Advanced Computational Methods for Time Series Analysis of Pharmaceutical Continuous Bulk Solid Feeding and Tableting

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Stefan Klinken
aus Mönchengladbach

Düsseldorf, Januar 2024

aus dem Institut für Pharmazeutische Technologie und Biopharmazie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Jörg Breitkreutz
Korreferent: Prof. Dr. Dr. h. c. Peter Kleinebudde

Tag der mündlichen Prüfung: 13.05.2024

“This being said, the question that remains is: what new scientific knowledge in the field of pharmaceutical sciences can be acquired by such a study. And after reading the article, the answer is: none.”

(Reviewer #2, 2023)

“The ai method applied performs remarkably outstanding on the dataset and provides a method for very good and precise interpolation in the design space! (...) The topic is of very high interest to the community and the industry.”

(Reviewer #1, 2023)

Table of content

List of publications

Scientific original papers:

1. Quodbach, J., Bogdahn, M., Breitkreutz, J., Chamberlain, R., Eggenreich, K., Elia, A. G., Gottschalk, N., Gunkel-Grabole, G., Hoffmann, L., Kapote, D., Kipping, T., Klinken, S., Loose, F., Marquetant, T., Windolf, H., Geißler, S., Spitz, T., (2022). Quality of FDM 3D printed medicines for pediatrics: Considerations for formulation development, filament extrusion, printing process and printer design. Ther. Innov. Regul. Sci., 56, 910 – 928.
2. Klinken, S., Quodbach, J., (2022). Sums of amplitudes analysis – A new non-parametric classification method for time series deviation evaluation in pharmaceutical processes. Powder Technol., 412, 118003.
3. Grumann, H. D., Klinken, S., Kleinebudde, P., (2022). Influence of temperature on the compression behavior of pharmaceutical excipients. Int. J. Pharm., 628, 122305.
4. Grumann, H. D., Klinken, S., Kleinebudde, P., (2022). Evaluation of in-die compression data for a deeper understanding of altered excipient properties upon temperature rise. AAPS PharmSciTech, 24, 89.
5. Kokott, M., Klinken, S., Breitkreutz, J., Wiedey, R., (2023). Downstream processing of amorphous solid dispersions into orodispersible tablets. Int. J. Pharm., 631, 122493.
6. Berkenkemper, S., Klinken, S., Kleinebudde, P., (2023). Investigating compressibility descriptors for binary mixtures of different deformation behavior. Powder Technol., 424, 118571.
7. Berkenkemper, S., Klinken, S., Kleinebudde, P., (2023). Multivariate data analysis to evaluate commonly used compression descriptors. Int. J. Pharm., 637, 122890.
8. Lura, V., Klinken, S., Breitkreutz, J., (2023). Challenges in the transfer and scale-up of mini-tablets: A case study with losartan potassium. Eur. J. Pharm. Biopharm., 192, 161-173.
9. Lura, V., Klinken, S., Breitkreutz, J., (2023). A systematic investigation of external lubrication of mini-tablets on a rotary tablet press with focus on the tensile strength. Eur. J. Pharm. Biopharm., 198, 114236.
10. De Bisshop, J., Klinken, S., (2023). Prediction of the tensile strength of tablets using LSTM networks on compression profiles. Int. J. Pharm., 645, 123280.
11. Lück, M., Klinken, S., Kleinebudde, P., (2024). In-line elastic recovery measurement in roll compaction for the prediction of ribbon solid fraction. J. Pharm. Sci., 4, 1020-1028

List of conference and seminar contributions

Poster presentations:

1. Klinken, S., Quodbach, J., Sums of amplitude analysis in frequency domain of bulk solid dosing processes, 13th World Meeting of Pharmaceutics, Biopharmaceutics and Pharmaceutical Technology, Rotterdam, 28.03.-31.03.2022.
2. Kokott, M., Klinken, S., Wiedey, R., Breikreutz, J, The interplay of two poorly soluble drugs in ASD combination products, 13th World Meeting of Pharmaceutics, Biopharmaceutics and Pharmaceutical Technology, Rotterdam, 28.03.-31.03.2022.
3. Kokott, M., Klinken, S., Wiedey, R., Breikreutz, J., Orodispersible tablets as promising alternative for paediatric HIV combination therapy, 14th European Paediatric Formulation Initiative Conference, Rome, 20.09-22.09.2022.
4. Klinken, S., Berkenkemper, S., Quodbach, J., Evaluation of the curve shape of compression profiles via curve stripping-based PCA, 14th European Conference on Pharmaceutics, Marseille, 20.03-21.03.2023.
5. Schulzen, A., Klinken, S., Niesbach, J., Klauke, P., Haase, I., Spitz, T., Quodbach, J., Continues 3D-printing of solid lipids - A proof of concept study, 14th European Conference on Pharmaceutics, Marseille, 20.03-21.03.2023.
6. Lück, M., Klinken, S., Kleinebudde, P., In-line determination of ribbon solid fraction in roll compaction using laser triangulation measurement, 4th Continuous Manufacturing Conference, Lisbon, 09.05-10.05.2023.
7. Klinken, S., Simulation driven concepts to optimize the run order in DoE for application in pharmaceutics and biopharmaceutics, 10th Polish-German Symposium on Pharmaceutical Science, Düsseldorf, 22.09-23.09.2023.

Oral presentations:

1. **Klinken, S.**, Berkenkemper, S., Quodbach, J., Multivariate non-parametric force-time data evaluations to determine material concentrations during tableting, 16th Pharmaceutical Solid State Research Cluster Symposium, Helsinki, 20.06-22.06.2022.
2. **Kokott, M.**, Klinken, S., Wiedey, R., Breitkreutz J., Challenges in downstream processing of orodispersible tablets containing ASDs, 16th Pharmaceutical Solid State Research Cluster Symposium, Helsinki, 20.06-22.06.2022.
3. **Klinken, S.**, Formulierungsaspekte beim Trockengranulieren. APV basics - Praktikum Trockengranulieren, Ennigerloh, 18.10.- 19.10.2022.
4. **Uth, L.-L. D.**, Klinken, S., Breitkreutz, J., Investigation of various factors affecting the wet film thickness in continuous oral film manufacturing, 17th Pharmaceutical Solid State Research Cluster Symposium, Cambridge, 29.08-01.09.2023.
5. **Klinken S., Muschert S., De Beer T., Markl D., Goodwin D., Van Den Ban S**., Discussion on formulation development beyond DoE, 17th Pharmaceutical Solid State Research Cluster Symposium, Cambridge, 29.08-01.09.2023.

List of abbreviations

ANN	Artificial neural network
API	Active Pharmaceutical Ingredient
CMA	Critical Material Attribute
CPP	Critical Process Parameter
CQA	Critical Quality Attribute
FDA	United States Food and Drug Administration
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
LSTM	Long short-term memory
NAMUR	Normenarbeitsgemeinschaft für Meß- und Regeltechnik in der chemischen Industrie
PAT	Process analytical technology
PCA	Principal component analysis
QbD	Quality by Design
ReLu	Rectified linear unit
SAA	Sums of amplitudes analysis

1 Introduction

1.1 The linkage of pharmaceutical research and data science

In pharmaceutical science, as in every other scientific discipline, data evaluation is crucial for deriving meaningful insights. The pharmaceutical industry has increasingly emphasized data science aspects in recent years. This trend can undoubtedly be linked to the growing interest in the topics of continuous manufacturing, process analytical technology (PAT), and quality by design (QbD) [1].

In pharmaceutical manufacturing, the continuous production of drug products has become increasingly prominent within the last two decades [2]. The term 'continuous manufacturing' in this context describes the production of pharmaceuticals were the raw materials, and intermediate products are charged into the system over the full process time. The product on the other hand is continuously discharged from the process. In contrast, in batch processes, the raw materials and intermediate products are fed into the manufacturing devices at the beginning of the process. After the process is finished, the product is discharged [3]. While the subprocesses in batch manufacturing are spatially separated, they are only temporally separated in continuous manufacturing. Up to now, batch processes remain the standard method for drug production in the pharmaceutical industry, while the continuous processing of goods is widespread in other industries, such as the chemical or food industry. The innovations into the pharmaceutical manufacturing can likely be attributed to the stringent regulatory barriers in the pharmaceutical sector. Nonetheless, postponing modernization of production is a risk for drug safety [4] and cost effectiveness. Initiated by the United States Food and Drug Administration (FDA), international regulatory bodies and organizations have implemented continuous manufacturing into their guidelines [5]. One example is the Q13 guideline of the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use [6]. Another significant advancement is the implementation of QbD aspects in pharmaceutical manufacturing. To ensure drug safety, it is essential to control the Critical Quality Attributes (CQA) of each produced batch before its release. In this context, the quality of a manufactured product refers to the conformity of the properties of the product with the pre-defined specifications. The suitability of the product for the intended application arises from the correct definition of the specifications. The specific CQAs strongly depend on the type of drug product [7]. Ensuring these required attributes is the goal of the overall pharmaceutical quality management. This is typically accomplished by thoroughly testing the manufactured drug products using pre-defined testing procedures. This quality-by-testing approach is time-consuming, costly, and lacks flexibility [8]. QbD refers to ensuring drug safety through thoroughly planned and monitored process management aiming at the real-time release of pharmaceutical goods. The use of PAT is indispensable in this regard [3]. PAT includes any form of sensing technology that can be integrated into manufacturing processes to provide information about the resulting drug products, intermediates and the status of the ongoing process [9].

As a result, the automatic and computational pre-processing and evaluation of the data, as well as the understanding of the properties of the data, become increasingly important. Due to the temporal continuity of the continuous manufacturing processes and thus the data, these possess special properties. Such data are referred to as time series. In the analysis of this type of data, not only classical statistical parameters such as mean and standard deviation of a sample are of interest, but also the temporal correlation of individual values [10]. The data series can be of different dimensionality and orders [11]. Zero-order time series are one-dimensional matrices, resulting from the acquisition of one measure over time. The measurement of multivariate data at multiple time points results in two- or three-dimensional matrices and thus first and second order time series. While in other fields of science time series analysis is attributed high importance, in some areas of data evaluation in the pharmaceutical context, preferably conventional methods are used. These conventional methods include basic hypothesis testing, simple or multiple linear regression or the consideration of measures of the spread

of the distribution of values within the data series. This may lead not only to the loss of information from the data, but also to incorrect conclusions about Critical Material Attributes (CMA) and Critical Process Parameters (CPP). While the utilization of data science aspects is common in drug design and development [12–15], pharmaceutical manufacturing and formulation development is still in its early stages in this context. Approaches demonstrate that the use of big data in the development of processes and formulations can be substantially beneficial [16]. Further examples of latest developments in the area of data handling pertain to the data fusion of different sensing systems [17], data integrity [18], big data handling and machine learning [19].

