second, the requirements for the collection of calibration data must be practical from the standpoint of time and cost [10].

It is now generally accepted that with suitable interfacing, robust analysis methods and proper interpretation, in-line NIR offers major advantages in reliability, efficiency and process understanding compared to traditional sampling. Availability of PAT data on process performance and variability offers opportunities to apply continuous process improvement strategies. From a risk perspective, direct blend uniformity control eliminates criticality of raw material properties, process parameters and their interactions therefore these characteristics stop having such a big impact in the overall final characteristics of the product and stop being that important for the process [4]. This can strongly reduce development efforts and input control requirements in commercial production. However, translating in-line NIR data to BU remains challenging. The development efforts for a sufficiently robust method are demanding. This applies to quantitative methods such as Partial Least Squares (PLS) models, achieving a robust calibration is the main challenge. Quantitative models may also have too large prediction errors for low-dose formulations, i.e., those with the most significant BU risk, for which monitoring and control are most needed.

Although NIR is more established as an analytical tool in pharmaceutical process and quality control, its limited chemical specificity restricts a broader applicability [18]. Transmission Raman spectroscopy arises as an alternative to overcome that restriction, as it provides much higher chemical specificity which is useful in more complex formulations [18]. TRS typically exhibits excellent specificity with many sharp and distinct features that can be assigned readily to individual components [19, 20].

Raman spectroscopy has different variants depending on where the collection and detection fibers are. To simplify, just the conventional Raman and Transmission Raman will be addressed [21].

In conventional backscatter Raman the signal collected is mainly from the superficial layers, either by scanning the sample's surface point-by-point or by a single wide-area illumination measurement. Both approaches assume the signal from the surface is representative of the whole sample and this cannot be considered true. When using real pharmaceutical formulations uniformity at a microscopic level may not be fully accomplished [18, 21, 22].

Meanwhile in the transmission method, the laser scatters through the sample generating a better approximation of the bulk-averaged signal, although there is still a slight bias toward the center and away from the edges [18, 22]. As shown in *Figure 2*, the laser illuminates the sample over an extended zone on one side and the signal is collected on the other[18]. By providing information on the volumetric content of the samples, this method reduces differences between TRS predictions and HPLC reference data, suppresses fluorescence from superficial layers and reduces sensitivity to matrix effects such as particle size and compaction force compared to other spectroscopic methods (particularly NIR spectroscopy) [18, 21–23]. TRS is commonly used as an off-line PAT [21].

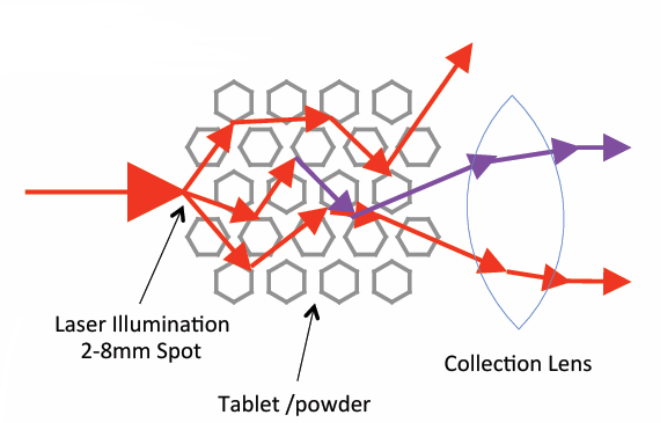


Figure 2: TRS principle [24].

The analytical versatility of Raman molecular spectroscopic techniques allows for investigating a wide diversity of transparent, translucent, opaque, and colored samples including solid dosage forms like tablets, capsules as well as powders [24, 25].

This method has many advantages similar to NIR, such as its non-destructive nature, simple workflow, no sample preparation necessary, no solvents used in the process, no waste to dispose of and no consumables [23, 24]. One of the biggest perks is undoubtedly the simple workflow, this

technique only needs a tray to be loaded with samples requiring little skill and time. Also to change between product methods, a new tray is loaded and a different method selected, meaning many different methods/products can be run, one after another on the same instrument, with very little downtime in between [24]. TRS also is very useful to analyze aqueous samples, as water is a weak Raman scatter resulting in negligible interference with the spectrum whereas IR spectra are strongly influenced [25, 26]. The technique's limitations include the inability to analyze uncomplexed ionic compounds (e.g. NaCl) and interference from fluorescence in cases where this overwhelms the Raman signal [27].

TRS is an optical technique that provides information on various aspects like identity, purity, and quantity as well as chemical composition and molecular structure of drugs and excipients. However, although the current Ph. Eur. edition recommended Raman spectroscopy in general chapters as an alternative technique to control microbiological quality, polymorphism, crystallinity and chemical imaging, up to now the Raman technique is not referred to in any single special monograph for the characterization of substances or formulations. Instead, infrared spectroscopy (IR) is the method of choice for identification and quantification of substances under standard conditions [24, 26]. So, although these techniques have been suggested as complementary, this last section showed why TRS can be an asset and improve content uniformity and quantitative analyses like API quantification.

TRS spectra are quite complex containing multiple peaks of the various Raman-active compounds in a sample, which overlap with each other [19]. Multivariate analysis is useful to analyze this data, the most common models include PCA or PLS that were already detailed in the NIR section of this introduction.

OBJECTIVES

The aim of this thesis is the monitoring of blend uniformity through Near-Infrared Spectroscopy focusing on the sensitivity of NIR to detect subtle concentrations changes in formulations. Also investigate Transmission Raman Spectroscopy as a HPLC alternative by developing a PLS method for API quantification to be used as reference analytics.

In order to achieve this goal, several trials were conducted and supported by off-line analysis, this will be explained in more detail in the next section.

2. Methods

This section delineates the experimental procedures applied to achieve the objectives of the study. The approach utilized to design, execute, and analyze the experiments is described to ensure reproducibility.

The experiments performed for the caffeine formulation are designated as (A) following by a respective number from lower to higher concentration, for example, the trial with 10% caffeine is A1, 12% is A2 and so on. The experiments performed with the API A from the development pipeline (hereinafter called API A) formulation are designated as (B) and the respective number as well. Regarding the validation experiments, are designated as AV and BV, respectively.

2.1. Blend composition

2.1.1. Blend composition of caffeine calibration experiments

For the investigation of the blending process with the Viavi MicroNIR PAT-W nine different blends with increasing API (caffeine) concentration from 10 to 20% drug load were manufactured. A large concentration range around the target formulation (15%) and ensure that all quantitative predictions would be reliable, preventing model extrapolation. The intended batch size was 1.2 kg. The changing concentration of caffeine in the batches was achieved by compensating in the amount of mannitol, so both components are negatively correlated.

To exemplify, the formulation of experiment A1 with 10% caffeine is shown in Table 1. 10 % of mannitol was substituted with caffeine. The tables with the remaining formulations are presented in the Appendix.

Table 1: Formulation of experiment A1.

Name	***Function***	***Content [%]***
Caffeine	API	10.00
Mannitol	Filler	67.26
MCC	Filler	17.56
PVP K 30	Binder	4.68
Magnesium stearate VS	Lubricant	0.50
Total		100.0

2.1.2. Blend composition of caffeine validation experiments

The validation trials were all conducted with a 15% drug load using the same formulation as the calibration trials.

Table 2: Formulation of experiments AV1 to AV4.

Name	***Function***	***Content [%]***
Caffeine	API	15.00
Mannitol	Filler	62.26
MCC	Filler	17.56
PVP K 30	Binder	4.68
Magnesium stearate VS	Lubricant	0.50
Total		100.0

2.1.3. Blend composition of API A calibration experiments

For the investigation of the blending process by the Viavi MicroNIR PAT-W five different blends with increasing API (API A) concentration from 18.94 to 35.17% drug load were manufactured. A large concentration range around the target formulation (27.05%) and ensure range and robustness. The intended batch size was around 750 g. The increasing concentration of API is achieved by substituting in the amount of both fillers, mannitol and cellulose microcrystalline. The ratio of mannitol/MCC is not kept constant because in order to build a more robust PLS model the excipients, and API should not be correlated.

Table 3: Formulation of experiments B1 to B5.

	B1	*B2*	*B3*	*B4*	*B5*
API [%]	18.94	22.99	27.05	31.11	35.17
Ratio MCC/Mannitol	1.44	1.53	1.50	1.48	1.37
Desintegrant / Lubricant / Flow agent	about 10 %	about 10 %	about 10 %	about 10 %	about 10 %

2.1.4. Blend composition of API A validation experiments

The validation trial was performed using a drug load of 27.05% and a batch size of 10 kg, so although the formulation is the same, quantities were adjusted.

2.2. Blend manufacturing process

2.2.1. Blend manufacturing process of caffeine calibration experiments

The raw materials were weighed individually and sieved through a 1.0 mm hand sieve. The magnesium stearate, caffeine and MCC were premixed in a polybag and sieved together. Polyvinylpyrrolidone K 30 was not sieved because was already very fine and no agglomerates were visible. Mannitol was sieved alone. Then the raw materials were added to the 5-liter container, in a sandwich scheme, in the following order:

1. Premix of caffein, MCC and magnesium stearate
2. Polyvinylpyrrolidone K 30
3. Mannitol

For a batch size of 1.2 kg, a filling level of the container of around 40-50% was observed for all batches. The filled 5-liter container was then mixed for 30 minutes at 12 rpm on the Servolift container blender.

2.2.2. Blend manufacturing process of caffeine validation experiments

In the validation experiments, the raw materials were weighted, sieved and filled into the container in the same manner as in the calibration trials. Then the blending process proceeded as follows for each trial,

a) Experiment AV1: Stop the blending when f-value is below threshold for one minute using a block size of 20, at 12 rpm.
b) Experiment AV2: Blending for 30 minutes at 12 rpm.
c) Experiment AV3: Blending for 3 minutes at 12 rpm, to verify that it is inhomogeneous.
d) Experiment AV4: Increasing the filling level to 75% and blending for 30 minutes at 12 rpm.

2.2.3. Blend manufacturing process of API A calibration experiments

The raw materials were weighed individually and sieved through a 1.0 mm hand sieve. Mannitol and flow agent were premixed in a polybag and sieved together. MCC and disintegrants were premixed in a polybag and sieved together. API and lubricant were sieved alone. Then the raw materials were added to the 5-liter container, in a sandwich scheme, in the following order:

1. Premix of Mannitol/flow agent
2. API
3. Premix of MCC /disintegrants

For a batch size of 750 g, a filling level of the container of around 40% was observed for all batches. The filled 5-liter container was then mixed for 10 minutes at 12 rpm on the Servolift container. The blending was stopped, the magnesium stearate was added and then blended for another 5 minutes at 12 rpm.

2.2.4. Blend manufacturing process of API A validation experiment

The validation trial was carried out in the same manner as calibration trials.

2.3. NIR spectra acquisition during the blending process

During the blending process, spectra were recorded through a lid with a sapphire window using the Viavi NIR spectrometer. In API A experiments NIR spectra were recorded in both blending steps.

With each rotation of the container one spectrum is recorded, which means that for a 30-minute blend at 12 rpm 360 spectra are expected. In the experiments A1, A2, A3, A5, A6 and A7 the NIR spectrometer didn´t record a spectrum at every rotation resulting in a lower number of spectra. In caffeine validation trials no bad spectra were observed and therefore none were deleted. In both B2 and B5 not every rotation recorded one spectrum resulting in a lower amount at the end. One bad spectrum will nearly always be collected because at the end of the blending time when the container blender is stopping, the last rotation is made in a very slow speed causing that no material is in front of the sapphire window. Other bad spectra may result from the measuring starting or finishing sooner than it´s supposed to and therefore there´s no powder in the sapphire window. The bad spectra were then removed before further analysis of the data. The overview is presented in Tables 4 to 7.

Table 4: Overview of total spectra and bad spectra for caffeine calibration experiments.

Experiment	***Date***	***Total spectra collected***	***Spectra to be deleted***
A1	22.02.2014	357	0
A2	22.02.2024	357	1
A3	22.02.2024	356	4
A4	22.02.2024	361	1
A5	23.02.2024	352	1
A6	23.02.2024	354	1
A7	23.02.2024	357	7
A8	23.02.2024	361	1
A9	23.02.2024	360	1

Table 5: Overview of total spectra and bad spectra for API A calibration experiments.

Experiment	***Date***	***Total spectra collected***	***Spectra to be deleted***
B1	10.06.2024	182	2
B2	10.06.2024	174	1
B3	10.06.2024	181	2
B4	10.06.2024	181	1
B5	10.06.2024	176	1

Table 6: Total number of spectra recorded for caffeine validation experiments.

Experiment	***Total number of spectra***
AV1	67
AV2	303
AV3	37
AV4	360

Table 7: Overview of total spectra and bad spectra for API A validation trial.

Experiment	*Date*	*Total spectra collected*	*Spectra to be deleted*
BV1	26.06.2024	182	1

The spectra collected during the blending process were evaluated by PCA using Unscrambler X 10.5. The following steps were executed to pretreat the spectra of both formulations:

- Importing the raw spectra from the MicroNIR pro 3.2 software to Unscrambler X 10.5 software
- Deleting the bad spectra from each batch
- Preprocessing the raw spectra by using the Savitzky-Golay derivative with following parameters: (Derivative Order :1 / Polynomial Order :2 / Smoothing Points :5 / Left Points :2 / Right Points :2 / Range: columns 1-125)
- Preprocessing the spectra by using the standard normal variate (SNV) with following parameters: (Range: columns 16-113 (1001,2-1601,87 nm))

The parameters for each PCA are detailed in the Results section.

2.4. Measurements performed on powder blends

After finishing the blending, for the caffeine trials both calibration and validation, three samples of around 10 g each were taken from different positions as showed in *Figure 3*. For API A trials five samples were collected as showed in *Figure 4*. This samples were then used to carry measurements like penetration depth of the NIR light, Transmission Raman, HPLC. Also, a sample of around 5 g was collected for particle size analysis for each batch. The remaining blend was placed in a polybag and an in-process control was performed, including loss on drying and bulk density.

Figure 3: Samples taken after blending in caffeine trials.

Figure 4: Samples taken after blending in API A trials.

2.4. 1. Bulk density

The apparatus consists of a 100 mL cylindrical vessel of stainless steel. Initially the cup was placed on the scale and tared.

Using a small paddle, the sample was transferred until the vessel was full a little above the limit. The excess powder was scraped carefully from the top of the cup by smoothly moving the edge of the blade of a spatula perpendicular to and in contact with the top surface of the cup, taking care to keep the spatula perpendicular to prevent packing or removal of powder from the cup. Any material from the side of the cup was removed and the mass was determined, giving bulk density in g/100 mL. The bulk density (g/mL) was then recorded as the average of three determinations using three different powder samples to minimize the error.

2.4.2. Particle size distribution

A particle size analysis was performed in each blend using multifunctional particle size and morphology analyzer (see *Figure 5*), Camsizer® X2 from Microtrac. The steps carried out will be briefly explained [28] and for more details consult [29]. A small mass of powder (around two spoons) is poured via a sample feeder into the stainless-steel chute. The X-Jet mode of the Camsizer® X2 was used to disperse the powder. With the equipment vibration the powder slowly passes along the stainless-steel canal leading the particles to fall out of the chute into a receiving chute. Finally, images captured by basic and zoom cameras during the free fall of the particles are processed by an image analysis software that produces basic data about the material, including particle size and morphological parameters. The 4.0 mm nozzle was used with the 4.0 mm sample feeder with a pressure of 25 kPa. A comb and a pin are used to restrict the amount of powder passing and being analyzed in the chamber.

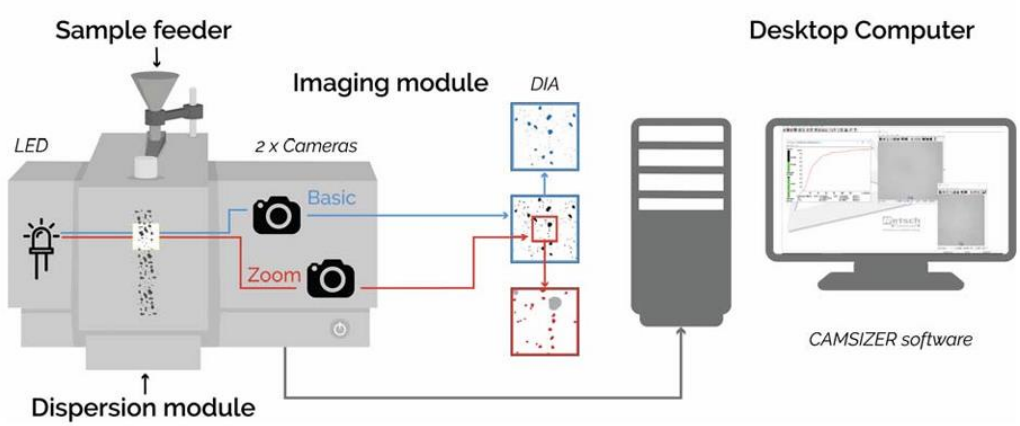


Figure 5: Camsizer X2 set[30].

2.4.3. Penetration depth of NIR light

Penetration depth of NIR light was only experimentally carried for three of the caffeine calibration experiments, namely the blends containing 10%, 15% and 20% of drug load. Since the concentrations of the mixtures are close between them, it was considered that the results for these three would be representative of the others. Penetration depth was not performed in validation trials as it has already been established for a drug load of 15% caffeine in first trials.

Penetration depth measurement was not carried for API A experiments. However, for data evaluation purposes the penetration depth determined for a previous trial with also 27.05% drug

load was used for further calculations. By obtaining the penetration depth, in combination with the bulk density, the number of revolutions per unit dose can be calculated and moving F-test models can be developed.

To carry this measurement the following steps were executed: the device with thinner layer depth (0.2 mm) was recorded empty to use as background reference; the samples were filled in the devices with layer depths of 0.2/0.4/0.6/0.8/1.0 mm, see *Figures 6* and *7*; to simulate the measurement on the container blender, the NIR probe was mounted on the container lid with sapphire window and placed on the sample device; five spectra per blend and sample device were recorded.

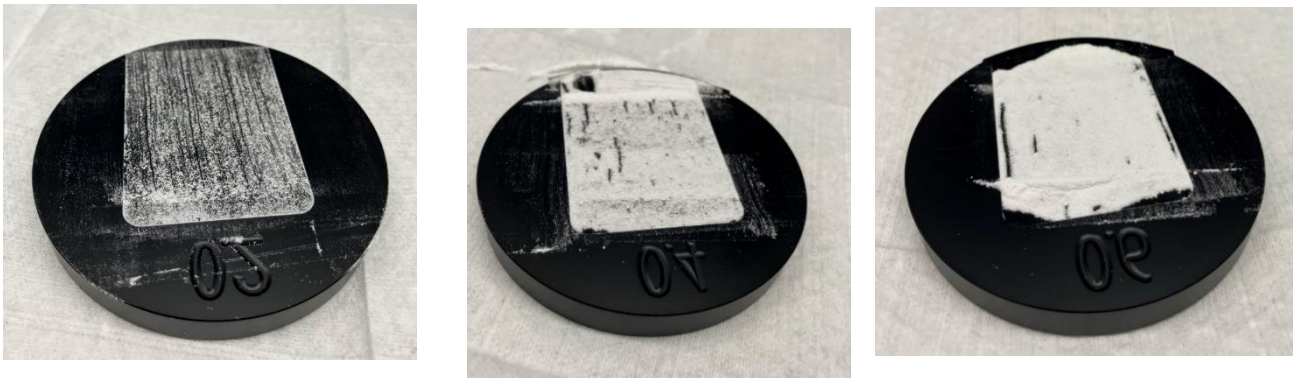


Figure 6: Penetration depth devices filled with powder with depths of 0.2/0.4/0.6 mm (from left to right).

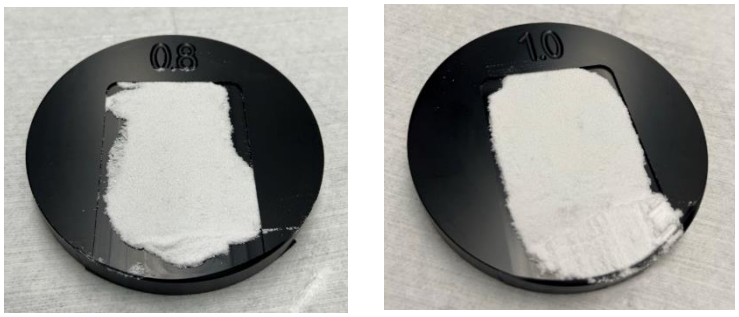


Figure 7: Penetration depth devices filled with powder with depths of 0.8/1.0 mm (from left to right).