Even though there has been an increase in the formulation of biologics in recent years, which usually have to be parenterally administered in liquid form, the proportion of solid oral dosage forms in the market remains substantial [20]. Beside capsules and granules for oral intake, tablets are by the far the most common solid dosage form for peroral administration of active pharmaceutical ingredients (API) [21]. Despite the widespread practice of pharmaceutical tablet manufacturing for over 100 years, research on tablet fabrication processes remains incomplete. This research primarily includes the study of properties and behavior of bulk solids like powders or granules. Bulk solids represent systems whose properties are not easy to characterize. Two examples are the flowability and the particle size distribution, for which numerous measurement methods exist. Each of these methods yields slightly different values, meaning that a ‘true value’ can only be given in reference to the specific measurement procedure. In tablet manufacturing, the flow properties of materials are crucial, but also the material's behavior during the actual compression phase is of high significance. Both, the analysis of dosing processes of bulk solids and of the evaluation of compression of bulk solid material results in time series data. This current thesis aims to demonstrate the development and implementation of advanced evaluation procedures for pharmaceutical time series data with focus on bulk solid flow and compression. In the following chapters, the fundamentals of bulk solid dosing and tableting, as well as the data analysis procedures, are discussed.

1.2 Continuous feeding of bulk solids

1.2.1 Introduction into continuous bulk solid feeding

The terminology used in this thesis regarding bulk solids refers to the book “Powders and bulk solids behavior, characterization, storage and flow” [22]. A bulk solid is characterized as a collection of individual solid particles, with these particles forming a dispersed phase within an ambient gas medium. This category encompasses for example bulks of powders, granules, pellets, or tablets.

Within the processing of bulk solids in the context of pharmaceutical continuous manufacturing, dosing of the solid materials is in most cases the initial step. Due to the nature of continuous production processes, the feeding of the material must be ensured continuously and homogeneously over the entire process time. Deviations in the input-stream of the powder into the later processes can influence the CQAs of the products [23]. For the homogeneous feeding of bulk solids, devices are used that are capable of continuously dispensing the stored material through a suitable mechanism. Frequently in continuous manufacturing processes, several of these bulk solid feeders have to be used, which feed either pure substances or pre-mixtures into the subsequent process [24–26]. Employing multiple dosing units and a continuous blender obviates the need for the cost-intensive pre-mixing of components [27]. Bulk solid feeders are not only found as the initial steps of continuous manufacturing lines. They can also be used within the production line for example to buffer the product stream between different processing units or to add additional components like lubricants or APIs to an intermediate product [28,29]. After a granulation step, extra granular components, like lubricants can be fed into the process. Within tablet presses, bulk solid dosing processes typically employ so-called force feeders. They use paddles to ensure a homogeneous transportation of the bulk solid within the feed-frame. The dies of the tableting machine are filled gravimetrically, depending on the density of the bulk solid within the feed frame. Various studies have been published investigating these types of feeders, including the influences of device geometries or settings [30–32]. There are also different approaches involving computer-aided simulation methods [33]. However, the following discussion will not consider these types of feeding devices. The overview is limited to pharmaceutically used continuous bulk solid feeders where the device doses free-falling bulk solids into a downstream process.

Continuous bulk solid dosing devices can be divided into different classes [34]. One classification is based on the mechanical method to dispense the material. While belt feeders are common in other industrial sectors, pharmaceutical dispensers usually work with dosing screws or vibrating chutes. Screw-based dosing systems can be further subdivided into single-screw and twin-screw feeders. They represent the most important group of bulk solid feeders in the continuous manufacturing of solid oral dosage forms. The geometry of the screws has a decisive influence on the transport of the goods and have to be adapted to the material properties. In the process of selecting the appropriate screw geometry for bulk solid dosing, multiple factors need to be considered. In general, the higher the free volume within the screw channels, the higher the throughput of material per rotation of the screws. The pitch of the flights is directly proportional to the free volume, making it a critical parameter. Alterations, such as deepening or widening of the channels can change the free volume. Some materials exhibit a tendency for compacting within the dosing screws, which can be counteracted by progressively increasing the free volume along the length of the screw. This can be accomplished either by varying the pitch over the length of the screws or by employing a conical shaft design. Additionally, the geometry of the screws directly affects the maximum possible dosing accuracy [34] and the maximum dosing rate. The smaller the screw diameters the lower is the maximum dosing rate but the higher is the accuracy of the dosing process. Figure 1 provides various examples of dosing screws of two different types of twin-screw feeders.

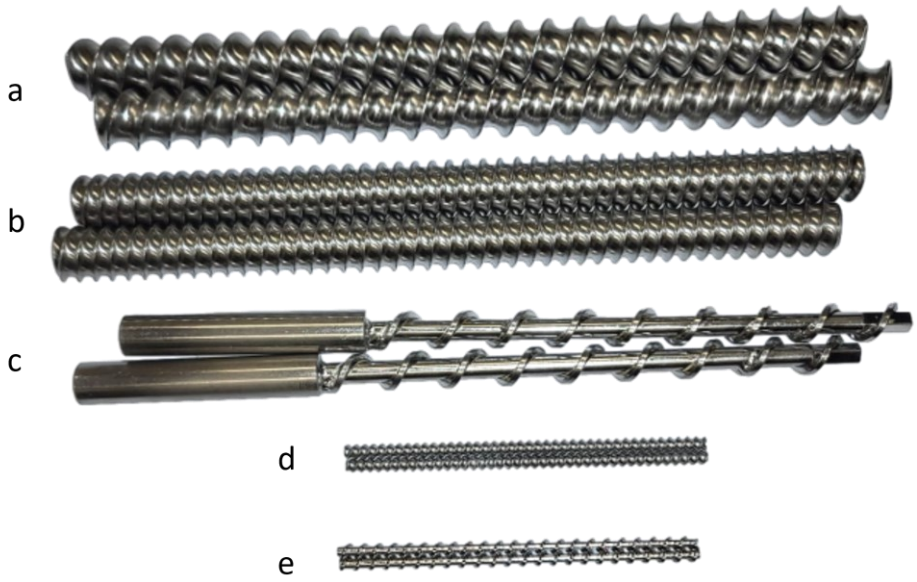

Figure 1 a-c: Dosing screws of a diameter of 20 mm, a: wide pitch double-flight intermeshing, b: narrow pitch double-flight intermeshing, c: single-flight non-intermeshing, d and e: dosing screws of a diameter of 5 mm, d: double-flight intermeshing, e: single-flight intermeshing with conical shafts

During the rotation of the screws, the material is released in fast frequent pulses, depending on the pitch of the dosing screws. These fluctuations can be reduced by the implementation of a discharge screen at the outlet of the dosing device. However, the utilization of such a dosing aid is often not feasible, primarily due to the tendency of many materials to compact within the dosing screws. The use of discharge screens can augment these issues, potentially leading to process interruptions. Static eliminators can improve the flowability of materials like colloidal silicon dioxide [29].

Vibration-chute-based dispensers are less common in manufacturing than in analytical equipment. They are frequently used, for example, in the measurement and sub-division of bulk solid samples. For the dosing of fine cohesive powders or low-dosed APIs a new subcategory of vibration-based feeders is described in the literature [35]. Similar studies describe the use of pneumatic systems [36] or scarping feeders [37,38]. However, up to now there is hardly any industrial application of the latter devices. Given the significance of screw-based dosing devices, future descriptions in the thesis will exclusively consider screw-based bulk solid feeders.

Apart from the classification of feeders by the transport mechanism, a distinction can be made between volumetric and gravimetric devices [29]. Volumetric bulk solid feeders dose the material based on a fixed set value, such as the rotation speed of the screws. Prior to each dosing process, the screw speed must be calibrated to align with the specified dosing rate. The actual dosing rate during the feeding process can-not be measured with volumetric devices. Changes in the material behavior within the storage vessel can lead to dosing irregularities. An example is the reduced material density in the dosing screws, resulting from the depletion of the storage vessel over extended periods of operation.

Gravimetric dosing devices have built-in mechanisms for measuring the mass flow of the dispenser. Usually, the dosing unit is located on a load cell. This internal balance continuously captures the weight of the device. In addition to the weight, data on screw speed and time are recorded. By calculating the derivative of the mass over the time, it becomes possible to calculate the flow rate of the dosed material. Internal controllers, usually with proportional and integral components, then enable the calculation of the necessary change in the setting of the screw speed. This results in the dispenser's adaptability to fluctuations in the process. By switching off the controller of the device, gravimetric dosing units can be used in a volumetric mode. Similar to volumetric feeding devices, gravimetric feeders have to be calibrated before every process. In the initial phase of a gravimetric dosing process the devices are driven in the volumetric mode. The reason for this is that at the start of the process, no data are available

for the controller of the device. Gravimetric feeders, therefore, initiate the process volumetrically and switch to gravimetric mode after a short operational period. Additionally, during long-lasting processes, it becomes necessary to refill the feeder's hopper. Since weight changes would directly affect the controller, the system switches back to volumetric mode in this instance. A value from a previously recorded calibration can be targeted, or running averages of the gravimetric dosage settings can be used for the set point during the refilling phase. Studies have shown that, particularly during the refilling phase, there are significant differences in the homogeneity of the dispensed bulk solid quantity compared to regular operation [39]. The refilling can be done either manually or automatically, e.g. through pneumatic conveying.

Gravimetric feeding units are most widespread in the application in pharmaceutical manufacturing lines [40]. In addition to the controllability of the process, the data collection in such devices enables the evaluation of events within the process, the investigation of material dependencies in dosing, and the correlation with product properties. The accuracies of the weighing cells used can be in the range of conventional analytical balances. However, not only due to the complexity of the flow process of bulk solids but also due to artifacts caused by the movement of the mechanics or external influences, such as air turbulence, the analysis of data from bulk solid feeders is by far more challenging. Additionally, the recorded data are usually subjected to multiple internal filter functions within the device. The applied data filtering functions are in most cases not known to the user, which is why gravimetric bulk solid feeders often represent black boxes in terms of the evaluation of the internally collected data. For this reason, the investigation of the dosing rate for bulk solid feeders is commonly conducted using an external catch scale, which is placed beneath the feeder's outlet. Over the course of the dosing process, the bulk solid accumulates on the plate of the external scale. The increase in mass over time can be evaluated accordingly. Figure 2 illustrates the setup of a flat-bottom feeder and the use of an external catch scale.