The spectra were then evaluated on Unscrambler X 10.5 and Excel as follows:

- Raw spectra from empty sample devices and filled sample devices imported into Unscrambler X 10.5.
- The empty sample device with 0.2 mm was selected as background reference for the subsequent PCA. As the differences in the spectra's intensities shall be evaluated, the sample device with 0.2 mm has lowest intensity and no preprocessing of the spectra was done.
- For all blends individual PCAs were established in the wavelength range 1001.2 to 1601.87 nm. (Algorithm used: NIPALS, Validation method: Leverage, Number of Calibration samples used: 28, Total number of components: 7, Components suggested by model: 1, Optimal number of components: 1)
- The scores values of PC-1 and PC-2 were used for further evaluation.
- The scores values of PC-1 were imported and plotted in Excel.
- Min-Max-normalization of the values for PC-1 (min = 0.2 mm sample device (empty); max = 1.0 mm sample device (filled)).
- Normalized values of PC-1 were plotted versus the layer thickness and a specific function was fitted to the data.
- The function "goal seek" in Excel was used to determine the penetration depth with a specific value of 1 - 1/e. For details see figures in *Appendix.*
- The mass per revolution was calculated according to: $m = \frac{\rho \pi s^2 d_p}{4}$

ρ= bulk density, s^2 = sampling area of 10 mm, d_p = penetration depth.

2.4.4. HPLC

For the caffeine trials both calibration and validation, to determine the amount of caffeine in the samples taken from each batch, high pressure liquid chromatography (HPLC) was performed as reference analytics. The amounts weighed of each sample were between 112.5 mg and 225 mg. The exact amounts weighted for each sample are in Appendix. The balance used for weighing was a Mettler Toledo. The solutions were prepared in 200 ml volumetric flasks, after weighing the flask

is filled with Mili-Q-water until half. Then sonicated for 5-10 minutes and filled with water until full. Then the volumetric flasks are shaken for three times. After shaking, each sample was filled in a 5 mL tube and centrifuged for 5 min with 15000 rpm. After centrifugating, the samples were filled up into vials and analyzed. The analysis was performed using an isocratic system. The eluent used was 4:1 water/acetonitrile plus 11,04 g NaH_2PO_4. The blank solution is Mili-Q-water. The standard solution was prepared the same way as the samples with a concentration of 0.225 mg/mL.

2.4.5. UPLC

For the API A trials both calibration and validation, to determine the amount of API in the samples taken from each batch, ultra pressure liquid chromatography (UPLC) was performed intended as reference analytics. The amounts weighed of each sample were approximately of 2 mg for an intended final concentration of 200 µg/mL. The exact amounts weighted for each sample as well as the volume of extraction medium used are in Appendix. The balance used for weighing was microbalance Sartorius. The solutions were prepared in snap cap vials, the solvent used was 50% water/50% acetonitrile. The volumes used were between 1.894 mL and 3.517 mL. The samples were stirred at 300 rpm for 1 hour with a magnetic stirrer and after that, each sample was filled in a 1.5 mL tube and centrifuged for 5 min with 15000 rpm. After centrifugating, 700 µL was filled up into vials and analyzed. The standard solution was prepared the same way as the samples by dissolving 2 mg of drug substance in 10 mL of extraction medium. Also, a control standard was prepared by dissolving 2.2-2.3 mg of drug substance in 10 mL of extraction medium.

2.4.6. Transmission Raman

Transmission Raman measurements were carried out in Agilent TRS100 equipment in the samples as well as raw materials. For each raw material the settings (laser power, time and accumulations) were individually adjusted. Two different measurement methods were created, one for caffeine formulation and one for API A. In both methods adequate settings were defined for the lower and higher concentration batches and then used for all the samples for comparison purposes.

The spectra were pre-processed with Whittaker baseline correction and averaged to have just one measurement per sample. The samples taken from the container were measured as described below for both formulations:

- Samples were filled into 3D printed sample holders (see Figure 8).
- The samples were measured on a TRS100 sample tray at four positions. The tray was measured 3 times, giving 12 measurements per sample.
- The spectra were imported into Unscrambler X 10.5.

A PCA was performed and the detailed parameters are in the Results section.

Figure 8: Sample holder used for Transmission Raman measurements and four positions measured.

2.4.7. Paired samples T-Test

To assess differences between two data sets, namely predictions from regression models (NIR/TRS) and reference values (HPLC/UPLC), a paired samples T-test was performed. A significance level of α=0.05 was adopted for all statistical tests. The analysis was performed using the Excel's Analysis Tool Pak.

3. Results and Discussion

This section presents the results and their interpretation within the context of the study. It is organized in two parts. The first part discusses the calibration and validation trials using caffeine as a model substance. The second part discusses the trials using API A as drug substance.

3.1. Caffeine Experiments

After the blending time, in all experiments was observed that small amounts of material were sticking to the sapphire window, this may be due the electrostatic properties of caffeine.

3.1.1. Particle size distribution

3.1.1.1. Calibration Experiments

Regarding the particle size distribution measurement that was carried in Camsizer X2 for each blend consult Table 8. From *Figure 9* it´s possible to see that particle size distribution curves are almost perfectly overlaying, which is expected since all blends are composed of the same raw materials and the raw materials came from the same batch. The narrow range in which we find Mv^3 values, 109.7 and 115.4, confirms what was stated before that particle size distribution is relatively consistent across the different blends.

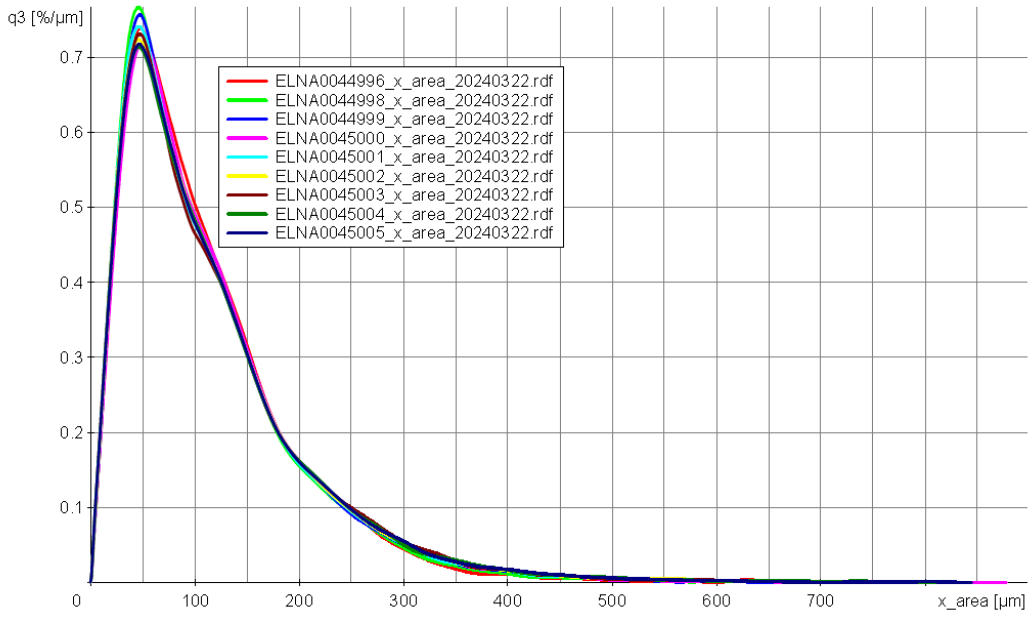


Figure 9: Overview of particle size distribution of experiments A1-A9.

Table 8: Particle Size Distribution Parameters: D_{10}, D_{50}, D_{90} Percentile Diameters and Mean Volume Diameter Mv^3 for experiments A1-A9.

Experiment	***D_{10} [µm]***	***D_{50} [µm]***	***D_{90} [µm]***	***Mv^3 (x) [µm]***
A1	31.5	89.1	217.3	109.7
A2	30.7	87.6	219.5	109.2
A3	31.3	89.2	225.7	112.3
A4	31.6	91.1	229.8	114.1
A5	30.5	89.1	228.6	113.1
A6	30.4	89.5	230.2	113.9
A7	30.3	89.7	233.3	114.3
A8	30.1	90.7	233.3	115.4
A9	30.7	90.4	232.9	114.9

3.1.1.2. Validation Experiments

Similar to the calibration trials, from *Figure 10* it´s possible to see that particle size distribution curves are almost perfectly overlaying, which is expected since all blends are composed of the same raw materials and the raw materials come from the same batch. Here the Mv^3 (Table 9) value varies even in smaller range than calibration trials, between 116.1 and 118.8, which is once again in accordance with particle size distribution consistency.

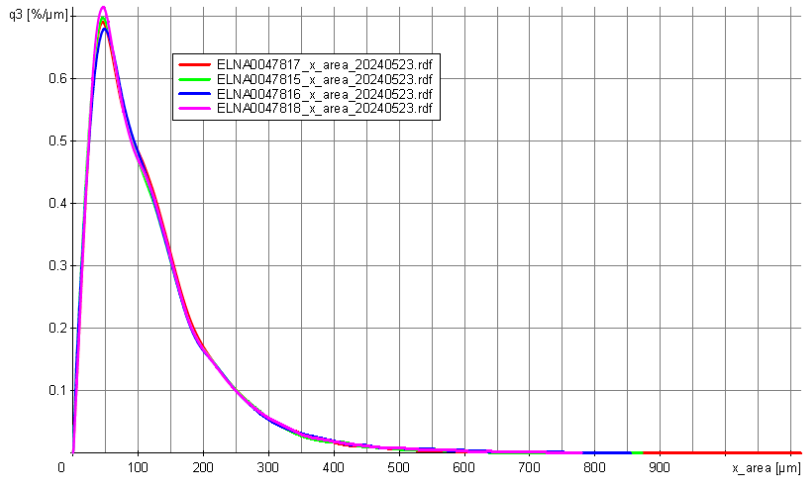


Figure 10: Overview of particle size distribution of experiments AV1-AV4.

Table 9: Particle Size Distribution Parameters: D_{10}, D_{50}, D_{90} Percentile Diameters and Mean Volume Diameter Mv^3 for experiments AV1-AV4.

Experiment	***D_{10} [µm]***	***D_{50} [µm]***	***D_{90} [µm]***	***Mv^3 (x) [µm]***
AV1	29.2	91.3	236.8	116.1
AV2	29.7	93.1	240.4	118.8
AV3	31.2	94.3	236.1	117.3
AV4	31.0	93.1	239.9	118.3

3.1.2. NIR spectrometry

3.1.2.1. Calibration Experiments

One objective of this study was to determine whether the different caffeine concentrations could be detected, so the spectra collected during the blending process were evaluated by PCA in this wavelength range 1001,2-1601,87 nm. The following parameters were used, Algorithm used: NIPALS, Validation method: Leverage, Number of Calibration samples used: 3198, Total number of components: 7, Components suggested by model: 4, Optimal number of components: 4.

As seen in *Figure 11*, the data clustering is shown in a scores plot (principal component 1 vs principal component 2). Here the batches are color coded, and each spectrum is represented as

a dot. The clustering of the different caffeine concentrations is significantly more observed along PC-1. Although a clustering along PC-2 is also observed, this principal component contributes only with 12%. In the beginning the measurements of each batch are scattered, but as time progresses it´s visible a clustering of the dots which means that the blending is becoming homogeneous.

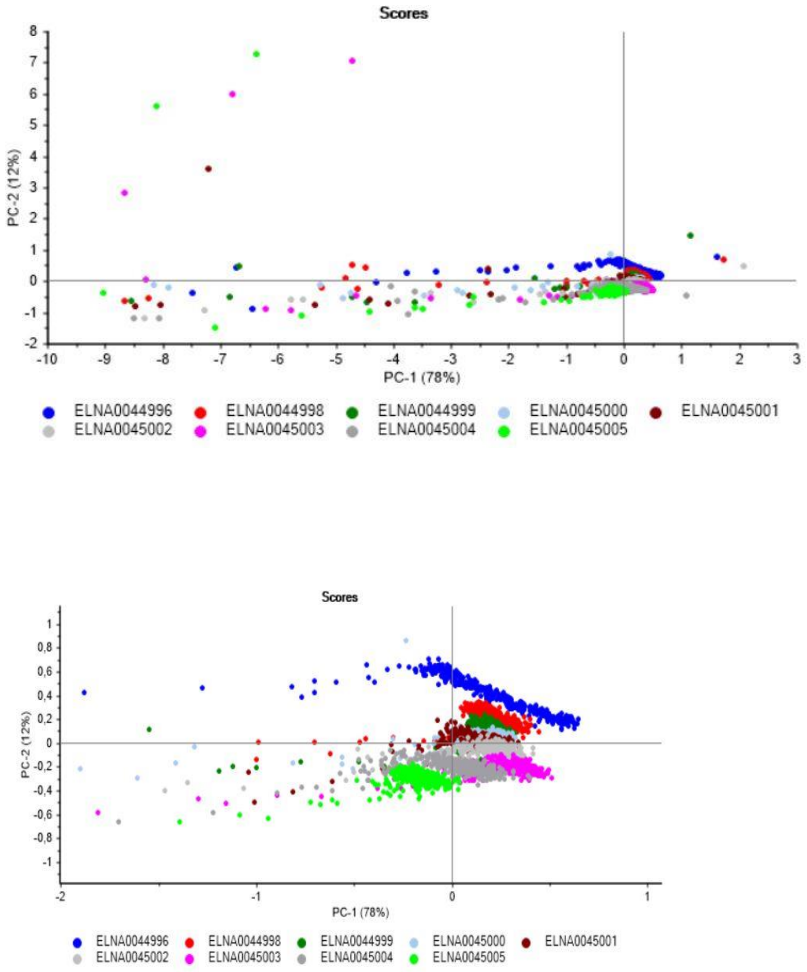


Figure 11: (a) Scores plot for experiments A1-A9 based on PCA. The x-axis and y-axis represent the first and second components, capturing 78% and 12% of the total variance, respectively. (b) Scores plot for experiments A1-A9 zoom in.

Even though the data are more distributed along the first component, both the loadings plot of PC-1 and PC-2 were further investigated. Each raw material of the formulation was measured in glass vials, see Figure 12. Three spectra per sample were collected and then averaged. The

averaged spectra were pretreated in the same way as for the PCA. The spectra of all raw materials are shown in the thesis' Appendix.

Figure 12: Example of off-line NIR measurement of raw material.

In *Figure 13*, it´s shown the overlaying of loadings of PC-1 with the preprocessed spectrum of caffeine. It is possible to observe that the PC-1 peaks are inversely correlated to the caffeine peaks, which means that in the batches with higher caffeine content, the concentration of other excipients like mannitol decreases and that higher caffeine concentrations can be found at lower/negative values on PC1.

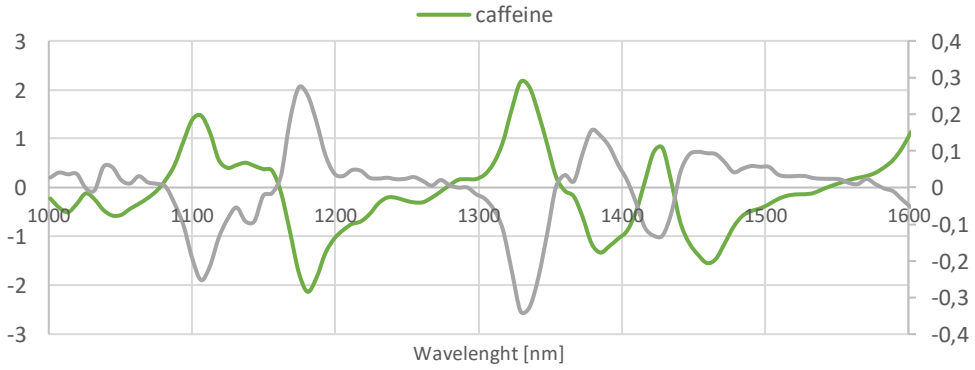


Figure 13: NIR loadings of PC-1 and pre-processed spectrum of caffeine for calibration experiments.

In *Figure 14* the loadings of PC-2 and the preprocessed spectra of caffeine and mannitol, the only raw materials which the concentration is changing between batches are shown. In the wavelength below 1400 nm caffeine appears to be inversely correlated with PC-2, whereas above that wavelength seems to have no correlation. Regarding mannitol above 1400 nm seems inversely correlated with this principal component but the peak between 1100 and 1200 nm matches the shape of PC-2 loadings.

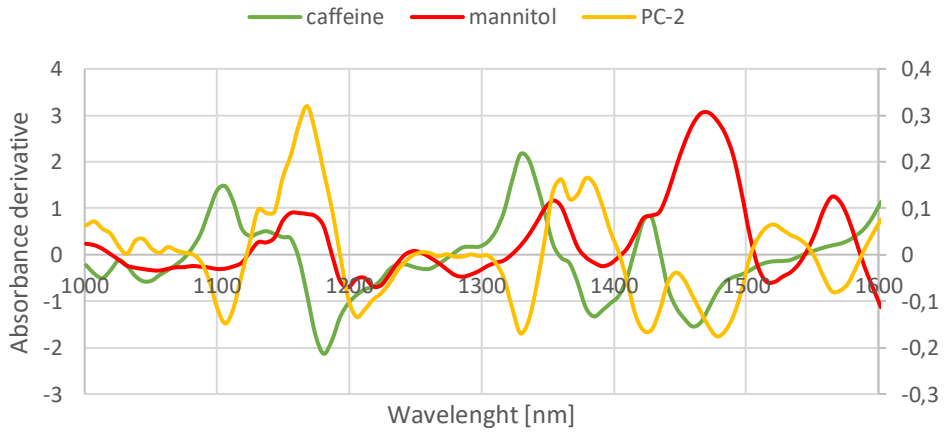


Figure 14: NIR loadings of PC-2 with pre-processed spectra of raw materials (caffeine and mannitol) for calibration experiments.

In *Figure 15*, it´s possible to see the scattering of the different caffeine concentrations of the blends at 1100 nm and 1329 nm, where caffeine has characteristic peaks. Also at 1471 nm, a characteristic peak of mannitol, we see a separation of spectra the ones with higher mannitol content and therefore lower caffeine concentration are above in the peak.

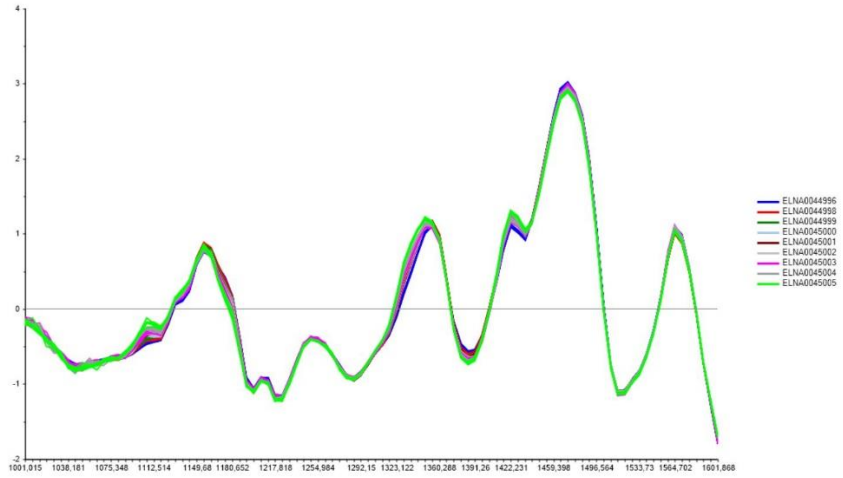


Figure 15: Line plot of preprocessed NIR spectra of experiments A1-A9 in wavelength between 1001,02 nm and 1601,87 nm.

The PCA only gives qualitative information about the blending process so to investigate further, a partial least squares (PLS) regression model was developed. This allows to monitor quantitatively the caffeine concentration in the blends. To develop this model the samples taken from each batch were measured with NIR (see *Figure 16*) and the caffeine concentration obtained from HPLC analysis was used as reference.

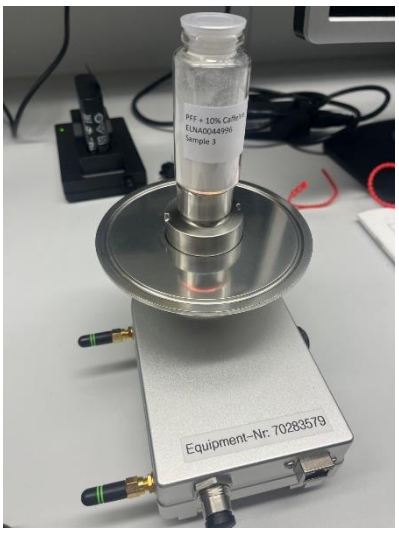


Figure 16: Example of off-line sample measurement with NIR spectrometer.

It was necessary to determine which data pretreatment was more suitable for the data to develop a better model. In Table 10, the different options that were tested are presented highlighting the RMSECV and R-square for each one. The chosen pretreatment is the one with the highest R-square and lowest prediction error, since both 1st derivative followed by SNV and 2nd derivative gave very similar results, predictions were made with both these models, but the first derivative followed by SNV gave better results. One factor was found to be optimal for the resulting model.

Table 10: Summary of Partial Least Squares (PLS) model performance with different pre-processing treatments applied to the data.

PLS pretreatments	***R-square***	***RMSECV [%]***	***Optimal number of factors***
No pretreatment (just absorbance)	0.93	0.79	5
1st derivative	0.91	0.84	2
SNV	0.94	0.79	4
1st derivative and SNV	0.96	0.63	1
2nd derivative	0.96	0.64	4
2nd derivative and SNV	0.94	0.72	3

When overlaying the weighted regression coefficients with the raw materials (*Figure 17*), the peaks are is the same position as the caffeine peaks, suggesting the prediction of the PLS model is strongly correlated with the caffeine concentration in the samples.