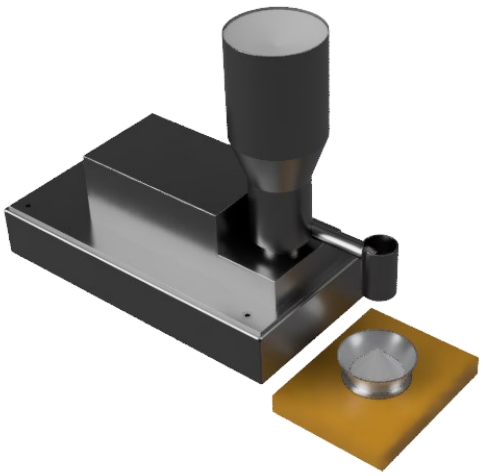

Figure 2: Flat-bottom dosing unit and external catch scale; sketch made using Autodesk Fusion 360

The understanding for the tools to qualify feeder equipment is crucial for the consideration of process deviations caused by the dosing unit. Therefore, in the following subsection, a more detailed overview of the performance evaluations of bulk solid feeders will be provided.

1.2.2 Evaluation of bulk solid feeding performance

The evaluation of bulk solid dosing processes is of high interest to ensure the uniformity of the continuous manufacturing processes right from the initial steps. In the literature, various approaches are found for the evaluation and handling of data from bulk solid dosing processes, as well as for experimental setups to measure these.

One of the most important guidelines for device manufacturers of dosing units regarding the measurement of fluctuations in bulk solid dosing is given by the "Normenarbeitsgemeinschaft für Meß- und Regeltechnik in der chemischen Industrie" NAMUR NA 40 [41]. This guideline briefly describes simple methods for measuring the dosing quality of the devices. The dispensed material is collected and weighed. From samples of a specific period of time, parameters such as the mean deviation from the setpoint or the standard deviation of the data are calculated.

References in the scientific literature describe setups in which an external catch scale is positioned under the outlet of the bulk solid feeder [23,29,39,42,43]. It is then read using a computer at a specific measurement frequency. This is justified with the already discussed significant noise component and the internal filter functions of the bulk solid doser's scale. The use of an external scale ideally allows minimizing the noise, as the vibrations of the device are not detected. Furthermore, there is an expectation that the data read from the external catch scale are processed with less aggressive black-box filtering functions compared to internal data of dosing devices. However, this approach has two distinct disadvantages for application in continuous processes. Firstly, an application in the ongoing operation of a continuous plant is impossible since the material is collected and therefore cannot proceed to the next processing step. Secondly, there is the risk of measuring artifacts, which may be caused by the flow of the bulk solid on the external scale.

For the evaluation of the collected data, many different approaches are described, some of which go far beyond the previously mentioned NAMUR NA 40. A comprehensive work by Engisch and Muzzio [29] details the properties of bulk solid feeding devices. The study give information about the properties of the collected data as well as the evaluation through the consideration of the relative standard deviation. Other investigations by the authors delved into detailed examination of the data [43]. They were able to show that the Fourier spectra of the dosing processes could be correlated with the properties of the materials. Also, in this study an external catch scale was used to measure the dosing process. An example on the evaluation of inline data of dosing devices is given by Johnson et al. [44]. In this study the authors propose a stochastic autoregressive-moving-average model to simulate the systematic deviations in the powder dosing process. With its high applicability in industrial settings this model is one of the most sophisticated approaches to qualify continuous dosing devices.

The application of PAT in bulk solid dosing processes rarely is described in literature. This is most likely due to the highly chaotic processes that occur during the dosing of powders, which complicate the quantification of the dispensed material using spectroscopic methods. An example is the use of videometric based mass flow control on microdosing processes by Madarász et al. [45] or the implementation of microwave sensors in large scale powder dosing [23]. While ultrasonic-based measurement methods exist for assessing the flow velocity of particles in suspensions [46], there is limited research available on similar measurement setups for falling powders [47].

While investigations into continuous dosing processes often utilize short process times, the runtime on an industrial scale is often much longer. Studies have shown that disturbing effects may accumulate over extended runtimes [48]. Depending on the properties of the materials, this may even lead to a required process stop. Emerging errors may include the clogging of dosing screws or rat-holding effects. Not only can material properties influence the dosing process, but it may also occur that the materials undergo changes in their properties as a result of the dosing process [49].

To assess the impact of dosing fluctuations on subsequent processes, it is crucial to consider the properties of these processes. Meier et al. [23] presented an approach for evaluating dosing fluctuations

in the context of continuous twin-screw granulation. They utilized data from an external catch scale and calculated the integral of dosing deviations from the setpoint over time. After dividing the absolute values by the setpoint, they further divided by the duration of the intervals. They used this relative weighted standard deviation over the time intervals for qualifying the fluctuations. However, it should be noted that the temporal dependency of fluctuations is lost within this approach. Moreover, the results are highly dependent on the sampling frequency. In the analysis of continuous blenders, more sophisticated approaches based for example on the convolution of dosing data are found [50]. The distribution of particle residence times in the mixer serves as the filter function, enabeling accurate predictions of the impact of powder dosing fluctuations. One issue is that analyses often rely on filtered internal data or offline data from an external scale. Additionally, the residence time of the material in the ongoing process must be known.

1.3 Manufacturing of tablets

1.3.1 Introduction into the tableting process

Tablets are defined by the European Pharmacopoeia as oral solid dosage forms, manufactured through an appropriate method [51]. The compression of powders and granules is the most common manufacturing method for tablets, even tablet manufacturing by other methods like lyophilization [52] or 3D printing [53] is possible. The process of tableting, in the context of this work, refers to the entirety of steps required to manufacture tablets from bulk solids via compression. The excipients used can differ significantly among the various manufacturing methods.

Materials used for the production of tablets through compression can have effects on both the process and the final product. Excipients, especially intended to simplify and/or enable process management, include lubricants and flow regulators. Lubricants reduce frictional forces during material's compression and tablet's ejection from the machine [54]. Lubricants can either be used within the mixture of other components or be applied to the punches tools using a suitable dispersion system. In this context, the terms "internal" and "external" lubrication are applied to distinguish between the two methods. Flow regulators can be used to ensure flowability by reducing particle-particle binding forces in the material to be compressed, which is crucial for the tableting process.

Excipients which are intended to influence the properties of the produced tablets are much more diverse [55]. Fillers are added to tablet formulations to increase the volume of the tablets, especially when the dosage of the API is low. Additional substances can be used to control the disintegration of tablets in aqueous media, influence drug release, mask the taste of bitter substances, or ensure the strength of the produced tablets. Materials that have a positive effect on the strength of tablets even at a low concentration in the mixture are referred to as binders. Binders contribute significantly to the shape stability of the tablet by forming strong bonds within the tablet during the compression process.

The materials are introduced into the tablet press as a mixture. This mixture can be a purely physical combination of the components. This procedure is referred to as direct compression. In contrast, individual materials or their mixtures can be granulated before the compression step. One of the advantages of direct compression is the evidently lower production costs [56]. However, undesirable characteristics of the starting materials, such as poor flowability or weak binding strength, may necessitate the use of granulation steps. Additionally, improvements in content uniformity can be achieved with a granulation pre-processing step, as well as a reduction in dust exposure for staff.

1.3.2 The compression of bulk solids into tablets

As previously mentioned, several steps can be distinguished in the tableting process. Figure 3 depicts exemplarily these steps on the example of data measured on a tablet press. During the filling phase, the material is transferred from a feed shoe into the die (b). This occurs at a very low position of the lower punch, ensuring an overfill of the die. Subsequently, the lower punch moves to the dosing height (c), with excess material being scraped off. After this, the lower punch typically descends again, marking the beginning of the movement of the upper punch. The compression phase starts when the upper punch first contacts the solid material (d). During this phase, the material is compacted, generating binding forces that ensure the tablet's physical integrity. After the upper punch reaches its lowest position, the decompression phase begins (e). The material expands due to elastic recovery, delaying the loss of contact between the upper punch and the tablet. Finally, by moving the lower punch upwards, the tablet is ejected from the die (f). When using external lubrication devices, an additional phase occurs just before the filling phase (a). During this lubrication phase, the tools are sprayed with the powdered lubricant.

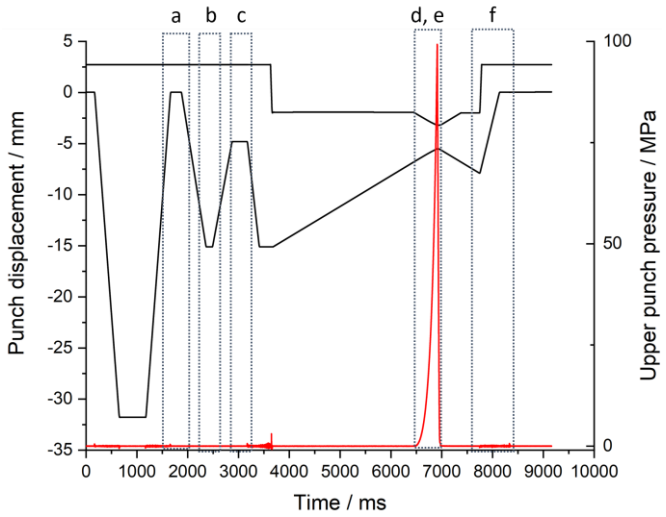

Figure 3: Example of data of a tableting process, black: punch displacement, red: compression pressure measured at the upper punch, a: external lubrication, b: die filling, c: dosing, d: compression, e: decompression, f: ejection of the tablet

During the compression phase, the material undergoes various states [57]. The compression pressure can be understood as the resilience of the particles to the reduction of the volume. Initially, there are rearrangement processes of individual particles as the volume decreases [58]. However, as the compression process continues, the particles start to deform. The deformation mechanism significantly depend on the properties of the compressed materials and their interactions as well as on the process parameters [59]. Several deformation mechanisms can coexist during this phase. It is essential to distinguish between elastic and plastic deformation processes. Elastic deformation does not contribute to the strength of the resulting tablets and leads to an increase in the tablets' volume during and after the relaxation phase. In contrast, plastic processes refer to the irreversible deformation of the material, where the applied plastically energy remains in the tablet, enhancing its strength. Additionally, various types of particle fractures occur. Particle fragmentation increases the surface area and thus the contact area between the particles. The adhesive forces created during compression, arising from, mainly van der Waals interactions, mechanical interlocking and solid bridges [60], account for the physical integrity of the tablets.