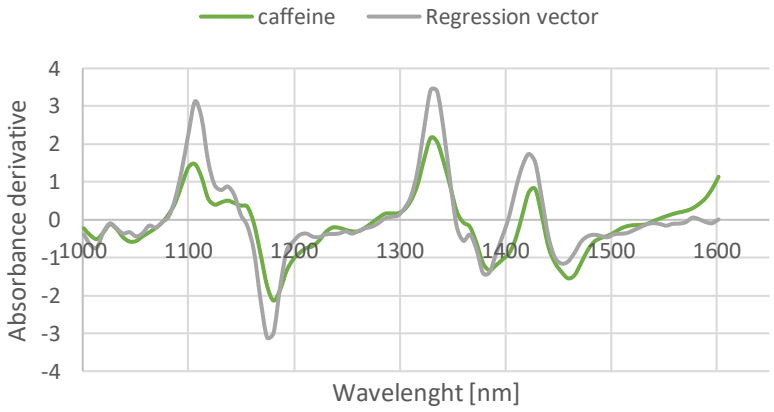


Figure 17: NIR PLS weighted regression coefficients and caffeine pre-processed spectrum for calibration experiments.

When looking at the predicted vs reference graphic (*Figure 18*), the results are valid for all batches because they are quite similar to the reference values. The model can be considered to have overall a high predictive accuracy and consistency, the points are mostly clustered along the line.

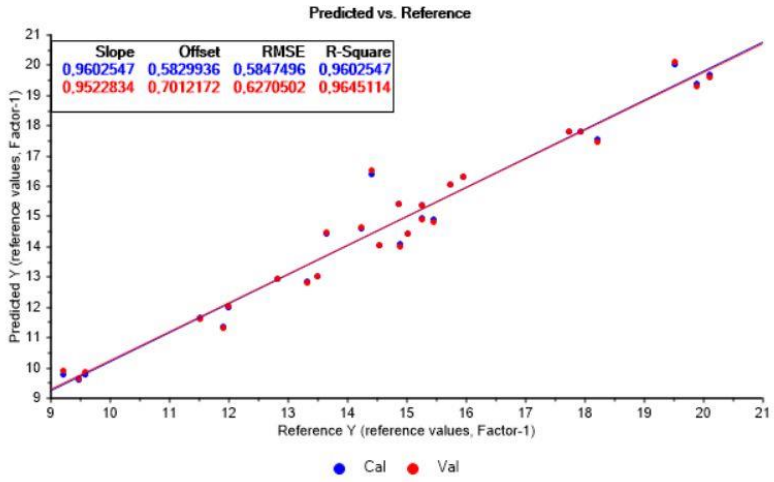


Figure 18: True vs predicted plot for caffeine calibration PLS model. The plot shows the relationship between actual (x-axis) and predicted (y-axis) values from the PLS model. The diagonal line represents perfect predictions (y = x). Points near the line indicate better model accuracy.

For each batch, predictions were made using the PLS model developed. Although these predictions are based on the same dataset used for model calibration, they will be further evaluated on an independent set of data. Therefore, we're essentially training how well the model fits the data it was trained on.

In *Figure 19*, the predictions for all batches are represented and for each batch individually are in Appendix. In the beginning there´s a higher variation but as blending progresses the predicted value becomes more constant.

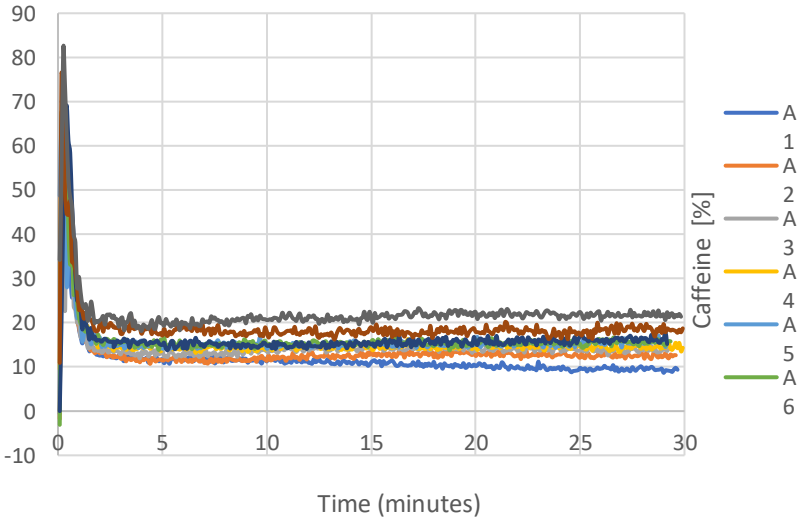


Figure 19: NIR PLS predictions for caffeine calibration experiments over time.

When comparing the results obtained by the predictions and HPLC analysis as well as the respectively standard deviations, the HPLC values are overall lower. Considering the HPLC is more accurate than predictions from PLS models, one can conclude that maybe the model is overestimating the samples concentration. In general, the suitability of PLS for the prediction of API concentrations during blending has been proven.

Table 11: Comparison between PLS predictions and HPLC results for experiments A1-A9.

Experiment	***Nominal values [%]***	***Prediction (mean last 10 spectra) [%]***	***Std prediction (last 10 spectra)***	***HPLC (average per batch) [%]***	***HPLC std***
A1	10	9.24	0.32	9.43	0.15
A2	12	12.52	0.32	11.81	0.20
A3	13.5	13.76	034	13.22	0.28
A4	14.5	14.59	0.65	14.36	0.52
A5	15	15.75	0.62	14.91	0.50
A6	15.5	15.61	0.45	15.13	0.18

A7	16.5	16.25	0.81	15.37	0.68
A8	18	18.29	0.31	17.96	0.20
A9	20	21.60	0.50	19.84	0.24

Additionally, a paired samples t-test was conducted, and the results are presented in the Table below. The mean of the predictions suggests that the model is slightly overestimating the actual reference values as described before. The p-value obtained as it is less than the significance level of 0.05 indicates that difference between predictions and actual values is statistically significant.

Table 12: Results of the paired t-test comparing NIR predictions of calibration data set with HPLC values (caffeine). The table presents the mean, variance, degrees of freedom (df), *t*-value and *p*-value. A *p*-value < 0.05 indicates statistical significance.

	Mean	*Variance*	*df*	*t-value*	*p-value (two-tail)*
Prediction	15.29	12.14	8	3.44	0.009
Reference	14.67	9.52			

3.1.2.2. Validation Experiments

The acquired spectra were used to compare results between the validation trials and to predict concentrations over time with the PLS model from the calibration trials.

The spectra collected during the blending time were evaluated in the same manner as in first trials, including the same data pretreatment.

Analyzing the line plot of the preprocessed spectra of all batches, looking specifically for the caffeine peak at 1106 nm, we can see a separation of the batches that are color coded. For AV3 that´s supposed to be inhomogeneous, a larger separation of the spectra would be expected but as we can see in *Figure 20* that doesn´t verify.

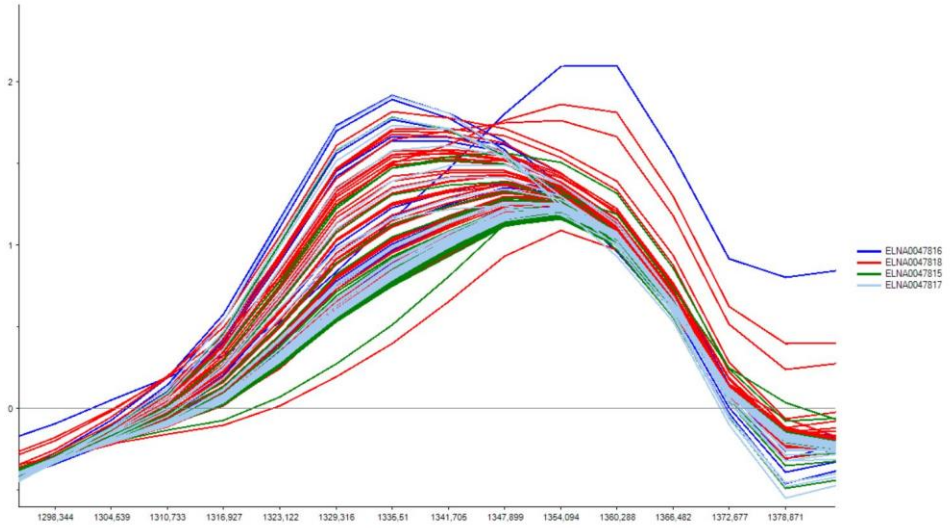


Figure 20: Line plot of preprocessed NIR spectra of validation trials showcasing caffeine peak.

Once again, the PCA (wavelength range 1001,2-1601,87 nm, Algorithm used: NIPALS, Validation method: Leverage, Number of Calibration samples used: 767, Total number of components: 7, Components suggested by model: 4, Optimal number of components: 4) shows the clustering of the data, indicating that as time passes the blend is becoming homogeneous. PC-1 was further investigated, and its loadings were compared to the preprocessed caffeine and mannitol spectra, see *Figures 22* and 23, respectively. Caffeine is inversely correlated with PC-1, so for positive values of caffeine there´s negative values of PC-1 and vice-versa. Above 1400 nm, mannitol is also correlated with PC-1 the spectrum matches the shape of the PC-1 loadings. The correlation is not as good as in first trials though, some of the peaks are not exactly in the same position.

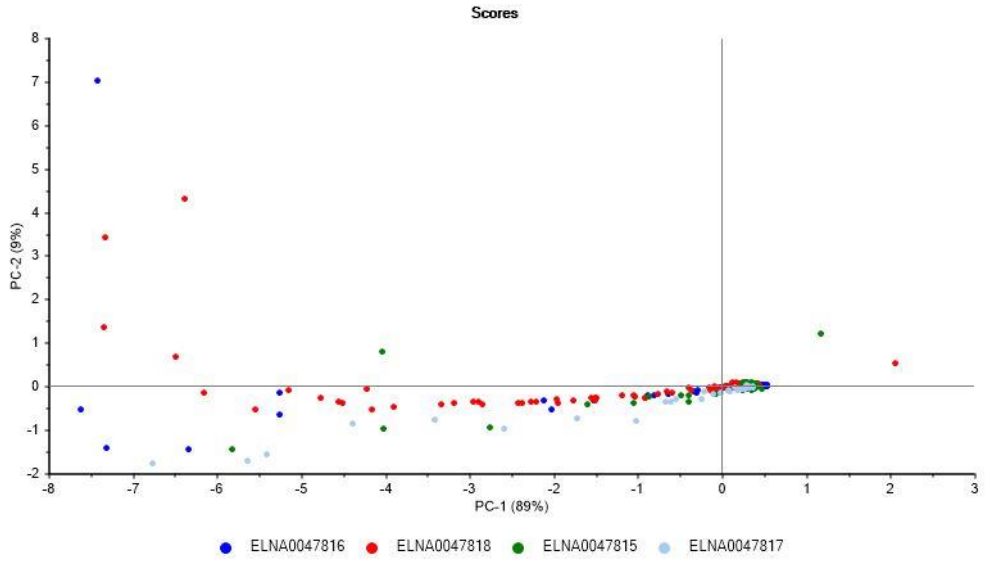


Figure 21: Scores plot for experiments AV1-AV4 based on PCA. The x-axis and y-axis represent the first and second components, capturing 89% and 9% of the total variance, respectively.

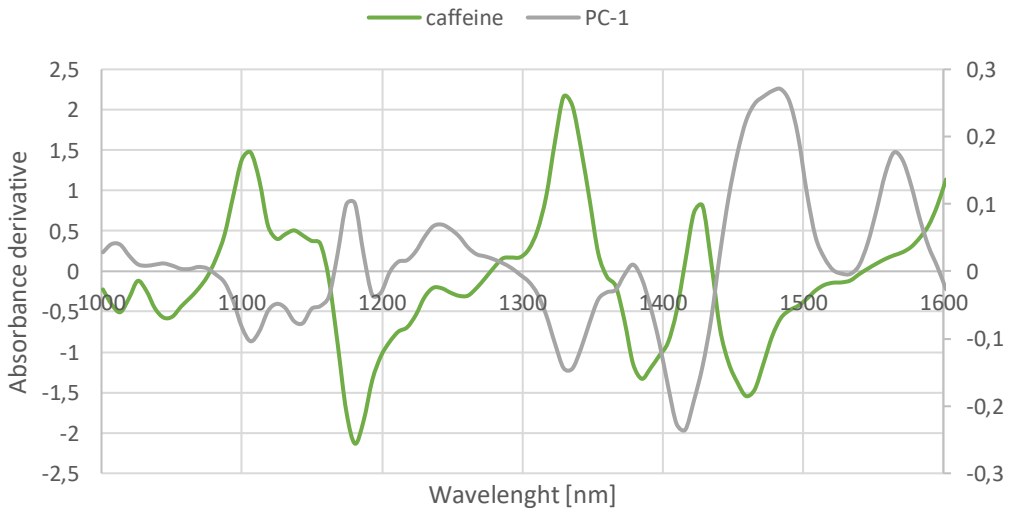


Figure 22: NIR loadings of PC-1 and preprocessed caffeine spectrum for validation experiments.

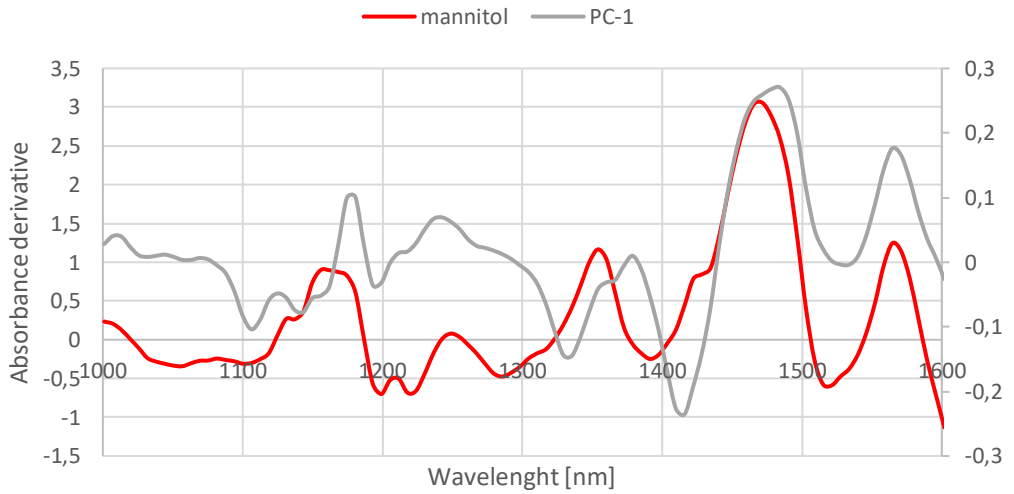


Figure 23: NIR loadings of PC-1 and pre-processed mannitol spectrum for validation experiments.

Using the raw data collected during the blending of the validation experiments and the PLS model previously developed, predictions of the caffeine concentration over time were made. In Table 13, predictions as well as HPLC results are shown and the respective standard deviations.

Looking at the prediction values, the blend AV3 shows the best result and closer to the expected value of 15% drug load. This doesn´t make much sense since the blending time of this experiment was the shortest and supposedly not enough to achieve a homogeneous blend. However, in batch AV1 the blending was stopped after less than 6 minutes, given that we waited for a minute after the f-value was below the threshold value. Meaning that around 5 minutes was already homogeneous, so after 3 minutes perhaps is not as inhomogeneous as we initially thought. Another unexpected prediction result is for the experiment AV2, with a blending time of 30 minutes the mixture is supposedly homogeneous and in previous trials the prediction gave a better result, nonetheless fits the expectation. As for experiment AV4, although the prediction shows a similar result to the other batches, the concentration determined by HPLC shows a much lower value, which may indicate that this was in fact not homogeneous. A larger batch size may need a longer blending time or increasing rotation speed. Also powder sampling from a container is always challenging so that should be taken into consideration.

Table 13: Comparison between PLS predictions and HPLC results for AV1-AV4 experiments.

Batch	***Nominal value [%]***	***Prediction (last 10 spectra) [%]***	***Prediction Std***	***HPLC (averaged) [%]***	***HPLC Std***
AV1	15	13.92	0.29	14.68	0.22
AV2	15	14.14	0.37	15.24	0.82
AV3	15	14.79	0.37	14.57	0.82
AV4	15	14.03	0.48	12.98	0.82

Additionally, a paired samples t-test was conducted, and the results are presented in the Table below. The p-value obtained as it is greater than the significance level of 0.05 indicates that difference between predictions and actual values is not statistically significant.

Table 14: Results of the paired t-test comparing NIR predictions of validation data set with HPLC values (caffeine). The table presents the mean, variance, degrees of freedom (df), *t*-value and *p*-value. A *p*-value < 0.05 indicates statistical significance.

	Mean	*Variance*	*df*	*t-value*	*p-value (two-tail)*
Prediction	14.22	0.15	3	-0.30	0.78
Reference	14.36	0.94			

3.1.3. Penetration Depth of NIR light

3.1.3.1. Calibration Experiments

The measurements carried in sample devices with layer depths of 0.2 and 0.4 mm were very difficult to perform. Due to being a very thin layer it was hard to get a smooth surface without dragging the material out and not have empty spaces.

The penetration depth and bulk density allowed to determine the mass per revolution for each blend, as described in the Methods section.

Table 15: Overview results of bulk density, penetration depth and determined mass per revolution for experiments A1, A5, and A9.

Experiment	***Nominal value [%]***	***Bulk density [g/ml]***	***Penetration depth [mm]***	***Mass per revolution [mg]***
A1	10	0.53	0.21	8.64
A5	15	0.55	0.24	10.56
A9	20	0.55	0.31	13.40

The penetration depth is increasing with increasing bulk density, and therefore with increasing caffeine concentration. The penetration depth varies between 0.21 and 0.31 mm. For an unit dose of 200 mg, the number of revolutions decreases with increasing bulk density and increasing caffeine concentration, varying between 8.64 mg and 13.40 mg.

Based on these results, for the evaluation of blend uniformity by F-test in validation trials considering a unit dose of 200 mg and a drug load of 15% was decided that a blocksize of 20 would be adequate to cover the 1-to-3-fold unit dose.

Table 16: Calculation of revolutions for unit dose of 200 mg for experiments A1, A5 and A9.

Experiment	***Number of revolutions for 200 mg***
A1	23
A5	19
A9	15

3.1.4. Transmission Raman

3.1.4.1. Calibration Experiments

The spectra of the raw materials were plotted, see *Figure 24*. MCC shows a characteristic fluorescence signal. Caffeine shows a strong and characteristic Raman signal at 555 cm^{-1}. A distinction between caffeine concentrations can be seen in this peak, see *Figure 25*.

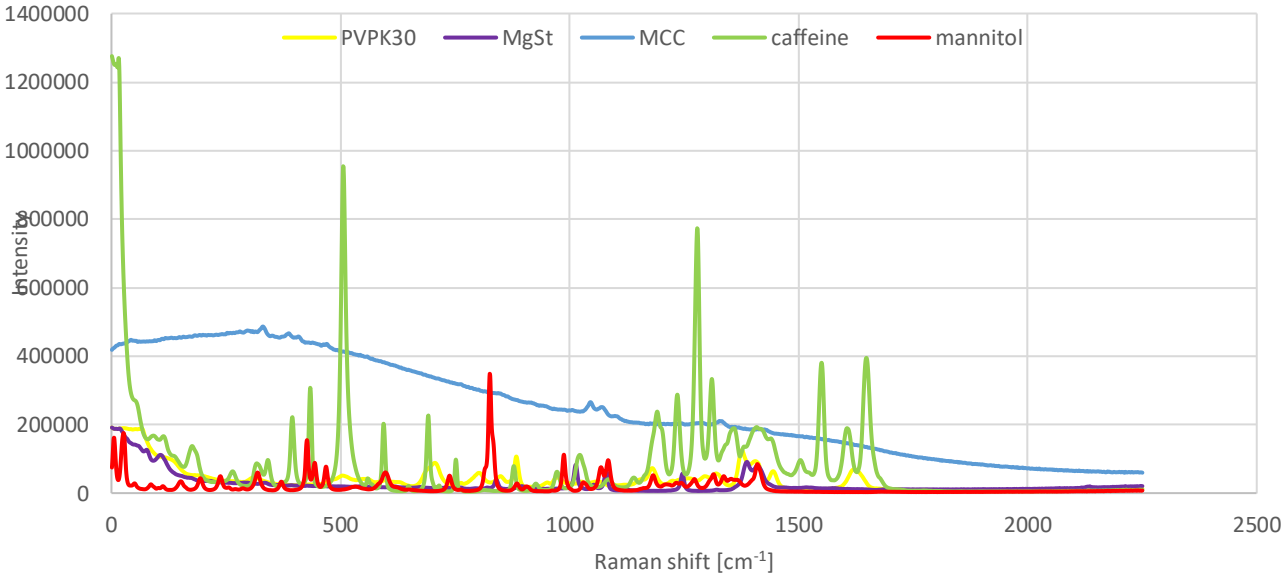


Figure 24: Transmission Raman spectra of raw materials for caffeine formulation.