Different types of equipment are available for tablet production, with rotary tablet presses being the most important ones for industrial production. They are capable of producing high quantities of tablets per time period. These machines operate on the revolver principle, where the mentioned processes, of die filling, dosing, compression and tablet ejection are separated spatially rather than contemporaneous. The station changes and the movements of the punches are caused by the rotating circular die table. In comparison, compaction simulators are mainly used in formulation development and scientific research. With these devices, there is relative free control over the movement of the punches. This is made possible with the use of hydraulic systems [61] or planetary roller screws [62]. Due to the given controllability of the movements of the punches, compaction simulators are suitable for simulating the punch movements of other tablet presses. This allows for the development of drug formulations and processes in small batch sizes and a subsequent transfer and scale-up to an industrial tablet press [61,63]. Compaction simulators can also be used to perform compressions with simple compression profiles, which can facilitate the study of materials and processes. In rotary tablet presses, the displacement profile of the punches during the compression phase is determined by the diameter of the compression roller and the geometry of the punch head [64]. The resulting profile is characterized by a linear acceleration of the punches [65]. Compaction simulators can operate at constant punch velocities. In the following this is referred to as the V-shape compression profile. However, the use of compaction simulators has shown that they are susceptible for providing inaccurate compression profiles at high tableting speeds [65]. Modern tablet machines are often equipped with PAT that allows

for information about the compression pressure and the position of the punches over the process time [60]. The measurement of the compression pressure is possible through piezoelectric elements or by using strain gauges. While the displacement profile of the punches in rotary presses is usually simulated in-silico, the measurement of the punch movement is possible in compaction simulators [66].

The settings and properties of the tablet machine, as well as the characteristics of the material being compressed, determine the CQA of the produced tablets. Among the CQA is the resistance of the tablet to breakage under diametral loading, which is expressed as the tensile strength. Since one of the focuses of the present work in the area of compression analysis lies in the examination of material behavior under compression pressure and the strength of the resulting tablets, the tensile strength will be explained in more detail in the following. The influences of process parameters and material properties on the tensile strength will be discussed. Additionally, the standard measurement procedure of the tensile strength of flat-faced tablets will be described. Due to the high interest in pharmaceutical research in predicting the tensile strength of tablets, a treatise on such models will be provided.

1.3.3 The tensile strength of tablets

As previously outlined, the tensile strength is of the CQAs of pharmaceutical tablets. It does not only correlate with the tablet disintegration and the release of the API but also serves as a vital parameter for the resistance of the tablet to stresses during downstream processing or handling by the patient.

For flat-faced tablets, the calculation of tensile strength TS has been described by Fell and Newton [67]. They utilized a testing procedure in which increasing lateral force is applied diametral to the tablet. Subsequently, the maximum force before the tablet breaks F_b is divided by a factor for the geometry of the tablet, which derives from the bridge height h_T and diameter d_T (Equation I), based on the distribution of stress in solid elastic bodies [68].

$$TS = \frac{2F_b}{\pi d_T h_T} \qquad I$$

This simple equation is notable not only for its empirical applicability in many studies [69–73] but also in simulation calculations [74]. More complex is the determination of the tensile strength of curved-face tablets [75] and tablets of other shapes [76]. Therefore, flat-faced tablets are commonly studied in research examining the influence of material properties and process parameters on tensile strength.

Semi-automatic tablet testers are available which do not only measure the breaking force but also the mass and dimensions of the tablets. Due to the destructive nature of tensile strength testing, in-line measurement is impossible for practical and economic reasons. Therefore, various predictive models have been extensively studied. Different prediction methods pursue various approaches. Numerous studies can be found predicting the tensile strength of tablets made from mixtures of different substances. [77–81]. These models are based on data of the tensile strength of tablets made from the respective pure materials and their mixtures. To reduce system complexity, experiments usually focus on two or three pure components within the mixtures. Grid-based experiments are conducted to investigate changes in tensile strength as the composition of the mixture varies. Subsequently, linear or geometric models [82] are typically used to describe the results. The investigations reveal that interactions between the different components can lead to significant deviations from the underlying models. Thus, these models have limited applicability and are only considerable for the studied conditions. However, for the description of the experimental spaces, they provide sufficient information. Beside this, the application of empirical equations for the prediction of the tensile strength is published [83,84]. The latter approaches primarily focus on the impact of the proportion of the lubricant in the power mixture on the tensile strength of the tablets.

Other models attempt to predict the tensile strength of tablets based on the cross-decomposition of material properties [85,86]. Recently, in this area, various machine learning models have been introduced [87–89]. Within all of this approaches, some also take into account the parameters of the tableting process [90,91]. Basically, the recently presented models are all connected by the requirement of extensive investigations of material properties or measurements of the behavior of the pure substances. Among all these published models, none is known that only utilizes the process parameters of the tablet press to predict the properties of the resulting tablets.

Apart from these approaches, implementations of additional PAT methods can be found. This includes the study of ultrasound probe implementation [92], as well as spectrometric methods [93]. Non-destructive offline measurements were proposed by Halenius et al. and Juban et al. using microscopic and nanoindentation procedures [94,95].

The prediction of the tensile strength of tablets is supported by the distinct dependence of the tensile strength on the maximum pressure during the compression phase. This relationship between compression pressure and tensile strength within one material is referred to as the tabletability. Together with the parameters of compressibility and compactability, tabletability describes the three-dimensional relationship among tensile strength, solid fraction, and compression pressure [96]. Due to the immense significance of these parameters on compression analysis, they will be explained in more detail in the following sections.

1.3.4 Tabletability, compressibility, compactability

The three terms, tabletability, compactability, and compressibility, denote two-dimensional correlations of the three parameters: compression pressure, solid fraction, and tensile strength (Figure 4).

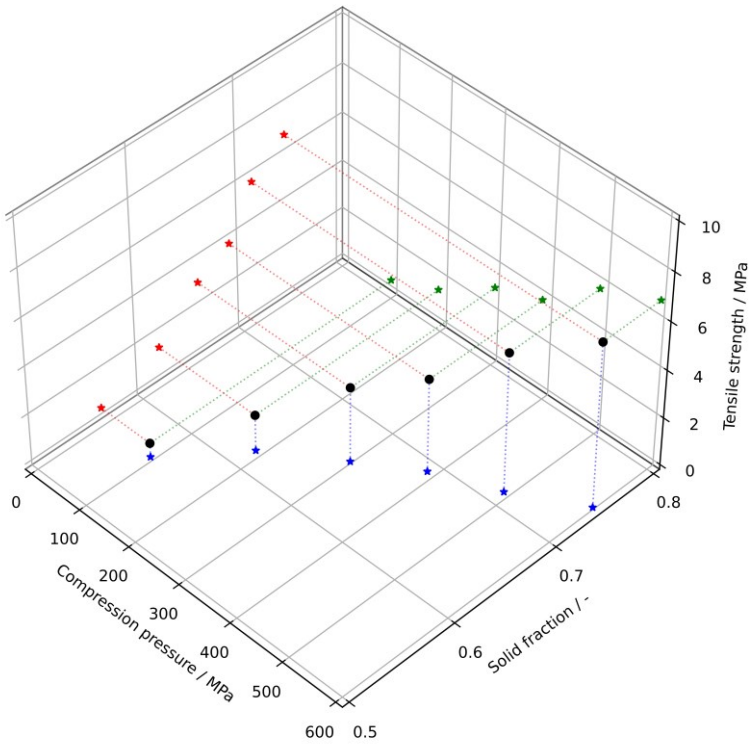

Figure 4: Blue: Compressibility, red: compactability and green: tabletability on the example of dibasic anhydrous calcium phosphate (DI-CAFOS® A 150), black points in 3D space

Every material or mixture of materials can be characterized within this parameter space. The investigation of these three parameters is one major focus in research on the behavior of pharmaceutical powders and granules within the tableting process. Compressibility refers to the plot of the solid fraction against the maximum compression pressure, and it is likely the most frequently described parameter. Various models exist for the analysis of compressibility, for example including equations by Heckel [97], Kawakita [98], Johanson [99], Kuentz and Leuenberger [100], and Sun [101]. The different equations utilize different portions of the collected data. While the models of Kuentz-Leuenberger and Sun are regressed over the full range of the data of the compression phase, models like Heckel or Walker are only applicable in a small linear portion of the transformed data. The procedure for the identification of the portions within the data is frequently not described in the literature. This seems to be a major drawback in terms of the reproducibility of the studies. Compressibility can be investigated using both in-die and out-of-die procedures. The term "in-die" refers to the utilization of PAT to monitor the solid fraction and compression pressure throughout the entire compression process of a single tablet. Controversially, "out-of-die" methods involve the production of different tablets at varying compression pressures. While compressibility can provide information about the deformation behavior of the material, it is less relevant in the consideration of the properties of the produced tablet. On the other hand, compactibility which correlates the solid fraction with the tensile strength, and tabletability, which involves tensile strength and compression pressure, are of higher interest in the practice of formulation development and production. However, fewer models are available to describe the course of these dependencies. One example is the equation according to Ryshkewitch [102] which is applicable to describe the tensile strength of tablets in correspondence to their porosity. Frequently, linear portions of tabletability profile are regressed, and the resulting slope is evaluated. A new approach combines the equations according to Kuentz-Leunberger and Ryshkewitch. The authors were able to demonstrate the applicability of the model for various materials in a range between approximately 10 and 500 MPa [103]. The evaluation of tabletability and compactability is only possible using out-of-die methods as the tablets must be destroyed to measure the tensile strength. Therefore, the shape of the compression profile, including both the solid fraction and the compression pressure over time, is hardly implementable in these evaluations. This contrasts with the understanding that the punch displacement profile during the compression phase plays a decisive role in the tensile strength of the resulting tablets [71]. One example is the influence of dwell time [104,105], which refers to the duration the punches remains at the lowest point in the die. The dwell time is an important feature of rotary tablet presses, resulting from the movement of the flat punch head on the compression roller [64]. The profiles of tabletability and compressibility consequently differ depending on the type and dimensions of the tablet machine used as the movement of the punch’s changes with the size of the compression rollers and the geometry of the punch heads [106].

Although the pressure exerted on the powder during tableting could theoretically be increased indefinitely, the tabletability profile of various substances suggests that beyond a certain compression pressure threshold, tensile strength does not improve. This observation is referenced within the literature as overcompression. After a threshold in the compression pressure, the tabletability profile of substances show a plateau or even a loss in tensile strength. With the examples of sodium chloride, sucrose and polyethylene this behavior was first demonstrated by Adolfson and Nystrom [107]. They considered the increase of elastic deformation at high pressure as the reason for the occurrence of overcompression. For some materials they describe that at pressures of around 800 up to 1200 MPa capping of the tablets was visible. The lamination and capping of tablets were found to be a common phenomenon of overcompression [108]. Besides the compression pressure, the material properties [109] and the speed of the compression [110] determine hereby the resulting tablet defects. In the investigation of tabletability profiles and dependencies of tensile strength with process data, the phenomenon of overcompression must therefore be considered. This implies that either tablet defects are incorporated into the computational models for more accurate predictions, or that the range of overcompression is deliberately omitted from the analytical scope.

Like bulk solid dosing data, compression profiles possess the characteristics of time series. However, both types of data also exhibit significant differences between them. Therefore, these two data series are excellent for exemplifying the advanced analysis procedures under considering the properties of these two distinct types of data. Following the presentation of the examined processes, an overview about the utilization of data series will be provided in the next chapter. Subsequently, methods that are relevant in the analysis of databases in the context of this current work will be introduced.

1.4 The analysis of data and data series

1.4.1 Properties of data series

When examining data, not only the individual values can provide information, but also the sequence of the individual values. When this is the case, we refer to them as data series. The properties of data series can significantly differ from the properties of unordered data in several aspects. The higher the informational content within the sequence of data, the more information is lost if this order is not taken into account [10]. For this reason, permutation can serve as an important tool in the evaluation of data series [111–113]. For data evaluation, merely examining distribution parameters might be sufficient for unordered data. For instance, location and dispersion parameters, like the arithmetic mean and standard deviation, are widespread in the evaluation of scientific data. An example is the study of the API content of single-dose drug dosage formulations following the chapter "Uniformity of dosage units" 2.9.40 of the European Pharmacopoeia [51]. This evaluation is adequate when making a statement about a batch of tablets. However, if one aims to determine whether the API content fluctuates too widely or follows a trend over an ongoing process, more sophisticated evaluation procedures should be applied.

A specific type of data series is the time series, where the sequence of data is sampled in chronological order. Thus, all data collected in-line or on-line using PAT can be understood as time series. Given the increasing significance of PAT, especially in the context of continuous manufacturing, analyzing data series has become of highest importance for pharmaceutical research.

Data series can be classified into many categories. For this work, the distinction between stationary and non-stationary series is especially significant. Within a stationary data series, the statistical measures of this data series are nearly constant. These data series thus have no change point. However, individual data points can fluctuate around a global mean. The dispersion of values can be either chaotic or in a certain pattern. A sinusoidal curve can therefore be considered as a stationary time series. Whether a time series is categorized as stationary or non-stationary depends on the observed data interval and the defined point of interest. This is evident from the example of different fraction of the sine function depicted in Figure 5. The subplot (a) depicts the sine function of $y = sin(x + 1.5\pi)$ over the range of $20\ \pi$ so that the character of the curve can be described as stationary. The narrower perspective (Figure 5 b) in the range 0 to π results in a non-stationary behavior of the data series.

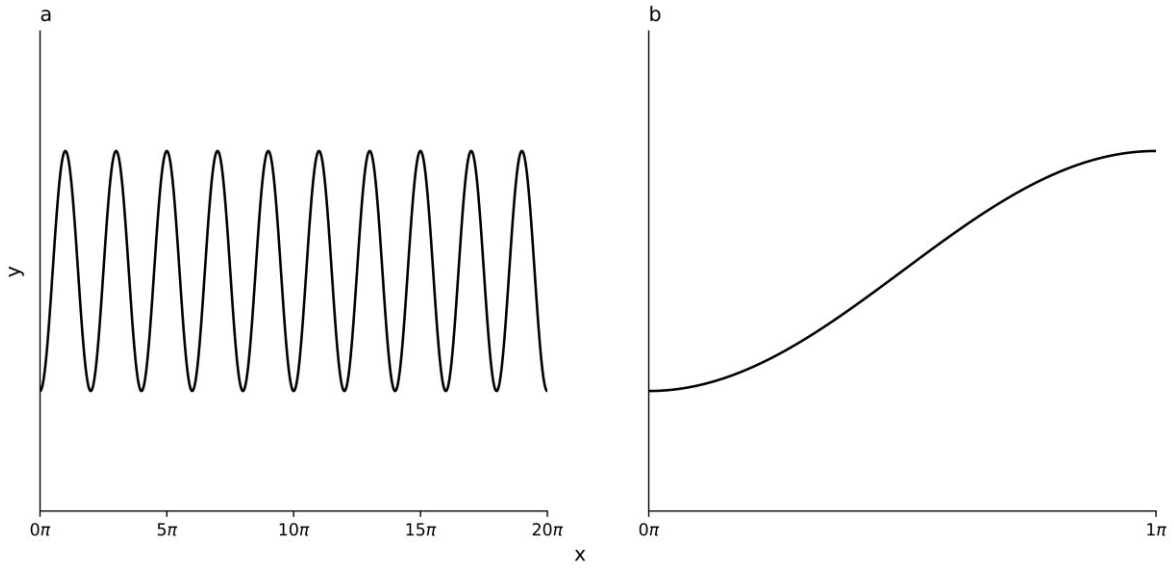

Figure 5: Different aspects of the sine function with $y = sin(x + 1.5\pi)$, a: stationary perspective and b: non-stationary perspective

In the focus of this work, the data of an ideal bulk solid dosing process can be given as an example for a stationary time series. Ideally, the dosing rate of the powder feeder fluctuates around a mean value, which in the best case is as close as possible to the target value. The fluctuations around this mean value can be considered almost constant over the process time. In contrast a single compression profile of tableting processes possesses a non-stationary characteristic. During the compression process of the material, the pressure increases, while simultaneously the distance between the punches decreases until the minimum of the distance is reached. Therefore, for these data, the location and dispersion parameters do not remain constant.

In the analysis of data series, the investigation of trends, change points, and correlations are of particular importance. Moreover, every real measurement carries an inherent error. Disturbances might also arise, appearing as noise in the data. The separation between measurement error and "real" curse of the data series is one field of focus within the analysis of time series. For time series, uniform intervals between values are of high relevance. Some evaluation methods are feasible only when there are consistent intervals between data points. One example here is given by Fourier analysis. Therefore, previous interpolation of the data might be required for their assessment. Given the significance of data evaluation methodology in this work, a subsequent introduction into the applied methods will be provided.

1.4.2 Fourier Analysis

The Fourier analysis allows for the description of discrete values and continuous functions through their simplification [114]. The Fourier analysis is used in many fields, such as in spectral analysis or in signal theory [28,115–118]. An introduction to the concept of Fourier analysis will be provided in the following. Every continuous function can be understood as the sum of a trigonometric function. The continuous Fourier transformation describes the calculation of this trigonometric function from the initial function $f(x)$ (Equation II) [119].

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-i2\pi\xi x} dx \qquad II$$

The "intermediate" result of this integral transformation is the complex function $\hat{f}(\xi)$ with ξ as the frequency of the trigonometric function. Since real data do not correspond to continuous functions but are discrete, this analytical mathematical function of Equation II is not applicable [120]. For the calculation with real data, the discrete Fourier transformation have to be applied (Equation III) with k and n as the indexes of the matrices $\hat{F}$ and F.

$$\hat{F}_k = \sum_{n=0}^{N-1} F_n e^{-\frac{i2\pi}{N} kn} \qquad III$$

In this case, F and $\hat{F}$ correspond to one-dimensional matrices of length N. $\hat{F}$ contains the result of the Fourier transformation as complex numbers. Each row of $\hat{F}$ corresponds to a sine function with the amplitude A and frequency ξ. The amplitude is calculated according to the Pythagorean theorem as the sum of the magnitudes of the real part a and imaginary part b (Equation IV).

$$A = \frac{1}{N} |a + ib| = \frac{1}{N} \sqrt{a^2 + b^2} \qquad IV$$

The frequencies ξ of the trigonometric function are determined by the sampling interval f_s of the input data F (Equation V).

$$\xi = \{0, \frac{f_s}{N}, \frac{f_s}{N-1} \dots, \frac{f_s}{2}\}$$

V

Using the amplitude and the frequency, the result of the transformation can be represented in a spectrum. This is then referred to as the frequency domain of the data. In contrast, the original data are located in the time domain [10]. The Fourier analysis allows for the identification and examination of regularly recurring signals in the frequency domain of the data. Fourier transformation is primarily used in the analysis of spectral data [121], but there have also been applications in the evaluation of data from powder dosing processes [43]. Figure 6 gives an example of the time domain (a) and frequency domain (b) of the Fourier transformation of an ideal sinusoidal of the frequency of 0.019 Hz. The subplots c and d depict mean centered and scaled data of a dosing process of Flowlac® 100. The main sinusoidal of the process is of a frequency of 0.019 Hz. As the sample frequency of both data series is 1 Hz, the upper limit of the calculated frequencies is 0.5 Hz according to Nyquist theorem [122].

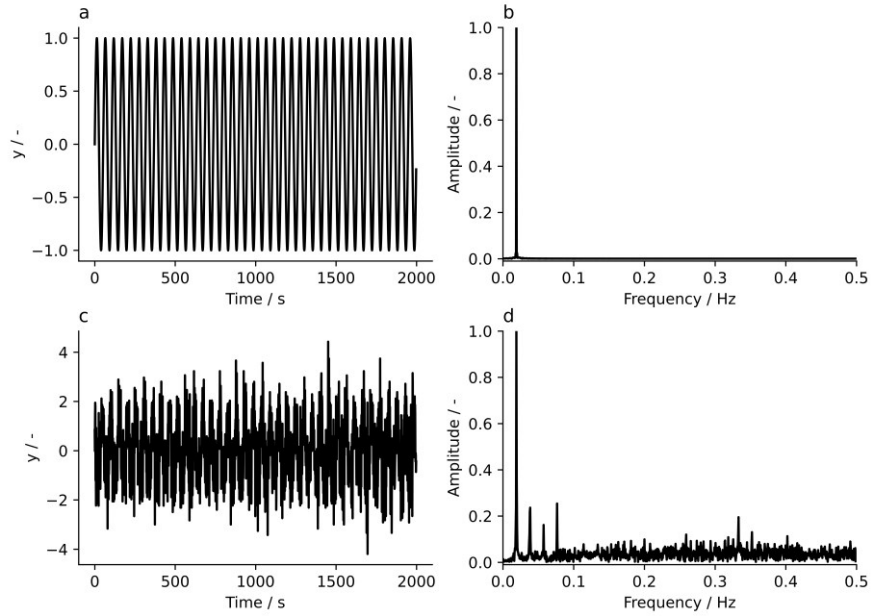

Figure 6: Time (a and c) and frequency (b and d) domain of an ideal sinusoidal and measured data of the frequency of 0.019 Hz.