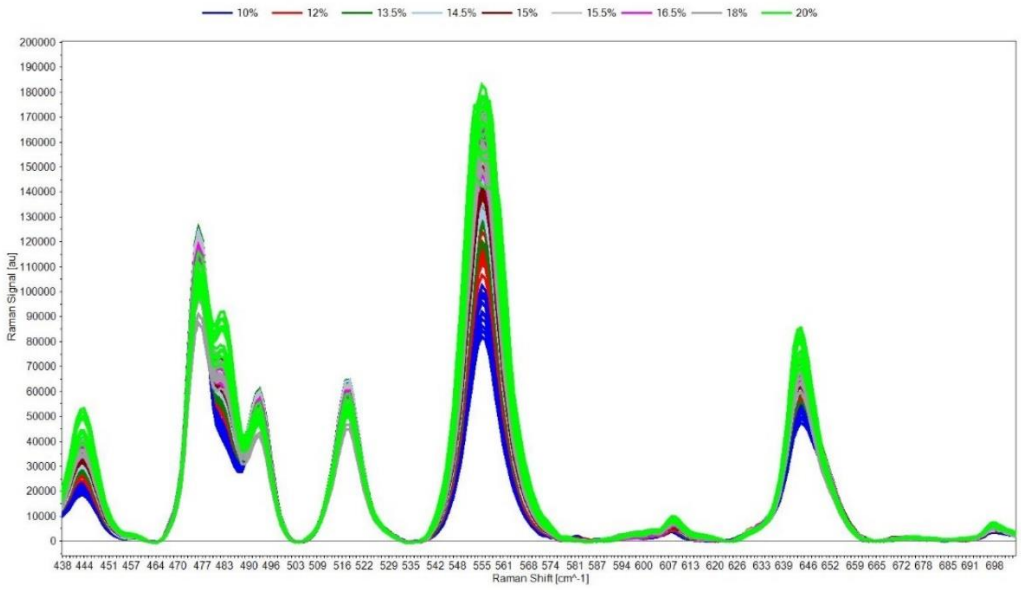


Figure 25: Transmission Raman spectra preprocessed with Whittaker for experiments A1-A9.

The objective of the PCA (wavelength range 219-1901 cm^{-1}, Algorithm used: NIPALS, Validation method: Leverage, Number of Calibration samples used: 28, Total number of components: 7, Components suggested by model: 3, Optimal number of components: 3) was to demonstrate that the differences in the caffeine concentrations can be detected with transmission

Raman. The clustering of the spectra is shown in the scores plot, see *Figure 26* (principal component 1 vs principal component 2). The clustering of the different caffeine concentrations is observed along PC-1, the batches are color coded. For further interpretation of the correlation of the caffeine concentration and the clustering of the batches the loadings plot of both PC-1 and PC-2 were investigated.

In *Figures 27* and *28*, the loadings of PC-1 and PC-2 are as well as the preprocessed spectrum of caffeine and mannitol respectively are plotted. The loadings of PC-1 are clearly correlated with caffeine as both show peaks in the same position, while PC-2 appears to be more correlated with mannitol.

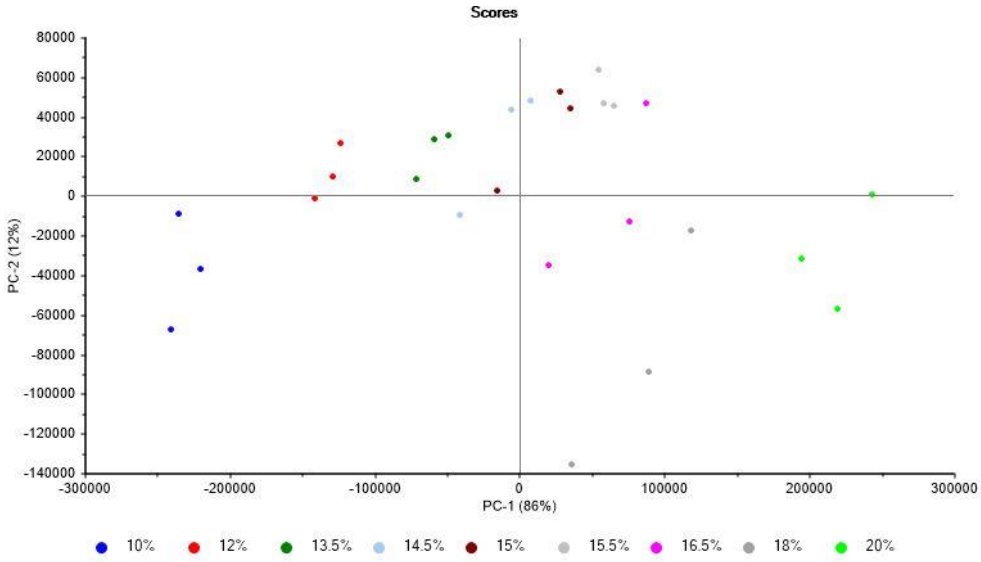


Figure 26: Scores plot with the Transmission Raman preprocessed spectra of the samples collected after blending for experiments A1-A9 based on PCA. The x-axis and y-axis represent the first and second components, capturing 86% and 12% of the total variance, respectively.

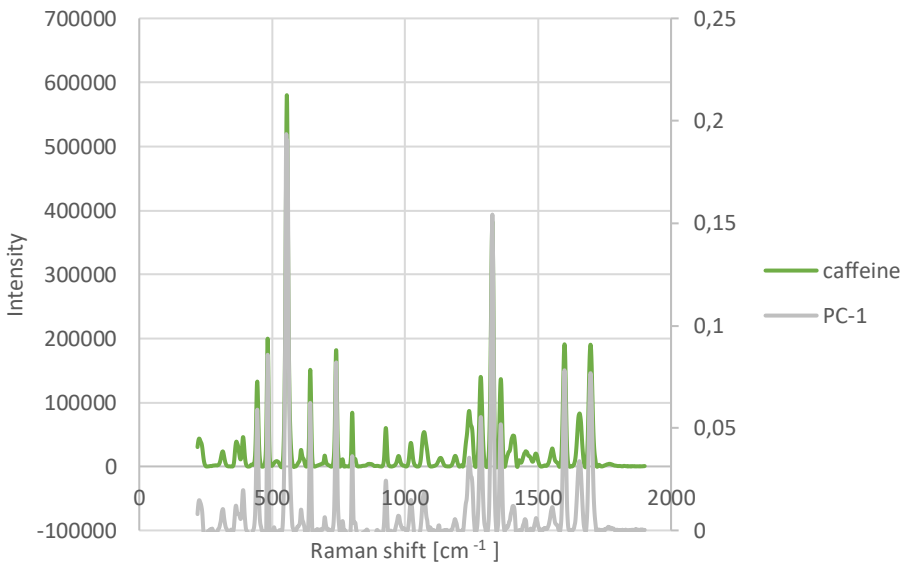


Figure 27: Transmission Raman loadings of PC-1 and the preprocessed spectrum of caffeine for calibration experiments.

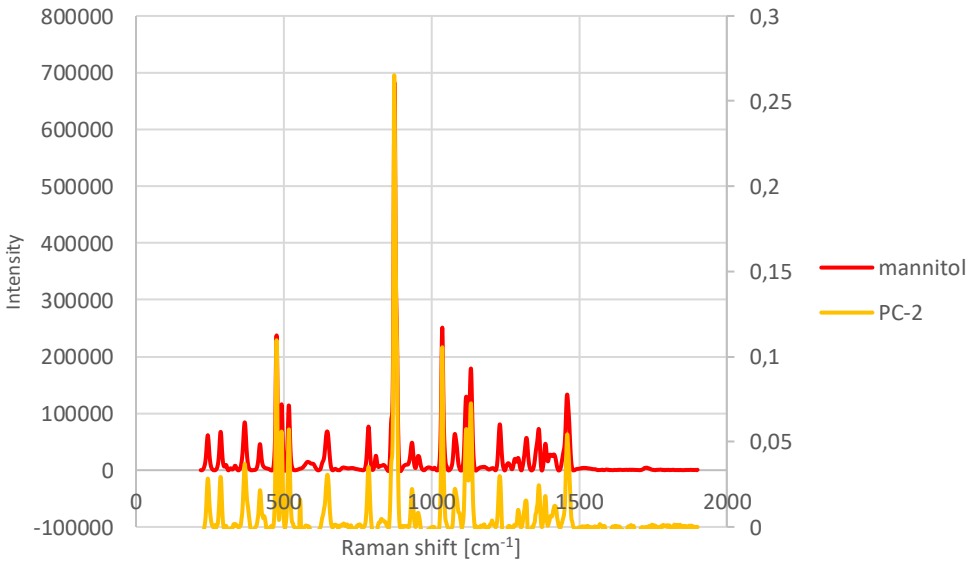


Figure 28: Transmission Raman loadings of PC-2 and preprocessed spectrum of mannitol for calibration experiments.

A quantitative evaluation by PLS regression was performed using TRS measurements of the samples and HPLC results as reference values. The following parameters were used:

- Algorithm used: NIPALS, Validation method: Cross validation, Cross validation method: Random with 20 segments, Number of Calibration samples used: 27, Total number of factors: 7, Factors suggested by model: 2, Optimal number of factors: 2

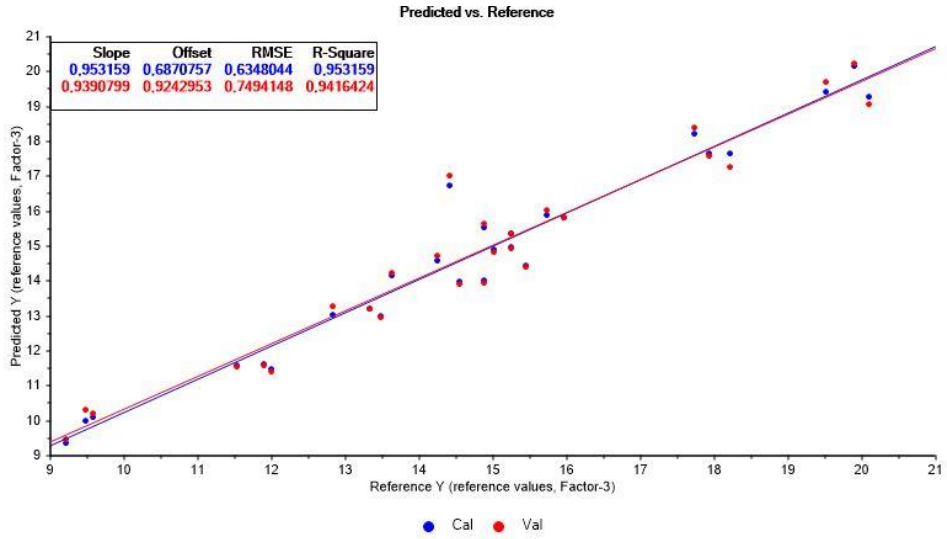


Figure 29: Predicted vs Reference plot of Raman PLS model caffeine calibration trials. The plot shows the relationship between reference (x-axis) and predicted (y-axis) values from the PLS model. The diagonal line (y = x) represents the ideal case where predictions perfectly match the actual values. Points near the line indicate better model accuracy.

The PLS model showed a R-square of 0.94 and a prediction error of 0.75%. When overlaying the PLS weighted regression coefficients with the caffeine pre-processed spectrum, the correlation is strong as the peaks match in position, as shown in Figure 30.

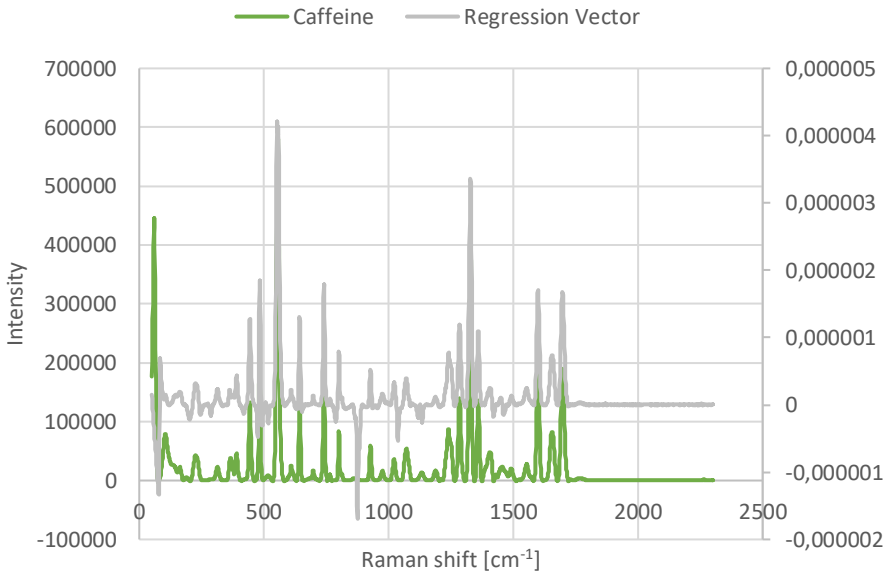


Figure 30: Transmission Raman PLS weighted coefficients and caffeine pre-processed spectrum for calibration experiments.

Predictions were made using the PLS model developed, for each batch there were three predictions, and these were averaged to have one result per batch and compared with HPLC, these results are summarized in Table 17.

The prediction results are very similar to the results obtained by HPLC analysis, which is a good indicator that Transmission Raman might be used as an alternative reference analytics technique to determine caffeine concentration.

Table 17: Comparison between Transmission Raman PLS predictions and HPLC results for experiments A1-A9.

Experiment	***Nominal values [%]***	***Prediction (average three samples) [%]***	***Prediction Std [%]***	***HPLC (average per batch) [%]***	***HPLC Std [%]***
A1	10	9.36	0.02	9.43	0.15
A2	12	12.11	0.16	11.81	0.20
A3	13.5	13.16	0.25	13.22	0.28
A4	14.5	14.19	0.61	14.36	0.52

A5	15	14.82	0.28	14.91	0.50
A6	15.5	14.99	0.18	15.13	0.18
A7	16.5	15.60	0.57	15.37	0.68
A8	18	17.88	0.15	17.96	0.20
A9	20	19.90	0.37	19.84	0.24

A paired samples t-test was conducted, and the results are presented in the Table below. The p-value obtained as it is greater than the significance level of 0.05 indicates that difference between predictions and actual values is not statistically significant. These results show that regarding the calibrations trials, Transmission Raman is closer to HPLC method than NIR for the whole range of concentrations. Also indicates that NIR model might not generalize well across different concentrations.

Table 18: Results of the paired t-test comparing Transmission Raman predictions of calibration data set with HPLC values (caffeine). The table presents the mean, variance, degrees of freedom (df), *t*-value and *p*-value. A *p*-value < 0.05 indicates statistical significance.

	Mean	*Variance*	*df*	*t-value*	*p-value (two-tail)*
Prediction	14.67	9.49	8	-0.04	0.97
Reference	14.67	9.52			

3.1.4.2. Validation Experiments

The samples taken after blending were measured in Transmission Raman using the same method developed for the first trials.

In the validation experiments all batches have the same caffeine concentration, therefore PCA does not give much information. The loadings of PC-1 show a mixture of caffeine and mannitol, as the loadings shape matches both raw materials spectra.

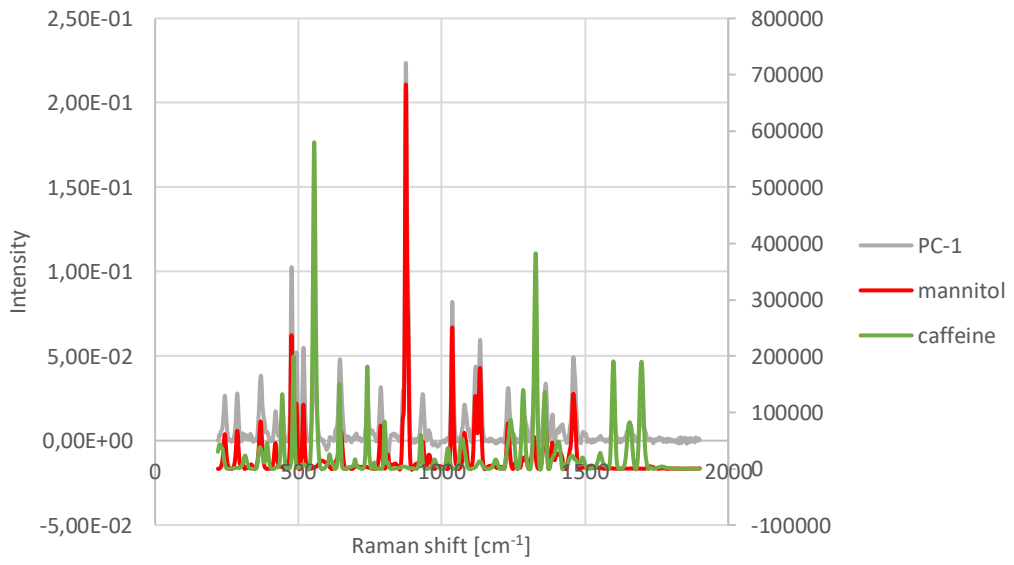


Figure 31: Transmission Raman loadings of PC-1 with pre-processed spectra of caffeine and mannitol for calibration experiments.

Predictions were made using the Raman PLS model previously developed, for each batch there was three predictions, and these were averaged to have one result per batch and compare with HPLC, these results are summarized in Table 19.

Regarding the Raman predictions, the mean result for each batch is fitting the expectation and is not significantly different from the HPLC result obtained, AV4 being the exception. The PLS model was created with all batches with the same size and may not be robust enough when faced with variations in batch size.

Table 19: Transmission Raman PLS predictions and HPLC results for experiments AV1-AV4.

Experiment	***Nominal value [%]***	***Raman Prediction (average per batch) [%]***	***Std [%]***	***NIR Prediction (last 10 spectra) [%]***	***Std [%]***	***HPLC (average per batch) [%]***	***Std [%]***
AV1	15	14.57	0.50	13.92	0.29	14.68	0.22
AV2	15	14.54	0.51	14.14	0.37	15.24	0.82
AV3	15	15.46	0.61	14.79	0.37	14.57	0.82

AV4	15	15.49	0.03	14.03	0.48	12.98	0.82

A paired samples t-test was conducted, and the results are presented in the Table below. The p-value obtained as it is greater than the significance level of 0.05 indicates that difference between predictions and actual values is not statistically significant. These results show that regarding the validation trials, both NIR and Transmission Raman perform similarly as both showed no significant differences from HPLC. However, this is only true for a specific concentration (15%) within the calibration range, so overall Transmission Raman showed to be more consistent and reliable.

Table 20: Results of the paired t-test comparing Transmission Raman predictions of validation data set with HPLC values (caffeine). The table presents the mean, variance, degrees of freedom (df), *t*-value and *p*-value. A *p*-value < 0.05 indicates statistical significance.

	Mean	*Variance*	*df*	*t-value*	*p-value (two-tail)*
Prediction	15.02	0.28	3	0.92	0.42
Reference	14.37	0.94			

3.2. API A Experiments

Following the results presented in the caffeine trials, the next step was to apply these principles to the API of interest that was API A. So, following the same line of thought as in previous trials, first a design of experiments considering the range 18.94% up to 35.17%.

3.2.1. Particle size distribution

3.2.1.1. Calibrations Experiments

A particle size distribution measurement was carried in Camsizer X2 for each blend. From *Figure 32* it´s possible to see that in particle size distribution curves only small differences are observed between the five batches, caused by the small changes in composition. This is expected since all blends are composed of the same raw materials and the raw materials are from the same batch.

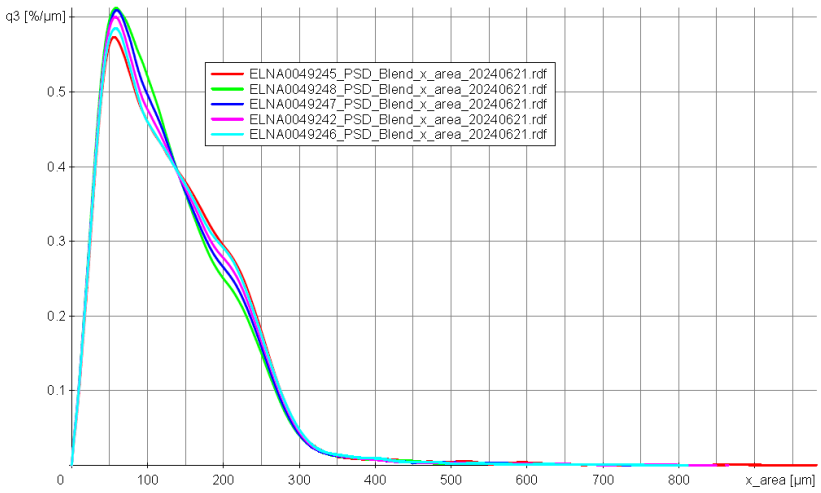


Figure 32: Overview of particle size distribution of experiments B1-B5.

Table 21: Particle Size Distribution Parameters: D_{10}, D_{50}, D_{90} Percentile Diameters and Mean Volume Diameter Mv^3 for experiments B1-B5 and BV1.

Experiment	***D_{10} [µm]***	***D_{50} [µm]***	***D_{90} [µm]***	***Mv^3 (x) [µm]***
B1	40.1	113.3	235.9	129.0
B2	40.5	116.0	237.7	130.6
B3	41.0	117.9	239.0	132.6
B4	39.9	110.5	233.3	126.9
BV1	39.4	110.2	231.6	126.4

3.2.1.2. Validation Experiments

A particle size distribution measurement was carried in Camsizer X2 for experiment BV1.The results are presented in Table 21 and are similar to the results found for the calibration trials, which is expected because the raw materials used on both came from the same batch.

3.2.2. NIR spectrometry

3.2.2.1. Calibration Experiments

To investigate the blending process with different API loads, the spectra collected during the blending process, already pre-processed as it was previously described in Methods section, were

evaluated by principal PCA (wavelength range 1001,2-1601,87 nm, Algorithm used: NIPALS, Validation method: Leverage, Number of Calibration samples used: 887, Total number of components: 7, Components suggested by model: 4, Optimal number of components: 4).

As can be seen in *Figure 33* the clustering of the data is shown in a scores plot (principal component 1 vs principal component 2). Here the batches are color coded and each spectrum is represented as a dot. The clustering of the different API concentrations is significantly more observed along PC-1, from low concentration (left) to high concentration (right). In the beginning the measurements of each batch are scattered, but as time progresses it´s visible a clustering of the dots which means that the blending is becoming homogeneous.

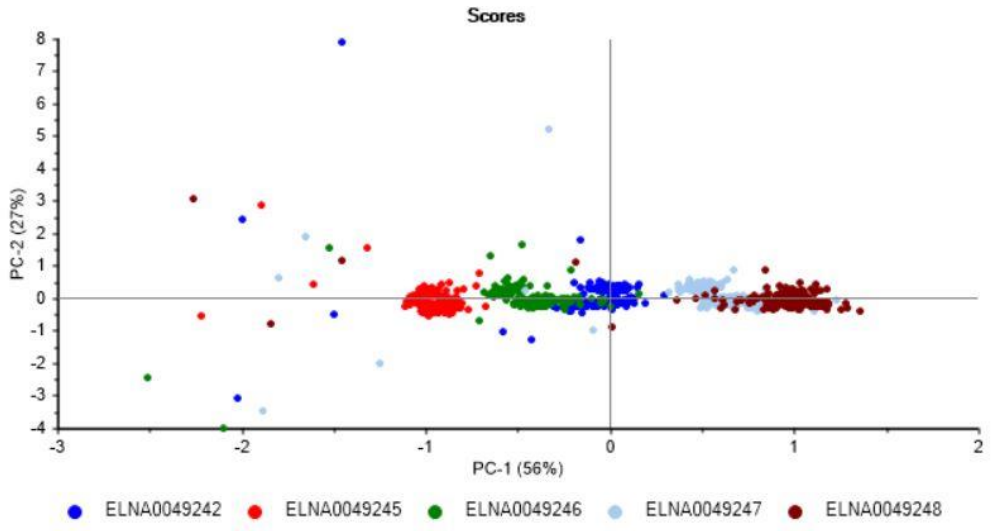


Figure 33: Scores plot for experiments B1-B5 based on PCA. The x-axis and y-axis represent the first and second components, capturing 56% and 27% of the total variance, respectively.

Both the loadings plot of PC-1 and PC-2 were further investigated. Each raw material of the formulation was measured in glass vials, as shown before. Three spectra per sample were collected and then averaged. The averaged spectra were pretreated in the same way as for the PCA.

In *Figure 34*, it´s shown the overlaying of loadings of PC-1 with the preprocessed spectrum of the API. It is possible to observe that the PC-1 peaks are inversely correlated to the API peaks, which means that in the batches with higher API content, the concentration of other excipients like MCC and mannitol decreases. API A presents a characteristic NIR peak at 1558 nm.

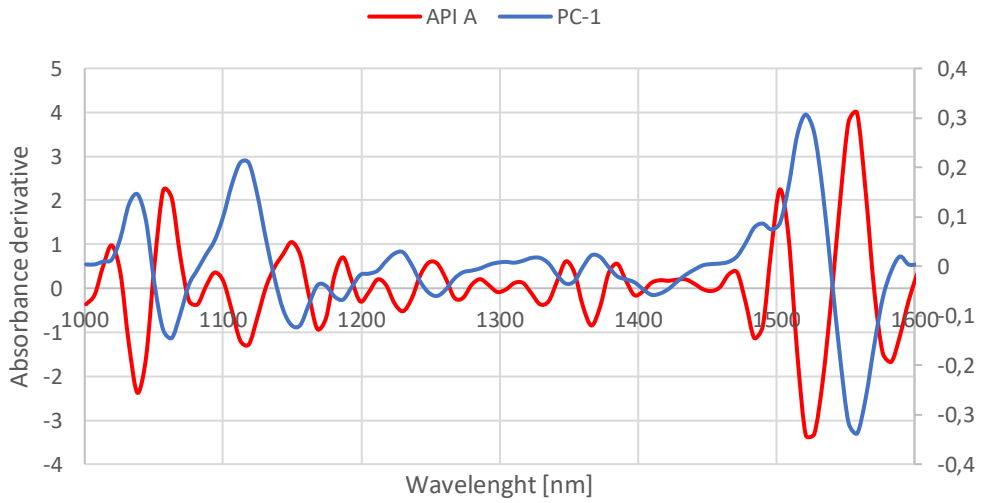


Figure 34: NIR loadings of PC-1 and pre-processed spectrum of API A for calibration experiments.

In *Figure 35*, the loadings of PC-2 and the pre-processed spectrum of mannitol are represented. The shape of the loadings slightly matches the shape of the spectrum so they´re correlated.

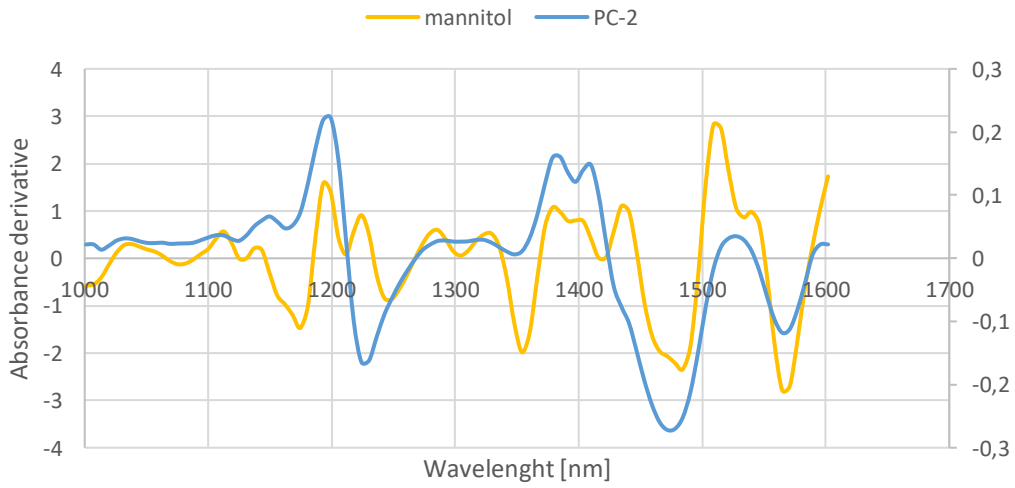


Figure 35: NIR loadings of PC-2 and pre-processed spectrum of mannitol for calibration experiments.

A partial least squares (PLS) regression model was developed. This allows to monitor quantitatively the API concentration in the blends. To develop this model the samples taken from each batch were measured with NIR and the API concentration obtained from UPLC analysis as reference values.

First, it was necessary to determine which data pretreatment was more suitable for the data to develop a better model. In Table 22, the different options that were tested are presented highlighting the RMSECV and R-square for each one. The chosen pretreatment is the one with the highest r-square and lowest prediction error, in this case the 2nd derivative. Two factors were found to be optimal for the resulting model.

Overall, all the models presented a low R-square and higher prediction error than would be preferable, which may indicate that the predictions made with this model could be less reliable. The discrepancies in API concentration obtained in UPLC between samples of the same experiment may be the reason for these results, see table in Appendix.

Table 22: Summary of Partial Least Squares (PLS) model performance with different pre-processing treatments applied to the data.

PLS pretreatments	***R-square***	***RMSECV [%]***	***Optimal number of factors***
No pretreatment (just absorbance)	0.86	2.32	4
1st derivative	0.89	2.08	2
SNV	0.82	2.33	1
1st derivative and SNV	0.88	2.27	1
2nd derivative	0.90	2.00	2
2nd derivative and SNV	0.88	2.24	1

In *Figure 36*, the weighted regression coefficients and API spectrum are overlayed. Although not perfectly, the curves match in shape meaning that the prediction is strongly correlated with API concentration. In *Figure 37*, it´s noticeable that above 1400 nm to 1500 nm the weighted regression coefficients match the shape of mannitol.

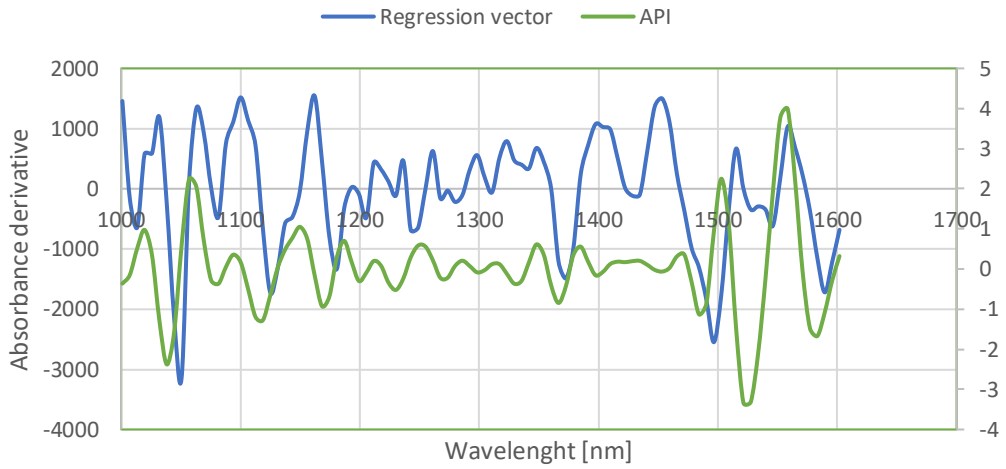


Figure 36: NIR PLS weighted regression coefficients and API A pre-processed spectrum for calibration experiments.

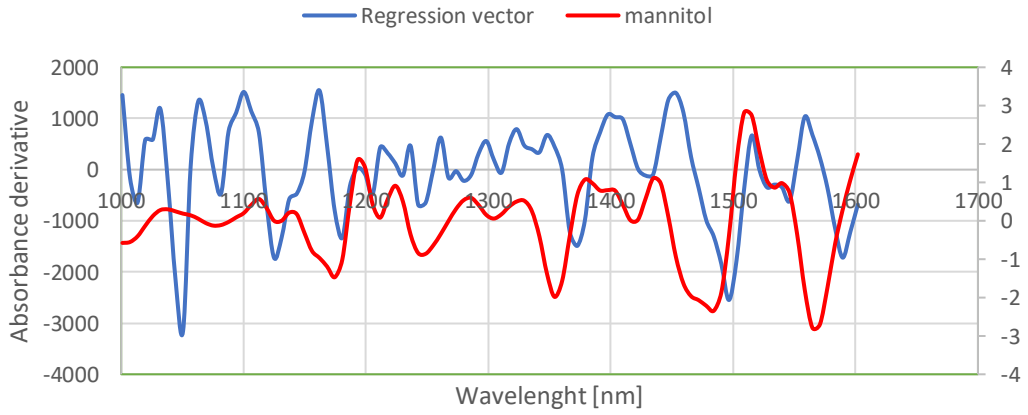


Figure 37: NIR PLS weighted regression coefficients and mannitol pre-processed spectrum for calibration experiments.

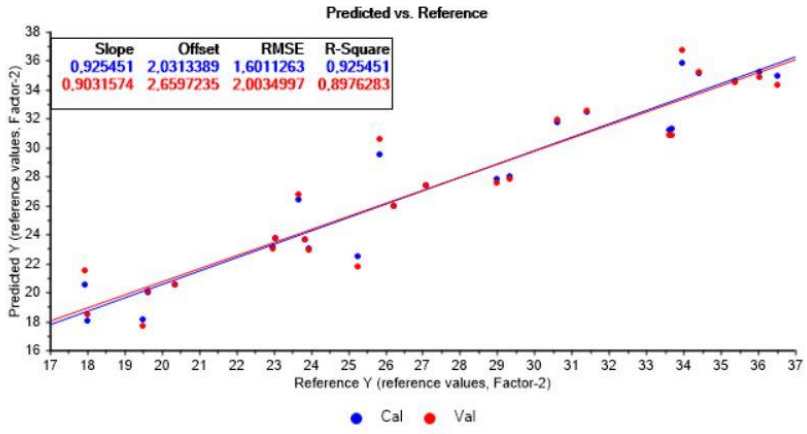


Figure 38: True vs predicted plot for NIR PLS model for API A calibration trials. The plot shows the relationship between reference (x-axis) and predicted (y-axis) values from the PLS model. The diagonal line (y = x) represents the ideal case where predictions perfectly match the actual values. Points near the line indicate better model accuracy.

For each batch, predictions were made using the PLS model developed. In *Figure 39*, the predictions for all batches over time are represented and for each batch individually are in Appendix. In the beginning there´s a higher variation but as blending progresses the predicted value becomes more constant. At 10 minutes the prediction goes to a lower value but then is again constant, this happens because it´s the pause between blendings 1 and 2.

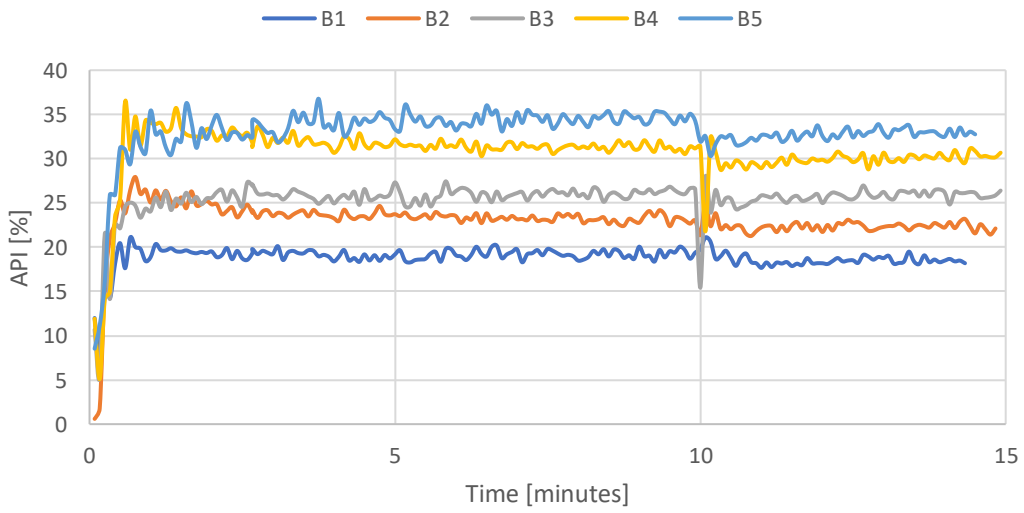


Figure 39: NIR PLS predictions for experiments B1-B5 over time.

When comparing the results obtained by the predictions and UPLC analysis as well as the respectively standard deviations, the prediction values are overall lower. Considering the UPLC is more accurate than predictions, one can conclude that the model is underestimating the samples concentration. In general, the suitability of PLS for the prediction of API concentrations during blending has been proven.

Table 23: Comparison between NIR PLS predictions and UPLC results for experiments B1-B5.

Experiment	***Nominal value [%]***	***Prediction (mean last 10 spectra) [%]***	***Std prediction (last 10 spectra) [%]***	***UPLC (average per batch) [%]***	***UPLC std [%]***
B1	18.94	18.46	0.27	19.08	0.95
B2	22.99	22.32	0.55	23.82	0.83
B3	27.05	25.99	0.29	27.05	2.06
B4	31.11	30.41	0.47	31.03	2.85
B5	35.17	32.91	0.36	35.26	0.96

A paired samples t-test was conducted, and the results are presented in the Table below. The p-value obtained as it is less than the significance level of 0.05 indicates that difference between predictions and actual values is statistically significant. This was previously observed with the caffeine formulation, suggesting that NIR model cannot generalize well across a large range of concentrations.

Table 24: Results of the paired t-test comparing NIR predictions of calibration data set with UPLC values (API A). The table presents the mean, variance, degrees of freedom (df), *t*-value and *p*-value. A *p*-value < 0.05 indicates statistical significance.

	Mean	*Variance*	*df*	*t-value*	*p-value (two-tail)*
Prediction	26.02	34.40	4	-3.80	0.02
Reference	27.25	39.25			

3.2.2.2. Validation Experiments

The spectra collected during the blending time were evaluated in the same manner as in first trials, including the same data pretreatment.

The validation batch was added to the calibration data set and a PCA with the following parameters was performed: (Algorithm used: NIPALS, Validation method: Leverage, Number of Calibration samples used: 1068, Total number of components: 6, Components suggested by model: 4, Optimal number of components: 4)

In *Figure 40*, the scores plot of the PCA is presented and it´s possible to see that the validation trial is close to experiment B3 as expected since both have a drug load of 27.05%.

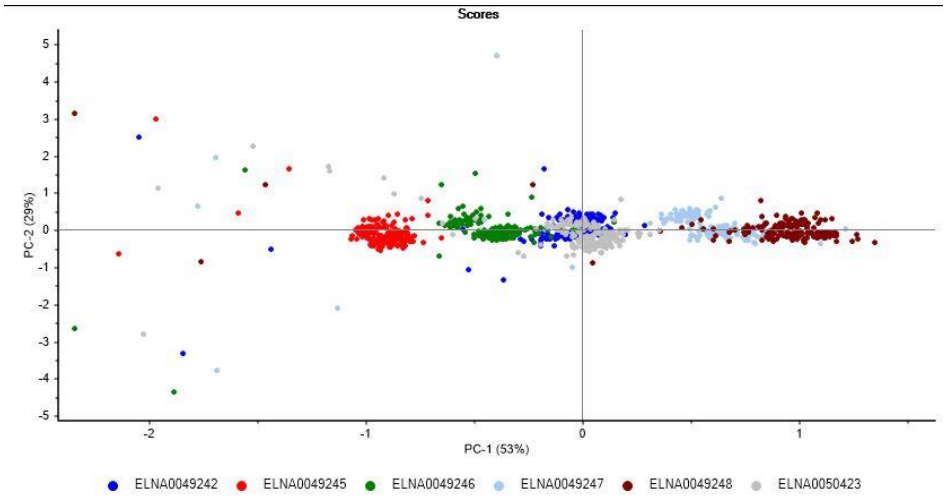


Figure 40: Scores plot for experiments B1-B5 plus validation experiment BV1 based on PCA. The x-axis and y-axis represent the first and second components, capturing 53% and 29% of the total variance, respectively.

Using the raw data collected during the blending and the PLS model previously developed, a prediction of the API concentration over time was made. In Table 25, the prediction as well as UPLC result are shown and the respective standard deviations.

Both the NIR prediction and UPLC obtained concentration are below the expected value but similar enough between them, which indicates that the model can accurately predict this set of data. The UPLC standard deviation is very high but as mentioned before powder sampling from the container is always challenging.

Table 25: Comparison between NIR PLS predictions and UPLC results for experiment BV1.

Experiment	*Nominal value [%]*	*Prediction (mean last 10 spectra) [%]*	*Prediction Std [%]*	*UPLC (average per batch) [%]*	*UPLC Std [%]*
BV1	27.05	26.13	0.46	25.99	3.86

3.2.3. Penetration depth

3.2.3.1. Calibration Experiments

Penetration depth measurement was not carried in these experiments. However, for data evaluation purposes the penetration depth determined for another experiment was used for the calculations. This experiment had also a drug load of 27.05% of API A.

Given that all experiments have approximately the same bulk density, the masses per revolution determined were between 9.19 mg and 9.39 mg. To reach a 200 mg unit dose around 22 revolutions are necessary.

Table 26: Overview results of bulk density, penetration depth and determined mass per revolution for experiments B1-B5.

Experiment	*Bulk density [g/ml]*	*Penetration depth [mm]*	*Mass per revolution [mg]*
B1	0.46	0.26	9.39
B2	0.45	0.26	9.19
B3	0.45	0.26	9.19
B4	0.45	0.26	9.19
B5	0.45	0.26	9.19

Table 27: Calculation of revolutions for unit dose of 200 mg for experiments B1-B5.

Experiment	***Number of revolutions for 200 mg***
B1	21
B2	22
B3	22
B4	22
B5	22

3.2.4. Transmission Raman

3.2.4.1. Calibration Experiments

The spectra of the raw materials were plotted, see *Figure 41* and *42*. API A shows a strong and characteristic Raman signal at 1576 cm^{-1}. A distinction between API concentrations can be seen in this peak, see *Figure 43*.