1.4.3 Change points

In statistics, change points are defined as the change of the distribution of data points in the progress of data series. The detection of change points is a subfield in the evaluation of data series [123]. The assessment of change points in the context of pharmaceutical manufacturing heavily depends on the specific research question at hand. In the evaluation of ideally stationary time series, the detection of change points hints changes in the underlying process. In these contexts, change points can signify undesirable events that indicate errors in the process. When considering powder dosing while accounting for refilling phases, these can also be understood as change points [39]. On the example of the refilling of dosing units it becomes evident that the precise definition and assessment of a change point is depending on the process and on the conditions and objectives of the specific experiments.

For compression profiles, reconsidering Figure 3 the occurrence of change points in the data series is evident. For the computational separation of the different events like die filling, dosing, compression and tablet ejection change points have to be identified. The precision of the separation is of highest importance as small variations in the detection of the compression phase can result in significant changes of the detected pressure in the begin of the compression process.

1.4.4 Multivariate decomposition

The goal of multivariate decomposition is to reduce the dimensionality of a dataset and separate the underlying information in as few dimensions as possible. In this context, the term "information" is synonymous with the spread or dispersion of the data. Multivariate decomposition can be understood as the shadow cast by a hyperdimensional object, such that the shadow contains information about the actual object.

In statistics, there are various methods that allow for the decomposition of input data. One widely used methodology is the Principal Component Analysis (PCA) [124]. This technique aims to reduce the dimensions of the input data by calculating new orthogonal dimensions known as principal components [125]. The principal components in PCA are vectors in the original coordinate system. The first principal component contains the highest proportion of information in the data. Each additional component contains progressively less of the total information. The first principal component is calculated by fitting a regression line that minimizes the sum of the magnitudes of the Euclidean distances of the data to this line (Figure 7). Each subsequent principal component is defined as to be orthogonal to the previous ones and is also the result of the similar regression procedure. The vertical projections of the data onto these components serves as coordinates in the new orthogonal space. These projections are referred to as the scores of the PCA. The slopes of the regressed principal components are called the loadings.

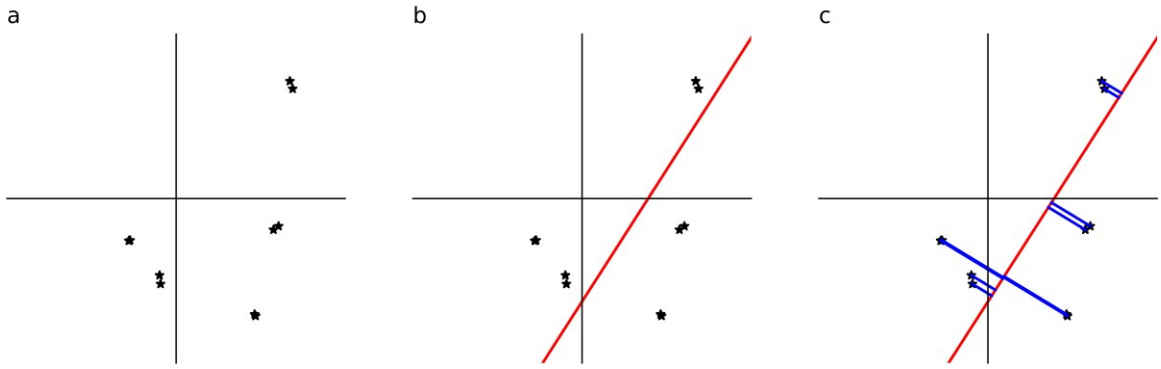

Figure 7: PCA regression procedure from a to c, black: data points, red: principal component, blue: projection lines of the data points

One of the commonly used calculation procedures of PCA utilizes the covariance matrix cov of the standard normal variate scaled input data matrix X of n dimensions according to Equation VI .

$$cov = \frac{1}{n-1} XX^T \qquad VI$$

The following step is an Eigen analysis, resulting in the diagonal matrix of the eigenvalues Λ and the quadratic matrix of the eigenvectors U of cov (Equation VII).

$$cov = U\Lambda U^T$$

VII

The scores of the principal components T are subsequent calculated by the product of U and X (Equation VIII) [126].

$$T = XU$$

VIII

Examining the scores of the first two principal components is typically sufficient to make a valid statement about the correlation within the data. The influence and correlation of every input dimension can be evaluated in the investigation of the loadings. PCA therefore is particularly well-suited for obtaining an overview of large databases and visually representing the underlying connections and differences between various datasets.

Although considering missing values is theoretically possible and often highlighted as an advantage of PCA, in practice, it comes with significant disadvantages. This can be illustrated using what is presumably the most modern and simplest implementation of this method. Podani et al. [127] present a procedure that utilizes a binary indicator matrix. The indicator matrix contains integer values of zero and one, where zero indicates missing values and vice versa. This indicator matrix is multiplied by the covariance matrix prior the Eigen analysis in the PCA regression calculation. This allows the necessary computational operations to be performed without all values being present in every dimension. However, they also describe that comparing values across different dimensions can introduce artifacts in data evaluation. For this reason, the omission of variables or samples is usually less problematic when values are missing in the input data frame.

Beside multivariate decomposition procedures for the handling of multidimensional problems in pharmaceutical data, machine learning seems to provide the most promising methodology. The almost limitless possibilities in the set-up of machine learning models provide powerful tools in the evaluation of the data. The following section gives an introduction to machine learning and its application in pharmaceutical science.

1.4.5 Machine learning

1.4.5.1 A short introduction to machine learning

Machine learning aims to enable computers to explore and solve problems with a high degree of autonomy. Applied are algorithms that can independently handle new inputs without the exact approach being predefined in advance. The algorithm assesses its own performance and, if necessary, modifies its behavior accordingly [128]. Commonly cited examples of machine learning include robotics [129] and autonomous driving [130]. In the context of pharmaceutics, machine learning can be applicated in data analysis [131], pharmaceutical formulation development [132] or process control [133]. Machine learning can be used here to recognize clusters in data, generate regression models, and even fully control production processes autonomously.

The use of machine learning has seen significant growth in the recent past. Notable examples of this are disruptive innovation in the field of large language models. The architecture of large language models can be thereby adapted to other fields of interest, such as for clinical tasks [134] or research of protein structures [135]. In various areas of research, the use of machine learning is widespread. Examples range from physics [136], over drug discovery [137] to evolutionary biology [138]. In the realm of pharmaceutical research, there are widespread applications of machine learning, such as in in-silico drug synthesis or the design of clinical trials. In the academic research related to drug

manufacturing, studies exist that utilize machine learning in grey-box or black-box models for process control [139]. However, there are few applications of machine learning for data analysis in the context of pharmaceutical manufacturing [87,140,141]. This might be partly due to the black-box nature of machine learning models. Often, empirical, semi-empirical or physics-based models are preferred. However, these models run the risk of oversimplifying complex relationships and risk the loss of information. Additionally, the size of pharmaceutical experimental databases is typically quite small. This limits the potential of machine learning, as there are often insufficient data available for training the model.

Training machine learning models is of paramount importance for the quality of the model. This does not fundamentally differ from traditional statistical models, but due to the size and complexity of the datasets, it is more challenging to train a machine learning model. The goal of training is to generate a model using training data, which can then be tested using test data. The better the model performs on the test data the higher is the predictive power and the value of the model.

There are various training approaches to differentiate [142]. These are characterized by different principles and methodologies during the training phase. In supervised learning, the model is fed by both input data and output data. The specific task can be either classification or regression. The goal of the training is to generate a model function that correlates the input variables with the output variables. In unsupervised learning only the input data is used in the training of the model. The aim of this method is to reveal previously unknown data clusters or other structures in the data and to examine connections between individual datasets. In a more specific sense, such models can also be listed among the decomposition methods previously described. The third category refers to reinforcement learning. In this approach, the model interacts with a reward function, which assesses and thus guides the learning of the model.

The selection of the training method to be used depends on the aim of the researcher and the type of model being employed. Typical models are based on decision trees, random forests, support vector machines, Bayesian kernel machine regression or artificial neural networks. Due to the versatile applicability of certain neural network architectures for time series data analysis, neural networks will be discussed in greater detail in the subsequent chapters.

1.4.5.2 Neural networks

Artificial neural networks (ANNs) are computer algorithms that are rudimentarily based on the structure of the animal brain. They serve as a key element in the field of artificial intelligence and machine learning [143].

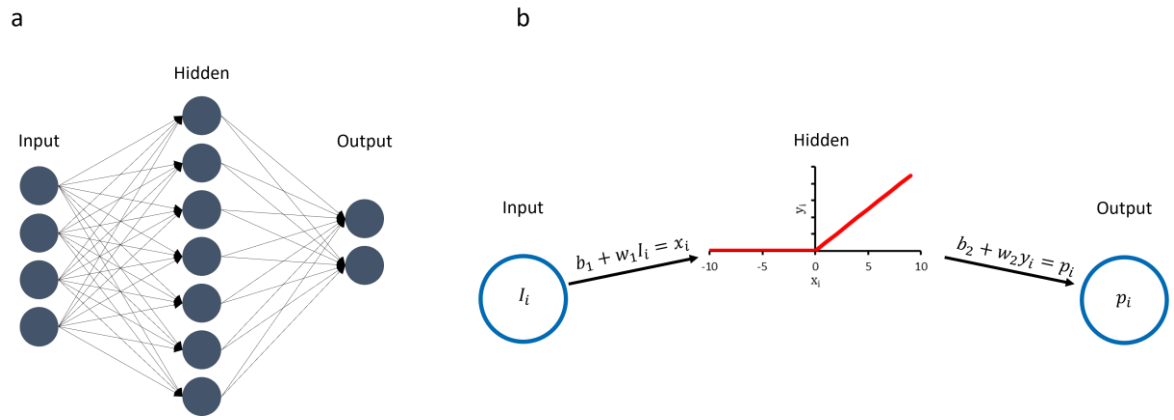

Figure 8 a: The structure of a simple neural network, b: depiction of a basic mathematical function model for the nodes and the connections.

ANNs consist of a series of layers composed of individual nodes (Figure 8 a). These nodes from different layers are interconnected in various ways, creating a network. The connections between individual nodes represent linear functions (Figure 8 b). The parameters of this linear functions are the slopes and the intercept. In the context of ANNs, the slopes are referred to as the weights of the functions and the intercepts are called the biases. In the training phase of the network the weights and biases are tuned to improve the prediction of the ANN. This will be addressed in more detail in a following chapter. The nodes in ANNs also correspond to functions that translate the input of the node into an output. The functions of the nodes are called activation functions. There are different types of layers in an ANN. The input layer accepts the input data, and its size corresponds to the size of the input data matrix. The input data matrix may contain only one value or may be very large or multidimensional. Furthermore, a network may have multiple individual input layers. The predictive power of an ANN arises from the use of non-linear functions within hidden layers [144]. These are attached to the input layer and can exhibit various substructures. The substructure of the hidden layers may specialize in different types of input data. For instance, convolutional layers are commonly applied in the evaluation of image data [145]. An ANN can have several consecutive interconnected hidden layers, all of which can have different substructures. Through the combination of appropriate layers, powerful models emerge for detecting hidden structures and information within the data [146]. For the application of ANNs with multiple hidden layers the term “deep learning” is often used. The final layer is the output layer. This layer produces the result, such as a prediction of a value or the probabilities for classification into various categories. The output of an ANN, in size and dimension, is theoretically unlimited.

ANNs are classified into overlapping categories based on their structures. These categories include convolutional networks [147], which are optimized for the analysis of grid data such as images, and generative adversarial networks [148] where two networks compete, leading to mutual improvements. Contrasting classes are formed by feedforward ANNs and recurrent ANNs. In feedforward ANNs, information flows unidirectionally from the input layer through the hidden layers to the output layer, without any feedback mechanism. In contrast, the information flow in recurrent ANNs occurs cyclically [149]. This enables the consideration of relationships between different data points within the input, making them suitable for the analysis of sequential data. They are utilized, for example, in predicting the trends of sequential data and their classification [150].

As of the author's best knowledge to this point, recurrent models have not been applied in the context of evaluating data from tablet manufacturing processes. Since recurrent ANNs potentially offer benefits in the analysis of pharmaceutical sequential data, their structure and properties will be explained in greater detail in a later chapter.

1.4.5.3 Application of machine learning in pharmaceutical research

Powered by the increase in the computational capacity over the recent decades, the application of machine learning in various fields of science is a growing tool in data handling and evaluation. In the scope of pharmaceutical science, the use of machine learning is widespread in drug discovery, pharmacodynamics and pharmacokinetics [137,151–153] (Figure 9 a and b). In contrast, in pharmaceutical development and manufacturing the application of machine learning tools is just starting to grow and will progress in the upcoming years.

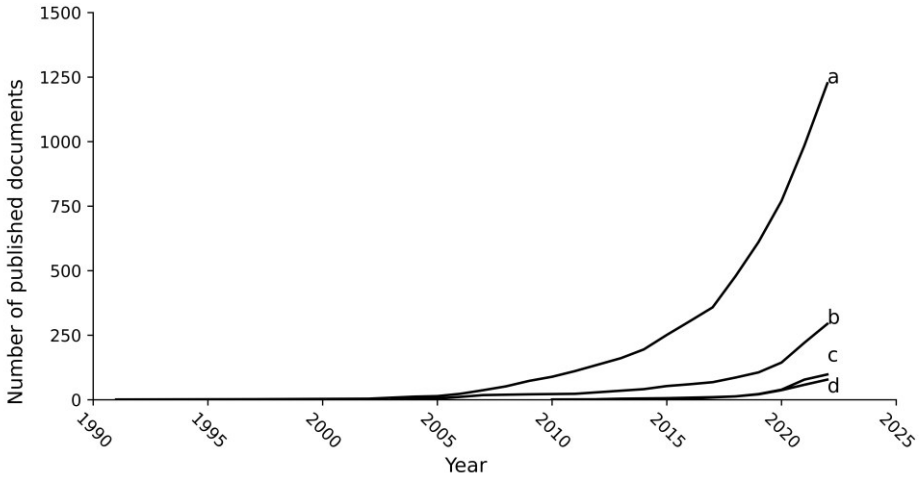

Figure 9: Number of published journal papers per year listed in the literature database Scopus® by Elsevier publisher limited to the field of pharmacology toxicology and pharmaceutics, searched for a: "machine learning drug discovery", b: "machine learning (pharmacodynamics OR pharmacokinetics)", c: "machine learning formulation development", d: "machine learning manufacturing" in article title, abstract or keywords

For instance, deep learning was used for the prediction of the in vitro performance of pharmaceutical formulations [154]. In this study the authors utilized parameters of the formulations and molecular descriptors of the APIs for the input in ANNs to predict parameters such as the f_2 value of the dissolution of the investigated formulations. Similar approaches have been provided by Ibric et al. [155] and Han et al. [155,156]. Other applications of machine learning in formulation development and pharmaceutical manufacturing include the field of additive manufacturing technologies [132,157–160], the development of nanoparticles [161,162], for avoiding of animal testing [163], or the investigation of food effects on bioavailability [164].

For the prediction of the CQAs of tablets using machine learning tools various studies can be found [87,156,165,166]. Akseli et al. [166] presented a neural network with the input of various parameters of the tablets. These included the geometry, porosity, weight and results from ultrasound measurements of the tablets. From these parameters they predicted the disintegration times and the breaking forces of the tablets. Djuris et al. [87] presented a comparable study. They utilized neural networks for regression and classification tasks to predict various measures of the tableting process, including the tensile strength of the tablets, the elastic recovery or the ejection force. For the input parameters the compression load, the content of APIs and excipients, and a nominal parameter were used. The nominal parameter was used to distinguish between co-processed excipients and physical mixtures of the components. On a test data set of around 15 data points, representing 15 % of the total data content, they demonstrated after an iterative training process of 20 neural networks high predictability of the optimal networks. Paul et al. [165] investigated the phenomenon of capping of tablets with multivariate decomposition and decision tree-based machine learning. A list of various measures of the material properties and descriptors of the tableting process including the plasticity or the die-wall pressure were utilized as input parameters. The data set of this study consists of 253 different measurements. Han et al. [156] applied two different ANNs for the prediction of the disintegrating time based on molecular descriptors of the API, API content and encoded categorial input for the excipients. In these examples, generating the input dataset involves significant experimental effort. The utilization of pre-existing data from the manufacturing process is not found in machine learning for tablet manufacturing.

The former examples are intended to demonstrate the diversity of application possibilities for machine learning within the field of pharmaceutical science. Understanding the properties of neural networks is crucial for avoiding errors during application, particularly in the training of the models. This will be elaborated upon in the following chapter.

1.4.5.4 Supervised training of artificial neural networks

ANNs have to be trained using sufficiently large databases. As this work focuses on the application of supervised learning, the discussion will be limited to this method. In practice, prior to the training of the model, the collected data is usually grouped into two categories. The training data is used to train the network, while the test data is utilized to evaluate the predictability of the trained model [167]. The split of the data into a training dataset and a test dataset can be done according to various criteria. Often, a random division is made.

Stochastic gradient descent is used for the training of most ANNs [168]. It describes the iterative statistical process of optimizing a loss function. The loss function is a function that evaluates the output of the ANN in relation to the observation. It can be simplistically understood as a grading method for the neural network. A commonly used loss function for regression problems is the sum of squared residuals SSR. It is defined as the sum of the differences between the model's predictions p and the observed values o to the power of two, depending on the weights and biases of the network. Instead of the sum the mean of the squared residuals can be used. The cross-entropy of prediction and observation is a typical loss function for classification tasks in ANNs. The gradient of the loss function is its derivative with respect to the change in a weight or a bias. In the initial step before training the network, the weights are usually set randomly according to the standard normal distribution. The biases are typically set to zero. For each weight and bias, the gradient is calculated. This gradient is then multiplied by a predetermined learning rate and serves as a correction term for the respective weight and bias by being added to them. This process is demonstrated in the following for a simple ANN with one hidden layer utilizing the Rectified Linear Unit (ReLU) as activation function as depicted in Figure 8 B. The gradient of the sum of the squared residuals regarding the second weight of the model w_2 is calculated according to Equation IX with y_i as the result of the ReLU activation function. The observation is given by o_i and the prediction by p_i, where i denotes the index up to the size of the data set n.

$$\frac{d\,SSR}{d\,w_2} = \sum_{i=1}^{n} -2(o_i - p_i)\,y_i \qquad IX$$

With a learning rate α, w_2 is corrected in this learning step according to Equation X.

$$w_{2,trained} = w_2 + \alpha \frac{d\,SSR}{d\,w_2} \qquad X$$

The training of the neural network mostly is conducted in the direction from the output layer towards the input layer. This procedure is called the backpropagation [169]. In this example, w_2 and b_2 are trained before w_1 and b_1. Based on this procedure the term backpropagation is used. This training is conducted in epochs, during which the training data are used multiple times in smaller batches to optimize the weights and biases. The training continues until a certain endpoint is reached. This endpoint can be set to a specific number of epochs or based on a set limit for the loss function. To track the predictive power of the neural network during training, an additional split of the data is often made within each batch. A portion of the training dataset, often around 90 %, is used for training the weights and biases, while the remaining data is exclusively utilized to evaluate the current state of the network. This is performed by applying the loss function to this validation data set. The value of the loss function for the validation data can also be used as an endpoint for the training.

The activation functions used in the network are of particular significance for training. This can be illustrated using the ReLU function, which is used in this example [170]. The ReLU function returns a value of zero for inputs less than zero. This introduces a threshold potential for the node. The output behavior of ReLU leads to a potential problem in training the neural network. If the input to the ReLU

function of a layer is less than zero, then its output will be zero. This output is then multiplied during backpropagation, resulting in a term for the gradient equal zero. Consequently, the corresponding weights and biases do not receive any further training. This problem is referred to as the "dying ReLU" problem, and it can cause nodes of the network to become non-responsive [171]. To address this problem, the leaky ReLU function can be employed [172]. This function has a linear slope α_{leak} for negative inputs, chosen to be slightly different from zero, and is the same as the ReLU function for positive inputs (Figure 10). By modifying the activation function in this way, the failure of the respective nodes can be prevented.