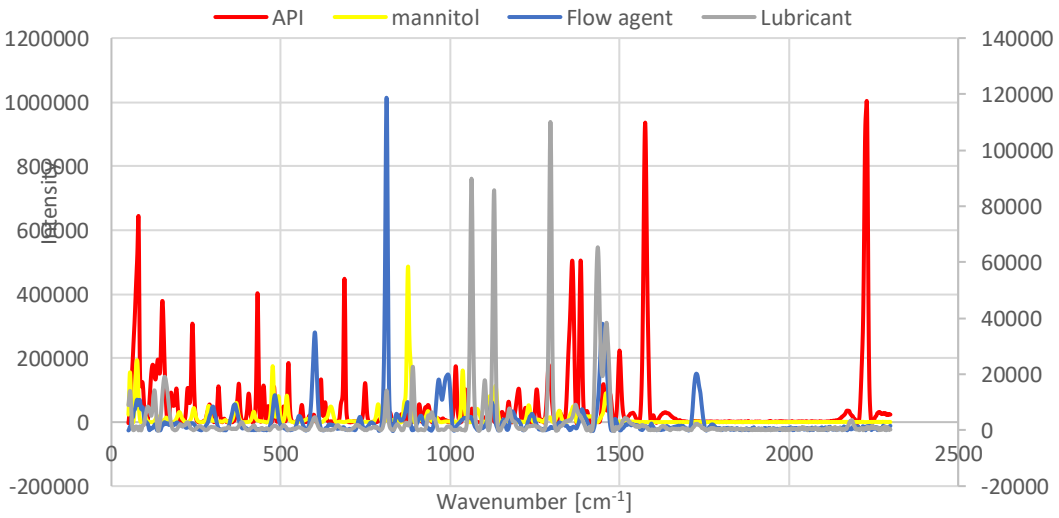


Figure 41: Transmission Raman spectra of raw materials (API A, mannitol, flow agent, lubricant).

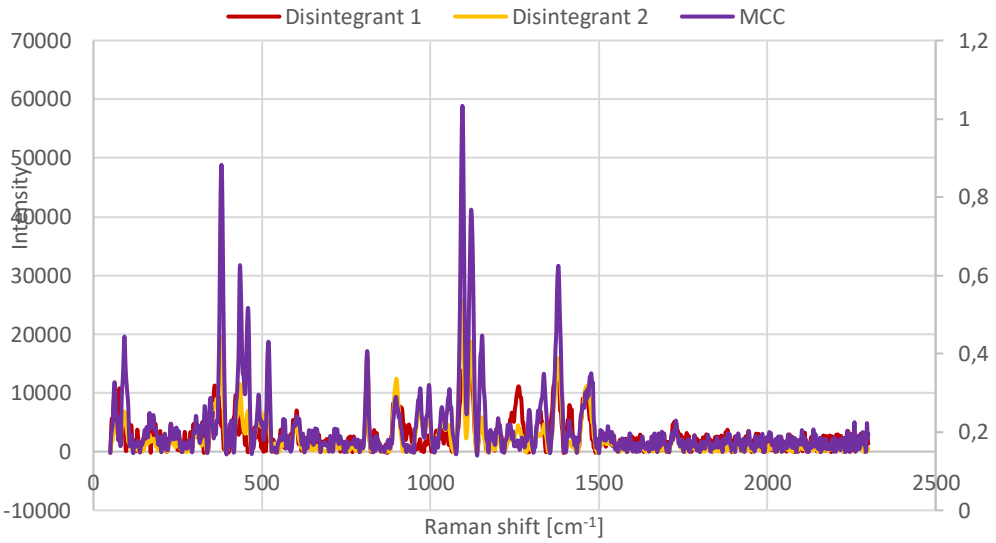
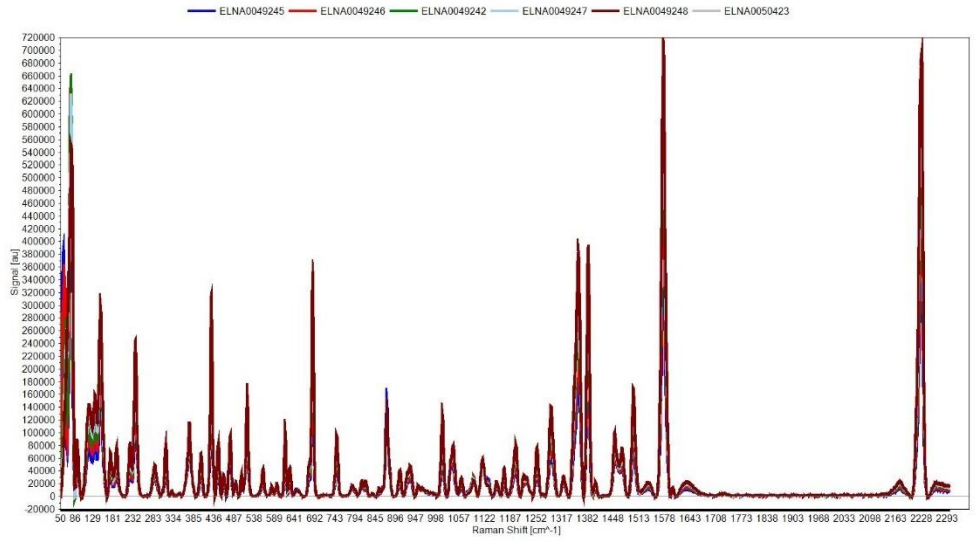


Figure 42: Transmission Raman spectra of raw materials (MCC, disintegrants 1 and 2).



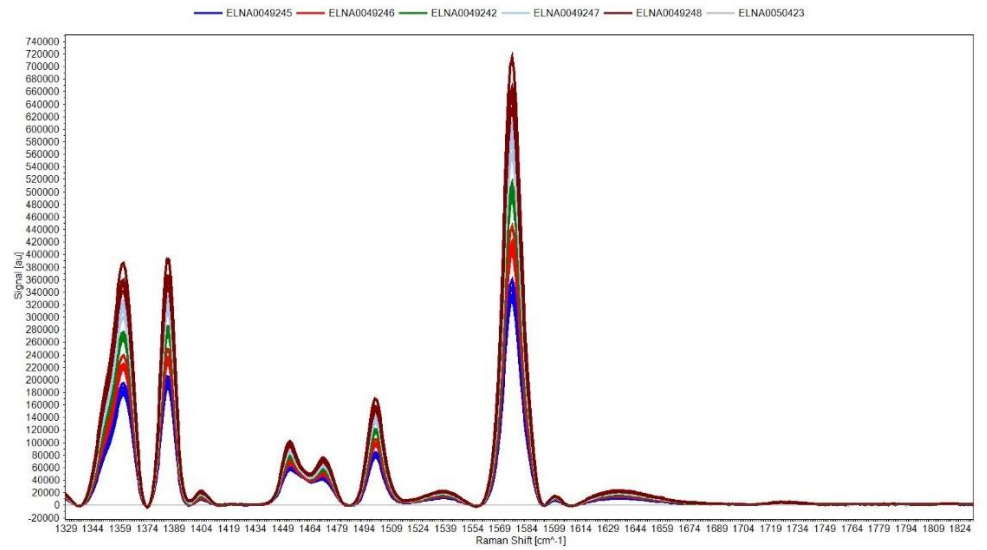


Figure 43: (a) Transmission Raman line plot of pre-processed spectra of experiments B1-B5 and (b) Transmission Raman line plot of pre-processed spectra of experiments B1-B5 zoom in API A peak.

The objective of the PCA was to demonstrate that the differences in the API concentrations can be detected with transmission Raman.

A PCA was performed with the preprocessed spectra with following parameters: (Algorithm used: NIPALS, Validation method: Leverage, Number of Calibration samples used: 25, Total number of components: 7, Components suggested by model: 3, Optimal number of components: 3)

The clustering of the spectra is shown in the scores plot, see *Figure 44* (principal component 1 vs principal component 2). The clustering of the different API concentrations is observed along PC-1, so the loadings of the component were further investigated.

In *Figure 45*, the loadings of PC-1 as well as the preprocessed spectrum of the API are plotted. The loadings of PC-1 are clearly correlated with the API as they match in shape and both show peaks in the same position.

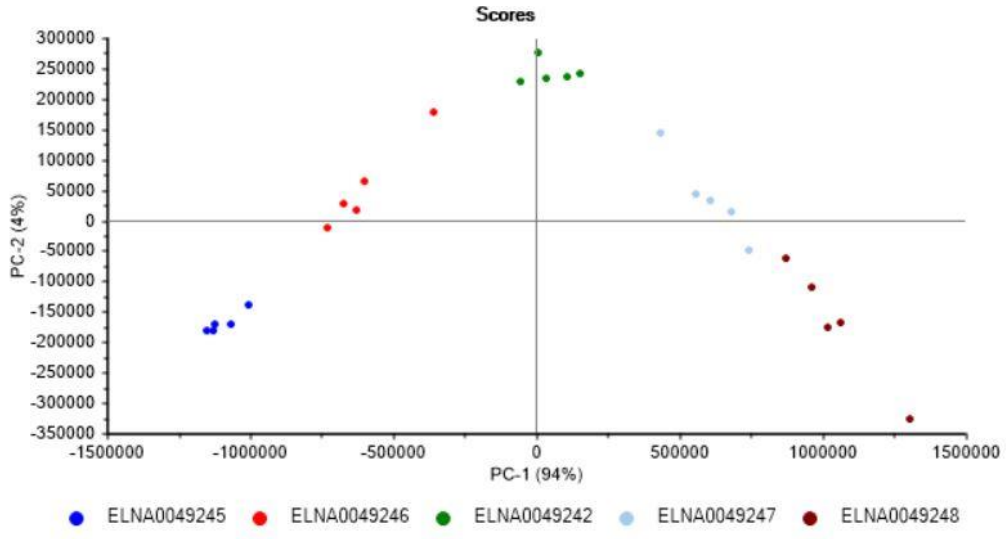


Figure 44: Scores plot for experiments B1-B5 based on PCA. The x-axis and y-axis represent the first and second components, capturing 94% and 4% of the total variance, respectively.

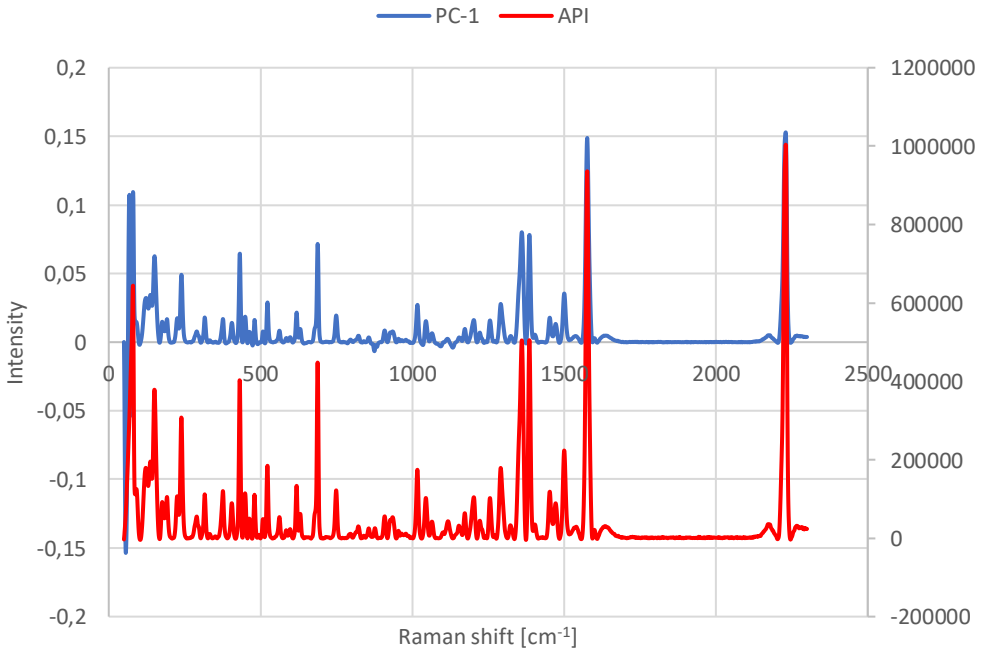


Figure 45: Transmission Raman PC-1 loadings and API A pre-processed spectrum for calibration experiments.

A quantitative evaluation by PLS regression was performed using TRS measurements of the samples and UPLC results as reference values. The following parameters were used:

- Algorithm used: NIPALS, Validation method: Cross validation, Cross validation method: Random with 20 segments, Number of Calibration samples used: 25, Total number of factors: 7, Factors suggested by model: 1, Optimal number of factors: 1

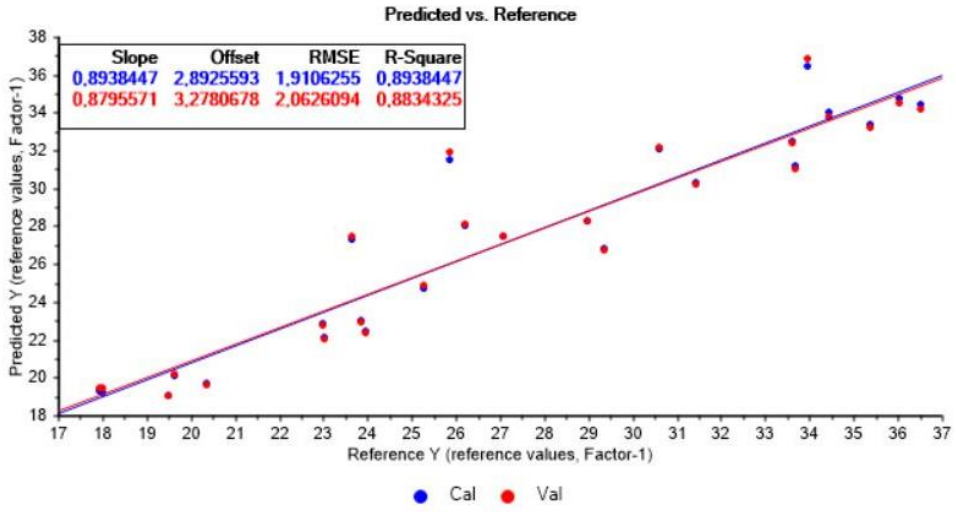


Figure 46: True vs predicted plot for Raman model for validation API A trials. The plot shows the relationship between actual (x-axis) and predicted (y-axis) values from the PLS model. The diagonal line represents perfect predictions (y = x). Points near the line indicate better model accuracy.

The PLS model showed a R-square of 0.88 and a prediction error of 2.06%.

When overlaying the PLS weighted regression coefficients with the API pre-processed spectrum (*Figure 47*), shows a correlation as weighted regression coefficients match API characteristic Raman peaks. The regression vector of PLS and the loadings of PCA have the same shape, this related to the uncorrelated composition of the five blends, allowing a to separate the different ingredients to different components.

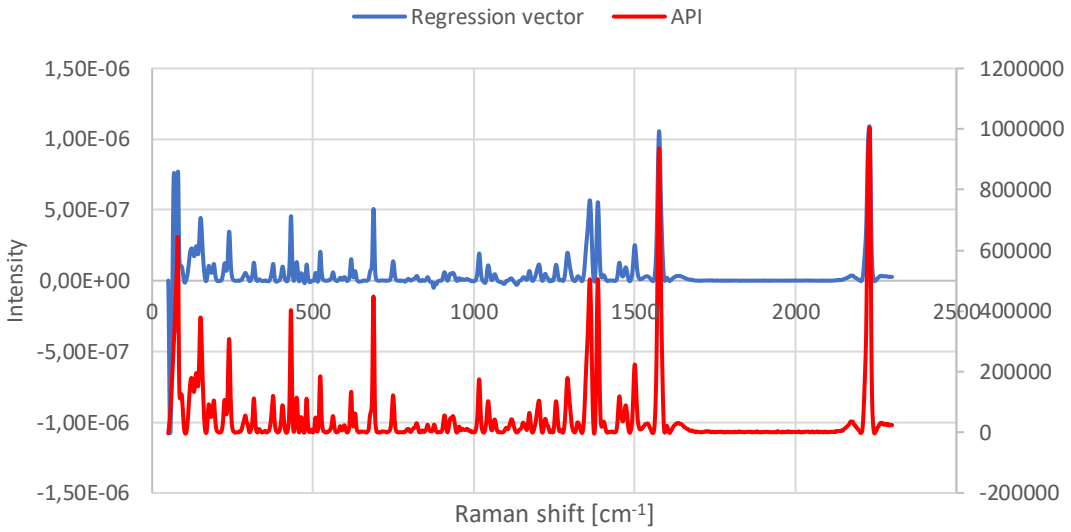


Figure 47: Transmission Raman PLS weighted regression coefficients and API A pre-processed spectrum for calibration experiments.

Predictions were made using the PLS model developed, for each batch there was three predictions, and these were averaged to have one result per batch and compared with UPLC, these results are summarized in Table 28.

The prediction results are very similar to the results obtained by UPLC analysis, which is a good indicator that Raman might be used as an alternative reference analytics technique to determine this API concentration.

Table 28: Comparison between Transmission Raman PLS predictions and UPLC results for experiments B1-B5.

Experiment	***Nominal value [%]***	***Prediction (average five samples) [%]***	***Prediction Std [%]***	***UPLC (average per batch) [%]***	***UPLC Std [%]***
B1	18.94	19.49	0.38	19.08	0.95
B2	22.99	23.03	0.90	23.82	0.83
B3	27.05	27.58	0.52	27.05	2.06

B4	31.11	31.52	0.75	31.03	2.85
B5	35.17	34.62	1.04	35.26	0.96

A paired samples t-test was conducted, and the results are presented in the Table below. The p-value obtained as it is greater than the significance level of 0.05 indicates that difference between predictions and actual values is not statistically significant. Regarding the calibration trials, these results show that once again Transmission Raman predictions are more accurate and closer to UPLC method for the whole range of concentrations.

Table 29: Results of the paired t-test comparing Transmission Raman predictions of calibration data set with UPLC values (API A). The table presents the mean, variance, degrees of freedom (df), *t*-value and *p*-value. A *p*-value < 0.05 indicates statistical significance.

	Mean	*Variance*	*df*	*t-value*	*p-value (two-tail)*
Prediction	27.25	37.67	4	-2.42×10^{-15}	1
Reference	27.25	39.25			

3.2.4.2. Validation Experiments

The samples taken after blending were measured in Transmission Raman using the same method that was created for the first trials.

Similarly to NIR, the validation trial was also added to the calibration data set and a PCA with the following parameters was performed (Algorithm used: NIPALS, Validation method: Leverage, Number of Calibration samples used: 30, Total number of components: 5, Components suggested by model: 3, Optimal number of components: 3).

In *Figure 48*, the scores plot of the PCA is presented and it´s possible to see that the validation trial appears close to experiment BV1 as expected since both have a drug load of 27.05%.

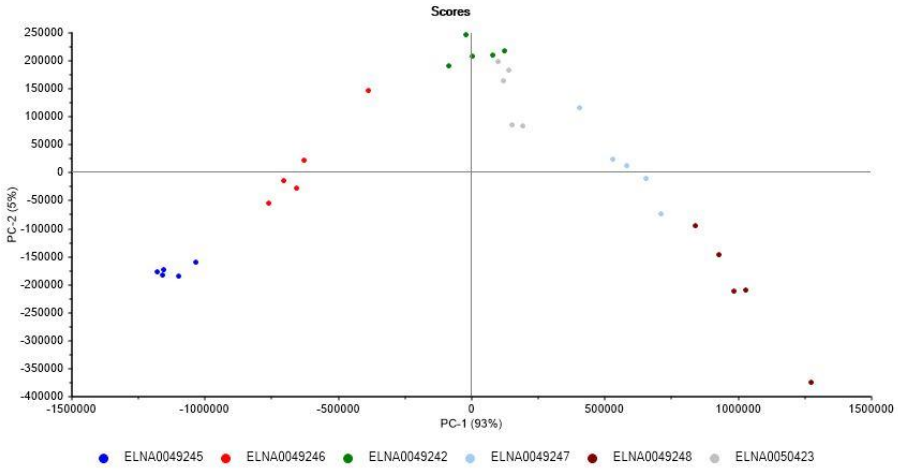


Figure 48: Scores plot for experiments B1-B5 plus validation experiment BV1 based on PCA. The x-axis and y-axis represent the first and second components, capturing 93% and 5% of the total variance, respectively.

A prediction was made using the Raman PLS model previously developed, there was five predictions, and these were averaged to have one result and compare with UPLC, these results are summarized in Table 30.

Regarding the Raman prediction, the mean result batch is fitting the expectation and is not significantly different from the UPLC result obtained. Also the prediction is closer to the nominal value.

Table 30: Comparison between Transmission Raman PLS predictions and UPLC results for experiment BV1.

Experiment	***Nominal value [%]***	***Prediction (average five samples) [%]***	***Prediction Std [%]***	***UPLC (average per batch) [%]***	***UPLC Std [%]***
BV1	27.05	26.45	0.57	25.99	3.86

4. Conclusions & Outlook

This thesis main goals were to investigate the sensitivity of NIR as a tool for on-line blend monitoring and Transmission Raman as reference analytics in API quantification. Also develop and validate robust PLS models using NIR and TRS data for predicting drug concentration variations and comparing them to reference HPLC/UPLC measurements.

This work demonstrated that NIR can successfully detect minor changes down to 0.5% between formulations through a wide range of caffeine concentrations.

For caffeine, the predicted concentrations using NIR ranged from 9.24% to 21.60%, while the reference HPLC measurements indicated values between 9.43% and 19.84%. The correlation coefficient for the calibration model was robust, demonstrating strong agreement between the NIR predictions and the actual concentrations, despite slight overestimation in lower concentrations. Transmission Raman predictions were found to be consistent with those obtained via NIR and HPLC, with minimal deviations. The calibration model developed for Transmission Raman spectroscopy showed that once established, it could provide reliable predictions without the need for ongoing recalibration.

For API A, although NIR could monitor blend uniformity, the prediction accuracy was lower than for caffeine, this was associated with the formulation being more complex. The predicted concentrations using NIR ranged from 18.46% to 32.91%, while the reference UPLC measurements obtained values between 19.08% and 35.26%. The Transmission Raman predictions fitted the expectation and were consistent with UPLC results.

Overall, blend uniformity was more accurately determined by Transmission Raman Spectroscopy. The predictions obtained by this method showed, in both formulations, not to be statistically different compared to HPLC/UPLC results across wide ranges of concentrations. Therefore, Transmission Raman presents more accurate and reliable results coming closer to reference analytical techniques.

Further research should be performed, for example, a more robust model could be developed for the API A formulation in order to achieve better predictions. An approach could be instead of monitoring the blend uniformity just by quantifying the API also quantify the excipients.

Overall, this study contributes to the progress of process analytical methods in pharmaceutical industry and will facilitate the implementation of continuous manufacturing.

5. Bibliography

1. Haack, D., Koeberle, M., Schiemenz, W.: Blend uniformity before tabletting and filling. (2014)

2. De Maesschalck, R., Cuesta, F., Nchez, S.Â., Massart, D.L., Doherty, P., Hailey, P.: On-Line Monitoring of Powder Blending with Near-Infrared Spectroscopy. (1998)

3. Koller, D.M., Posch, A., Hörl, G., Voura, C., Radl, S., Urbanetz, N., Fraser, S.D., Tritthart, W., Reiter, F., Schlingmann, M., Khinast, J.G.: Continuous quantitative monitoring of powder mixing dynamics by near-infrared spectroscopy. Powder Technol. 205, 87–96 (2011). https://doi.org/10.1016/j.powtec.2010.08.070

4. Besseling, R., Damen, M., Tran, T., Nguyen, T., van den Dries, K., Oostra, W., Gerich, A.: An efficient, maintenance free and approved method for spectroscopic control and monitoring of blend uniformity: The moving F-test. J Pharm Biomed Anal. 114, 471–481 (2015). https://doi.org/10.1016/j.jpba.2015.06.019

5. Blanco, M., Cueva-Mestanza, R., Cruz, J.: Critical evaluation of methods for end-point determination in pharmaceutical blending processes. Analytical Methods. 4, 2694–2703 (2012). https://doi.org/10.1039/c2ay25379h

6. Flåten, G.R., Ferreira, A.P., Bellamy, L., Frake, P.: PAT within the QbD Framework: Real-time end point detection for powder blends in a compliant environment. J Pharm Innov. 7, 38–45 (2012). https://doi.org/10.1007/s12247-012-9119-9

7. Adam, S., Suzzi, D., Radeke, C., Khinast, J.G.: An integrated Quality by Design (QbD) approach towards design space definition of a blending unit operation by Discrete Element Method (DEM) simulation. European Journal of Pharmaceutical Sciences. 42, 106–115 (2011). https://doi.org/10.1016/j.ejps.2010.10.013

8. Step: Committee for Human Medicinal Products ICH guideline Q8 (R2) on pharmaceutical development. (2017)

9. Kim, E.J., Kim, J.H., Kim, M.S., Jeong, S.H., Choi, D.H.: Process analytical technology tools for monitoring pharmaceutical unit operations: A control strategy for continuous process verification, (2021)

10. Sulub, Y., Konigsberger, M., Cheney, J.: Blend uniformity end-point determination using near-infrared spectroscopy and multivariate calibration. J Pharm Biomed Anal. 55, 429–434 (2011). https://doi.org/10.1016/j.jpba.2011.02.017

11. Roggo, Y., Martinez, L., Peinado, A., Matero, S.: Near infrared spectroscopy for blend uniformity monitoring: An innovative qualitative application based on the coefficient of determination. J Near Infrared Spectrosc. 30, 322–336 (2022). https://doi.org/10.1177/09670335221130430

12. Mohan, S., Momose, W., Katz, J.M., Hossain, M.N., Velez, N., Drennen, J.K., Anderson, C.A.: A robust quantitative near infrared modeling approach for blend monitoring. J Pharm Biomed Anal. 148, 51–57 (2018). https://doi.org/10.1016/j.jpba.2017.09.011

13. Puchert, T., Holzhauer, C. V., Menezes, J.C., Lochmann, D., Reich, G.: A new PAT/QbD approach for the determination of blend homogeneity: Combination of on-line NIRS analysis with PC Scores Distance Analysis (PC-SDA). European Journal of Pharmaceutics and Biopharmaceutics. 78, 173–182 (2011). https://doi.org/10.1016/j.ejpb.2010.12.015

14. Luypaert, J., Massart, D.L., Vander Heyden, Y.: Near-infrared spectroscopy applications in pharmaceutical analysis, (2007)

15. Wargo, D.J., Drennen, J.K.: PHARMACEUTICAL AND BIOMEDICAL ANALYSIS Near-infrared spectroscopic characterization of pharmaceutical powder blends. (1996)

16. Blanco, M., Coello, J., Iturriaga, H., Maspoch, S., De La Pezuela, C.: Near-infrared spectroscopy in the pharmaceutical industry, (1998)

17. Igne, B., Juan, A. De, Jaumot, J., Lallemand, J., Preys, S., Drennen, J.K., Anderson, C.A.: Modeling strategies for pharmaceutical blend monitoring and end-point determination by near-infrared spectroscopy. Int J Pharm. 473, 219–231 (2014). https://doi.org/10.1016/j.ijpharm.2014.06.061

18. Matousek, P., Andrews, D., A. Griffen, J., W. Owen, A.: Recent Advances in Pharmaceutical Analysis Using Transmission Raman Spectroscopy. Spectroscopy. 32, 37–43 (2017)

19. Zhao, X., Wang, N., Zhu, M., Qiu, X., Sun, S., Liu, Y., Zhao, T., Yao, J., Shan, G.: Application of Transmission Raman Spectroscopy in Combination with Partial Least-Squares (PLS) for the Fast Quantification of Paracetamol. Molecules. 27, (2022). https://doi.org/10.3390/molecules27051707

20. Corredor, C.C., Vikstrom, C., Persson, A., Bu, X., Both, D.: Development and Robustness Verification of an At-Line Transmission Raman Method for Pharmaceutical Tablets Using Quality by Design (QbD) Principles, (2018)

21. Esmonde-White, K.A., Cuellar, M., Uerpmann, C., Lenain, B., Lewis, I.R.: Raman spectroscopy as a process analytical technology for pharmaceutical manufacturing and bioprocessing, (2017)

22. Andrews, D., Geentjens, K., Igne, B., McGeorge, G., Owen, A., Pedge, N., Villaumié, J., Woodward, V.: Analytical Method Development Using Transmission Raman Spectroscopy for Pharmaceutical Assays and Compliance with Regulatory Guidelines—Part I: Transmission Raman Spectroscopy and Method Development. J Pharm Innov. 13, 121–132 (2018). https://doi.org/10.1007/s12247-018-9311-7

23. Griffen, J., Owen, A., Matousek, P.: Comprehensive quantification of tablets with multiple active pharmaceutical ingredients using transmission Raman spectroscopy-A proof of concept study. J Pharm Biomed Anal. 115, 277–282 (2015). https://doi.org/10.1016/j.jpba.2015.07.019

24. Andrews, D., Romer, M.: Implementing Transmission Raman Spectroscopy for Fast Content Uniformity Testing: From Feasibility Evaluation to a Validated Release Method. (2021)

25. Paudel, A., Raijada, D., Rantanen, J.: Raman spectroscopy in pharmaceutical product design. Adv Drug Deliv Rev. 89, 3–20 (2015)

26. Silge, A., Weber, K., Cialla-May, D., Müller-Bötticher, L., Fischer, D., Popp, J.: Trends in pharmaceutical analysis and quality control by modern Raman spectroscopic techniques, (2022)

27. Griffen, J.A., Owen, A.W., Burley, J., Taresco, V., Matousek, P.: Rapid Quantification of Low Level Polymorph Content in a Solid Dose Form using Transmission Raman Spectroscopy 2 3. (2016)

28. Li, T., Nogueira, R., de Brito, J., Liu, J.: Quantitative analysis of the influence of fine aggregate's grading on mortar's rheology. Journal of Materials Research and Technology. 25, 310–318 (2023). https://doi.org/10.1016/j.jmrt.2023.05.236

29. Microtac: Microtac MRB. Camsizer X2- partcile size and shape analyzer-microtac

30. Buckland, H.M., Saxby, J., Roche, M., Meredith, P., Rust, A.C., Cashman, V., Engwell, S.L.: Measuring the size of non-spherical particles and the implications 1 for grain size analysis in volcanology. (2020)

6. Appendix

Table 1: Formulation of experiment A2.

Name	***Function***	***Content [%]***
Caffeine	API	12.00
Mannitol	Filler	65.26
MCC	Filler	17.56
PVP K 30	Binder	4.68
Magnesium stearate VS	Lubricant	0.50
Total		100.0

Table 2: Formulation of experiment A3.

Name	***Function***	***Content [%]***
Caffeine	API	13.50
Mannitol	Filler	63.76
MCC	Filler	17.56
PVP K 30	Binder	4.68
Magnesium stearate VS	Lubricant	0.50
Total		100.0

Table 3: Formulation of experiment A4.

Name	***Function***	***Content [%]***
Caffeine	API	14.50
Mannitol	Filler	62.76
MCC	Filler	17.56
PVP K 30	Binder	4.68
Magnesium stearate VS	Lubricant	0.50
Total		100.0

Table 4: Formulation of experiment A5.

Name	***Function***	***Content [%]***
Caffeine	API	15.00
Mannitol	Filler	62.26
MCC	Filler	17.56
PVP K 30	Binder	4.68
Magnesium stearate VS	Lubricant	0.50
Total		100.0

Table 5: Formulation of experiment A6.

Name	***Function***	***Content [%]***
Caffeine	API	15.50
Mannitol	Filler	61.76
MCC	Filler	17.56
PVP K 30	Binder	4.68
Magnesium stearate VS	Lubricant	0.50
Total		100.0

Table 6: Formulation of experiment A7.

Name	***Function***	***Content [%]***
Caffeine	API	16.50
Mannitol	Filler	60.76
MCC	Filler	17.56
PVP K 30	Binder	4.68
Magnesium stearate VS	Lubricant	0.50
Total		100.0

Table 7: Formulation of experiment A8.

Name	***Function***	***Content [%]***
Caffeine	API	18.00
Mannitol	Filler	60.26
MCC	Filler	17.56
PVP K 30	Binder	4.68
Magnesium stearate VS	Lubricant	0.50
Total		100.0

Table 8: Formulation of experiment A9.

Name	***Function***	***Content [%]***
Caffeine	API	20.00
Mannitol	Filler	58.26
MCC	Filler	17.56
PVP K 30	Binder	4.68
Magnesium stearate VS	Lubricant	0.50
Total		100.0

Table 9: Caffeine concentration determined by HPLC for experiments A1-A9.

Experiment	**Mass weighted [mg]**	**Caffeine concentration [%]**
A1_sample 1	223.20	9.22
A1_sample 2	224.74	9.48
A1_sample 3	224.44	9.59
A2_sample 1	187.7	11.90
A2_sample 2	186.03	11.99
A2_sample 3	187.16	11.53
A3_sample 1	169.51	13.34
A3_sample 2	165.13	12.83

A3_sample 3	164.67	13.49
A4_sample 1	154.21	14.88
A4_sample 2	153.92	13.64
A4_sample 3	154.28	14.54
A5_sample 1	149.69	14.25
A5_sample 2	148.66	15.46
A5_sample 3	150.35	15.01
A6_sample 1	144.71	15.25
A6_sample 2	144.01	15.26
A6_sample 3	146.23	14.88
A7_sample 1	135.16	15.73
A7_sample 2	134.81	14.42
A7_sample 3	135.24	15.96
A8_sample 1	124.28	17.93
A8_sample 2	124.2	17.73
A8_sample 3	124.51	18.21
A9_sample 1	110.75	20.10
A9_sample 2	111.91	19.90
A9_sample 3	113.93	19.52

Table 10: Caffeine concentration determined by HPLC for experiments AV1-AV4.

Experiment	**Mass weighted [mg]**	**Caffeine concentration [%]**
AV1_sample 1	149.17	14.38
AV1_sample 2	149.88	14.77
AV1_sample 3	149.04	14.89
AV2_sample 1	149.46	15.31
AV2_sample 2	150.14	15.33
AV2_sample 3	149.44	15.09
AV3_sample 1	149.55	15.00
AV3_sample 2	149.38	14.40
AV3_sample 3	149.24	14.31
AV4_sample 1	149.58	13.16
AV4_sample 2	149.98	14.17
AV4_sample 3	149.93	11.61

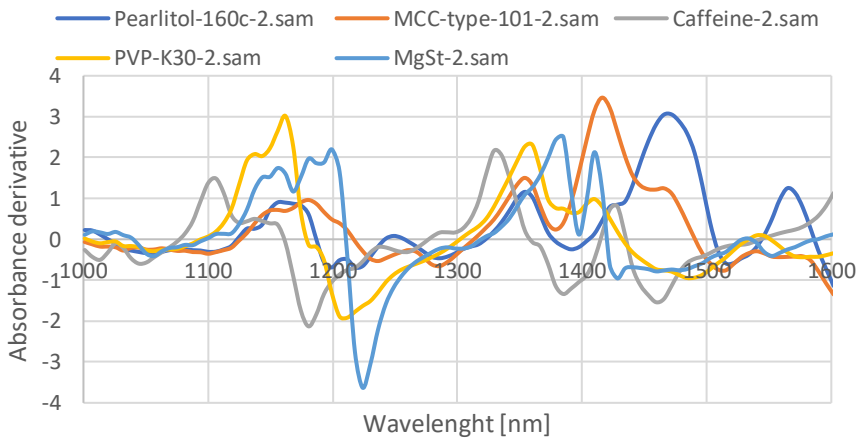


Figure 1: Preprocessed spectra of raw materials measured with Viavi MicroNIR PAT-W (caffeine formulation).

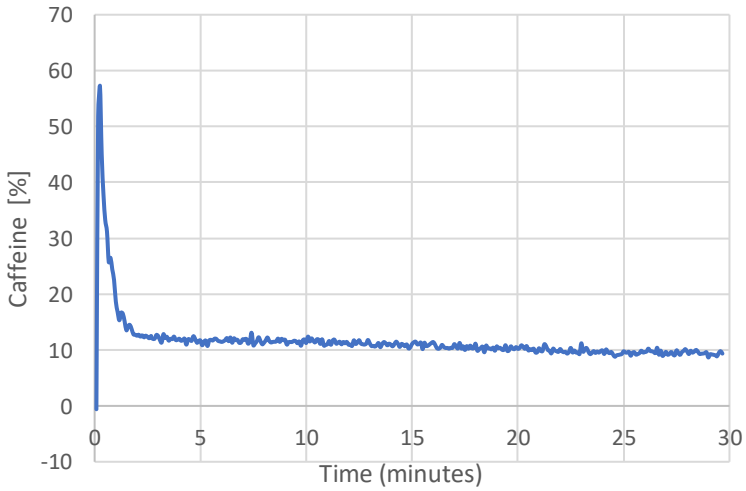


Figure 2: Prediction of the caffeine concentration for experiment A1.

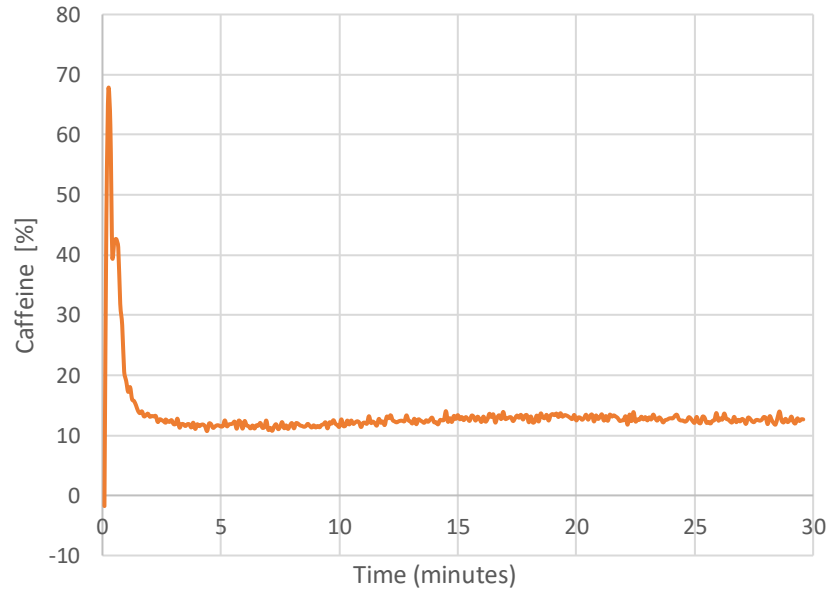


Figure 3: Prediction of the caffeine concentration for experiment A2.

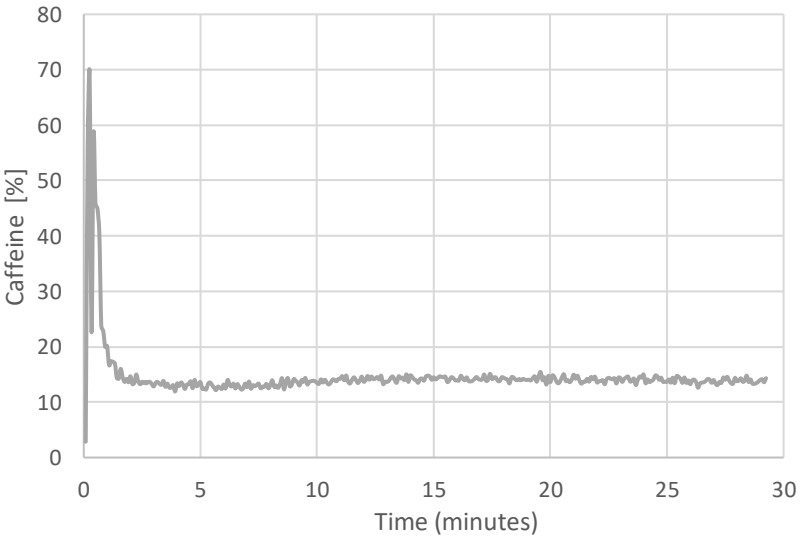


Figure 4: Prediction of the caffeine concentration for experiment A3.

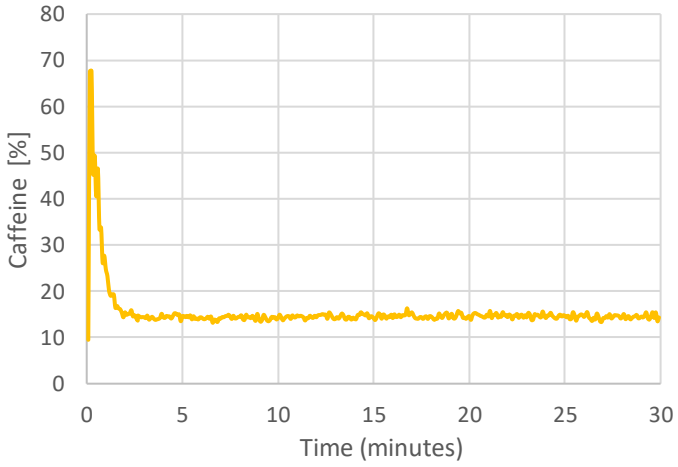


Figure 5: Prediction of the caffeine concentration for experiment A4.

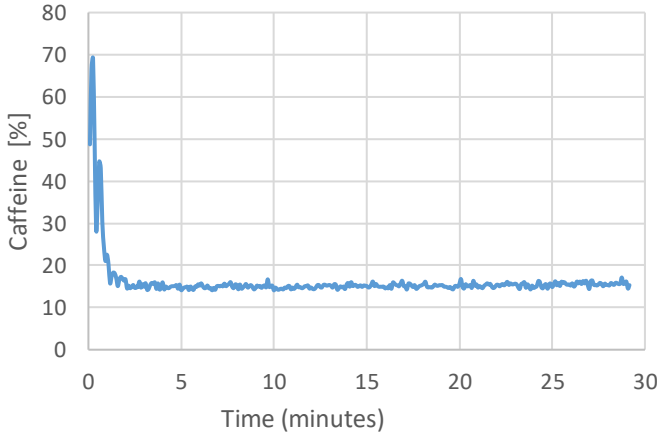


Figure 6: Prediction of the caffeine concentration for experiment A5.

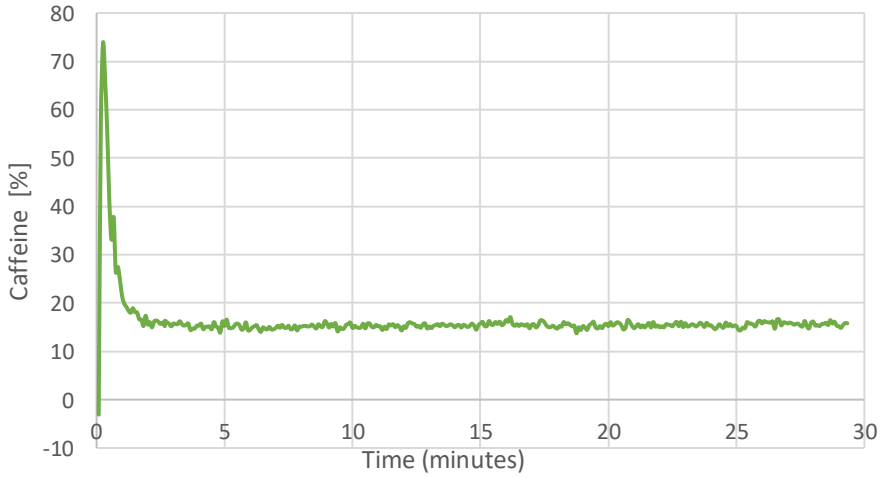


Figure 7: Prediction of the caffeine concentration for experiment A6.

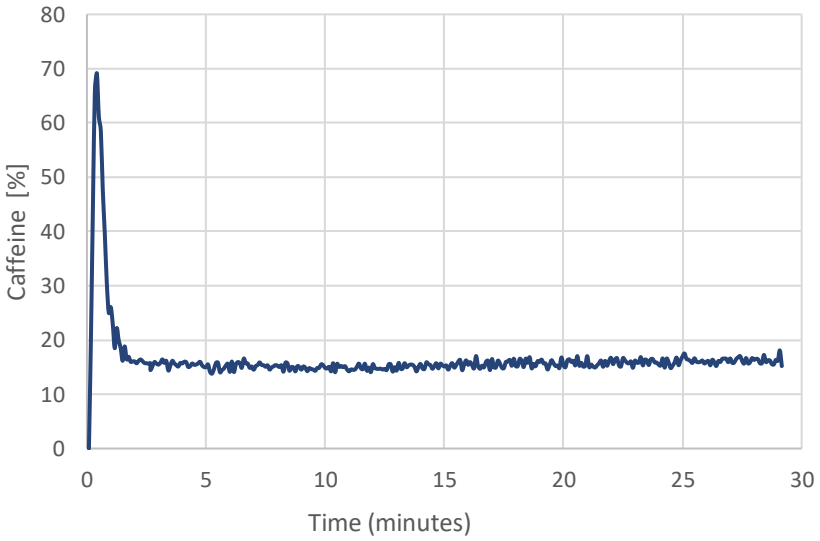


Figure 8: Prediction of the caffeine concentration for experiment A7.

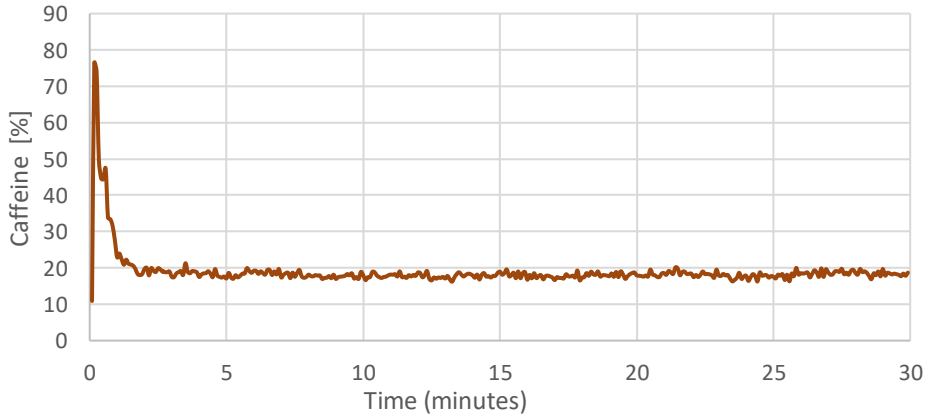


Figure 9: Prediction of the caffeine concentration for experiment A8.

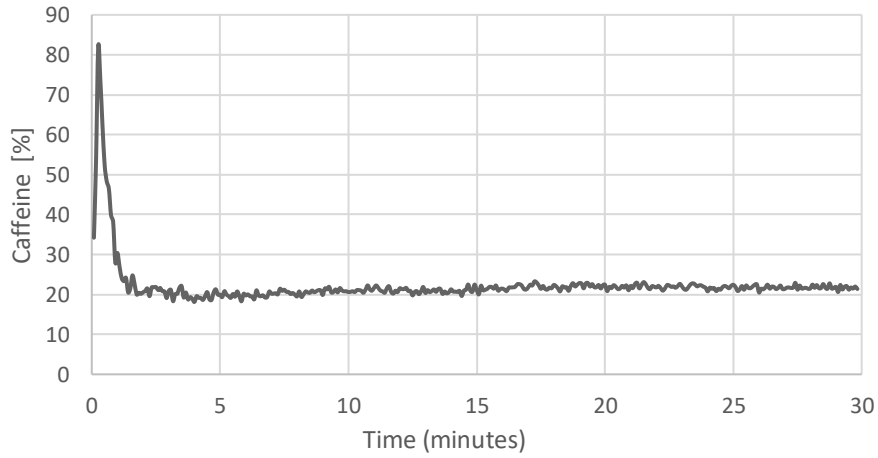
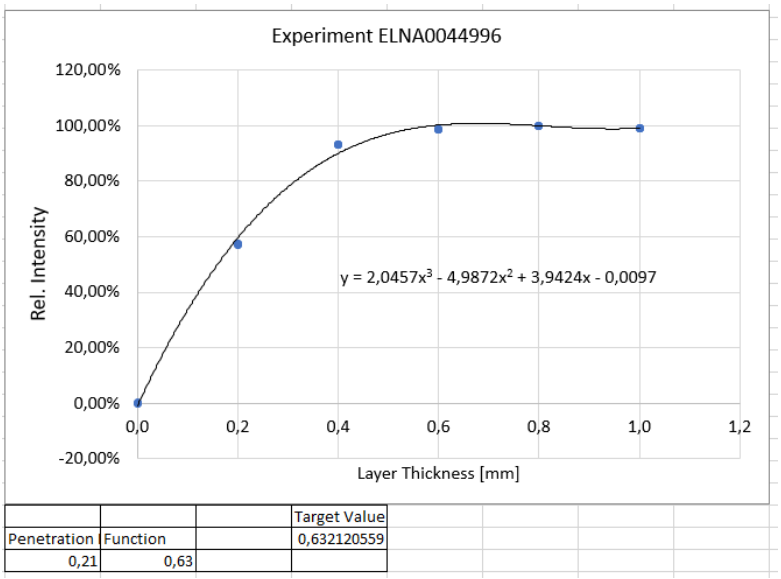
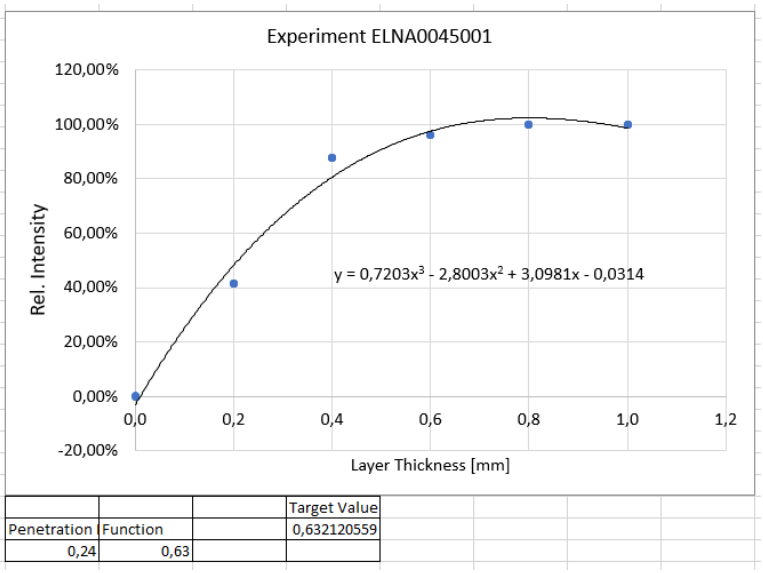


Figure 10: Prediction of the caffeine concentration for experiment A8.



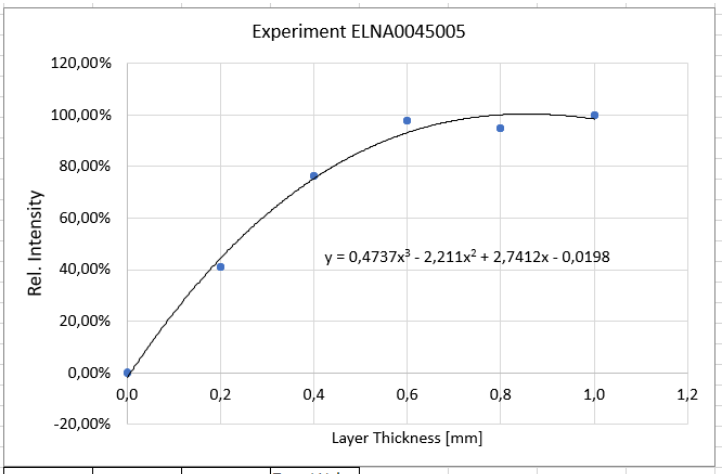
			Target Value
Penetration	Function		0,632120559
0,21	0,63		

Figure 11: Penetration depth of experiment A1.



			Target Value
Penetration	Function		0,632120559
0,24	0,63		

Figure 12: Penetration depth of experiment A5.



			Target Value
Penetration	Function		0,632120559
0,31	0,63		

Figure 13: Penetration depth of experiment A9.

Table 11: API concentration determined by UPLC for experiments B1-B5.

Experiment	**Mass weighted [mg]**	**API concentration [%]**
B1_sample 1	2.012	19.63
B1_sample 2	2.092	17.94
B1_sample 3	2.019	19.48
B1_sample 4	2.025	18.36
B1_sample 5	2.047	18.00
B2_sample 1	2.047	23.95
B2_sample 2	2.002	23.03

B2_sample 3	2.007	22.98
B2_sample 4	2.012	23.84
B2_sample 5	2.000	25.27
B3_sample 1	2.014	26.22
B3_sample 2	2.012	28.98
B3_sample 3	2.024	27.07
B3_sample 4	2.040	29.35
B3_sample 5	2.072	23.65
B4_sample 1	2.061	33.60
B4_sample 2	1.997	30.59
B4_sample 3	2.062	25.86
B4_sample 4	2.040	33.67
B4_sample 5	2.001	31.41
B5_sample 1	2.039	34.42
B5_sample 2	2.006	36.03
B5_sample 3	2.002	35.39
B5_sample 4	2.018	33.96
B5_sample 5	2.008	36.53

Table 12: API concentration determined by UPLC for experiment BV1.

Experiment	**Mass weighted [mg]**	**API concentration [%]**
BV1_sample 1	2.052	26.25
BV1_sample 2	2.060	24.53
BV1_sample 3	2.019	27.34
BV1_sample 4	2.072	31.85
BV1_sample 5	2.019	19.98

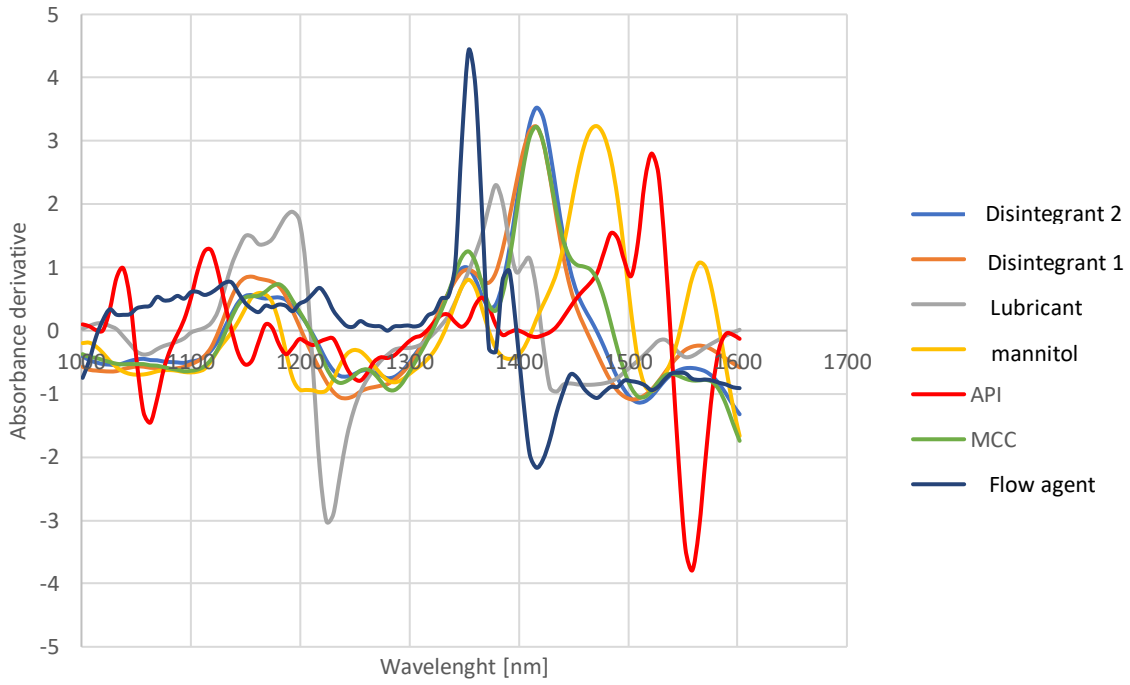


Figure 14: Preprocessed spectra of raw materials measured with Viavi MicroNIR PAT-W (API A formulation).

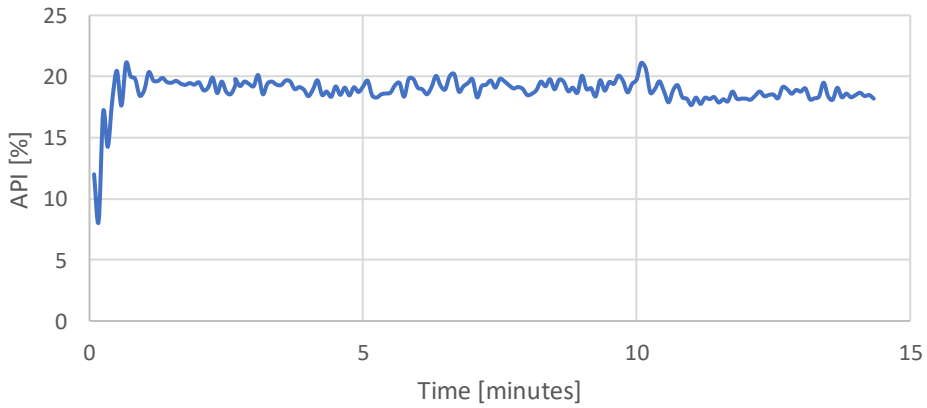


Figure 15: Prediction of the API A concentration for experiment B1.

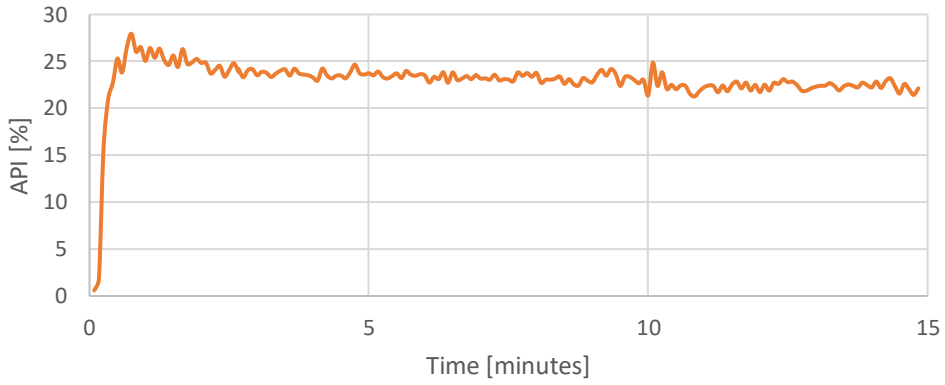


Figure 16: Prediction of the API A concentration for experiment B2.

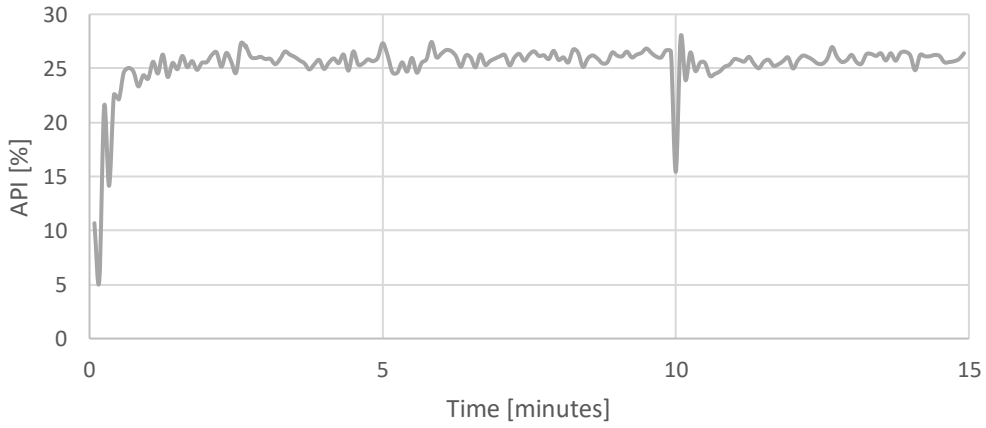


Figure 17: Prediction of the API A concentration for experiment B3.

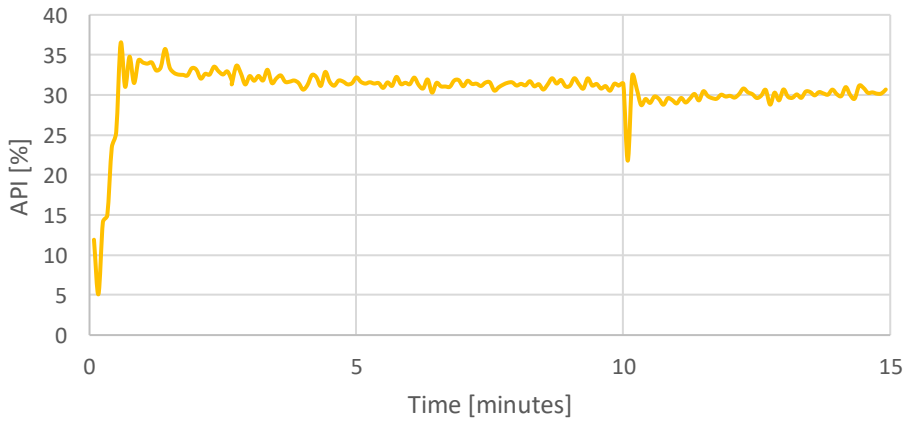


Figure 18: Prediction of the API A concentration for experiment B4.

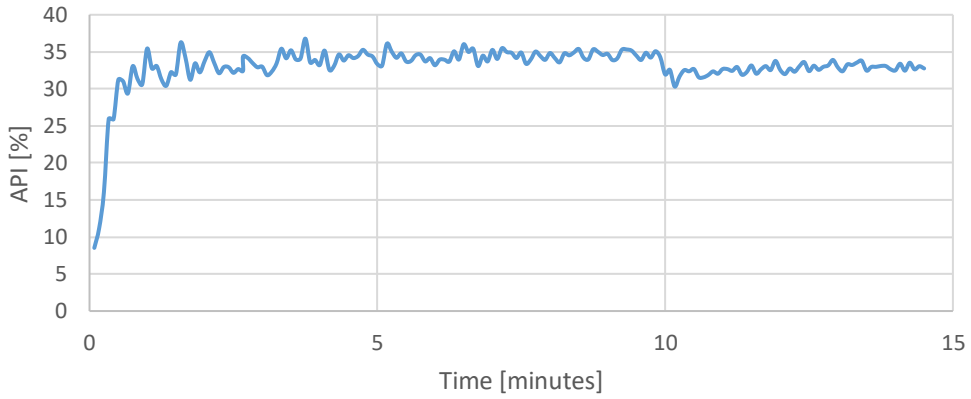


Figure 19: Prediction of the API A concentration for experiment B5.