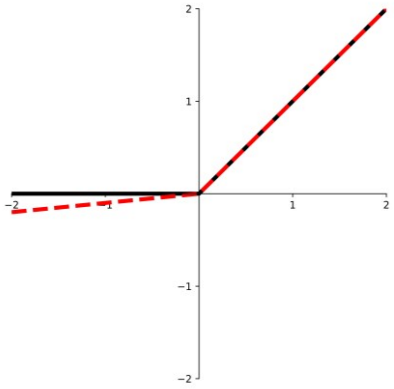

Figure 10: black: Rectified Linear Unit (ReLU) function and red: leaky ReLU with $\alpha_{leak} = 0.1$ for $-2 \leq x \leq 2$

Another method to prevent the failure of nodes is known as dropout [173]. In this technique, specific nodes are randomly omitted during the training phase of the computation. The proportion of these omitted nodes relative to all the nodes in the layer is determined by the dropout rate. As a result, other nodes may receive new types of inputs, enabling them to restore them. Moreover, employing dropout can reduce the computation time of networks, diminish overfitting of training data, and lessen co-adaptations between multiple nodes. The term "co-adaptations" refers to the occurrence of identical outputs from nodes when they interact similarly with the input. This, in turn, reduces the overall complexity and increases the power of prediction of the network.

In the training of neural networks, practitioners often rely on pre-defined algorithms that are optimized for high computational speeds. An example of this is the Adaptive Moment Estimation optimization algorithm [174], which is frequently used in practice [175–177]. The correct selection of algorithms allows the use of optimized hardware for the specific computing systems. The use of graphics processing units is often preferred over central processing units in this context.

The aim of the training of an ANN is to optimize the predictability of the model. Due to the complexity of the underlying architecture, the use of an appropriate test dataset, along with the meticulous examination of predictions through the analysis of residuals, is essential.

1.4.5.5 Recurrent neural networks

Recurrent ANNs, as previously described, are characterized by the fact that the flow of information within the network partially occurs in cycles. This enables the network to store states of various inputs and combine them, allowing recurrent neural networks to specialize in handling sequential data [142]. The networks gain a sort of memory for previously processed input through their structure. In this memory process, the input is a data series of length k. The individual values of the data series are fed sequentially into the recurrent layer. The output of the first input into the recurrent layer is then multiplied by a weight. The result serves as a weighting for the second input of the layer. This is continued in a chain process up to the last element of the data series. During this process, each individual output of each input can also passed on to the subsequent layers. This type of layer can also

be represented in what is known as an "unrolled network". It represents the chronological component of a recurrent layer, separated in different virtual nodes. It must be noted that in reality, only one node exists, and the input are fed into it sequentially. A simple form of an unrolled recurrent network is shown in Figure 11. The first input of the data series I_1 is fed into the layer and results in the output p_1. For the second input, the y_1 value of the first input is multiplied with the weight of the recurrent "memory" w_3 and considered in the calculation for p_2. The similar procedure is performed for p_3 regarding the value y_2.

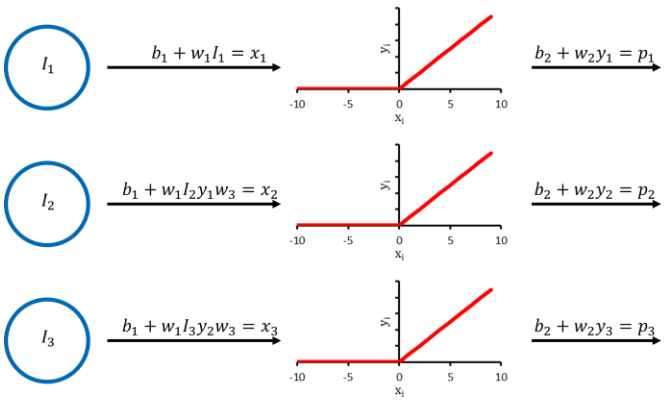

Figure 11: Unrolled recurrent layer with a data series of length 3 as the input and ReLU activation function

Since the calculation of each output of the recurrent layer can be understood as a multiplication of the previous outputs, the phenomenon of 'exploding gradients' or 'vanishing gradients' can easily occur. This situation takes place with longer data series of length k if the value of w_3 is different to one. The weight of w_3 enters then the result of the last input with ${w_3}^k$. During the backpropagation in training, if $w_3 > 1$, a large factor for ${w_3}^k$ will prevent the calculation of small changes in the weights. There will be large changes in the weights and biases, and their optima cannot be found. This is referred to as the 'exploding gradients' phenomenon [178]. A similar situation can occur if w_3 is smaller than one. In this case, for long data series, the rate of change of the weights may approach zero. Again, the optimization of the weights and biases becomes impossible. This latter situation is known as the 'vanishing gradient' problem.

Various approaches have been described to address this issue, and they can be applied either individually or in combination [179]. One solution is to limit the size of the gradients to a specific maximum value. Additionally, the initially chosen values for the weights of the recurrent layers can have a significant influence on the progress during training. However, the highest degree of improvement is achieved through the use of special architectures of the recurrent layers. In this method, the recurrent layer is equipped with gates, which enable the network to 'decide' whether an output should be retained or discarded. Of particular importance in this context of gated neural networks is the long short-term memory (LSTM) architecture [180], which will be explained in detail in the following section.

1.4.5.6 Long short-term memory

Long short-term memory (LSTM) networks are a type of recurrent neural networks that were developed to solve the problem of vanishing and exploding gradients. LSTMs manage this by implementing special structures called gates, which regulate the flow of information, allowing for more stable training and enhanced performance in capturing long-range dependencies within the data.

The LSTM layer combines two different types of structures (Figure 12). One is designed to act as short-term memory, while the other serves the function of long-term memory. The long-term memory is

facilitated by a structure called the 'cell state'. The cell state does not possess its own weights or biases, contributing to the layer's resistance to both vanishing and exploding gradients. The value for the cell state is given in Figure 12 as $m_{L,i}$. The so-called 'hidden state' represents the short-term memory, denoted as $m_{S,i}$ in Figure 12. The cell state and hidden state are interconnected through three gates. The forget gate enables for abolishing parts of the information in the cell state, achieved through the utilization of the sigmoid activation function. The result, which lies between zero and one, is multiplied with the current value in the cell state, potentially reducing it. The input gate, on the other hand, permits the addition of new information to the long-term memory. It utilizes two separate blocks: one employing the tanh activation function and the other a sigmoid activation function. The multiplied result of both blocks is then added to the cell state. The last gate is the output gate, where information from the cell state and hidden state converges. Both sigmoid and tanh functions are used again in this gate. The output of this gate serves as the new value for the short-term memory and/or as the output of the layer. In contrast to the cell state, the individual gates do possess weights and biases that can be trained. These weights and biases are depicted with w and b. The input in the layer (I_i) influences all three gates.

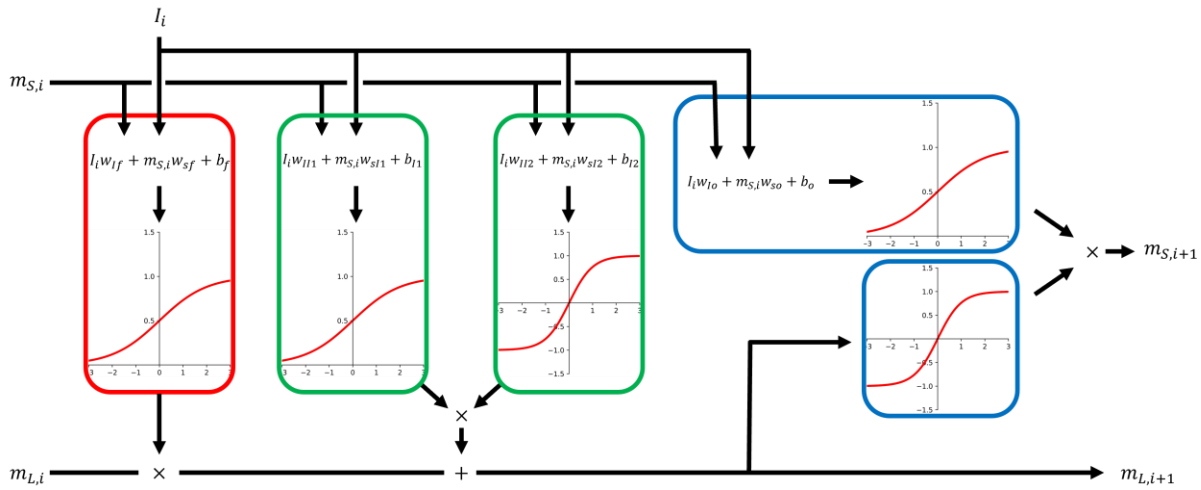

Figure 12: Long short-term memory (LSTM) layer, red box: forget gate, green boxes: input gate, blue boxes: output gate, multiplication operations for results of activation functions are given as ×, additions as +

The utilization of the forget gate and the input gate prevents the amplification of long-term memories within the network by multiplication operations. Instead, the long-term memory is given by an addition of the different values of the cell state. This solves the problem of both vanishing and exploding gradients. This flexibility in the architecture allows for more nuanced control over the information flow within the network, and can be tailored to specific tasks or objectives, enhancing the overall effectiveness of the model.

LSTM layers are frequently used in models that work with sequential data. These applications can include, but are not limited to, series forecasting [181,182], the creation of artificial music [183], the recognition and generation of handwriting and speech [184,185], or pharmacokinetic/pharmacodynamic modeling [186]. Due to the demonstrated exceptional suitability of LSTM-based networks for processing sequential data, they hold promise for a benefit in their application on compression profiles.

1.5 Aims of the thesis

As described in the preceding chapter, the investigation of experimental data from the processing of powders and granules, especially in view of their serial character, promises new perspectives in the description of processes and material parameters. The recent developments in data science and computational capacity provide powerful tools for advanced methods for data analysis. The aim of the present work is to utilize suitable data handling procedures to examine data from bulk solid feeding and compression processes to get novel insights into the 'hidden' information within the data. The developed and applied methods should be suitable for the in-line analytics of the processes, without the need for implementation of additional PAT. Regarding the need to handle ever-increasing amount of data, focus was placed on the development of algorithms and computer-assisted evaluation methods which reduce the individual source of human error.

The following overview presents a sequence of the individual targets of the work:

- Since the examination of fluctuations in pharmaceutical processes is of interest, the focus in the first part of the work was set on the study of these deviations. The first target was to achieve a better understanding of the properties of the stationary data series from bulk solid feeding devices. A secondary aim in this project was the investigation of the suitability of external catch scales as a characterization method for powder dosing units.

- The investigation of compression profiles in pharmaceutical tableting requires appropriate data handling and the establishment of a well-maintained database. Constructing such a database was the primary objective of the second project phase. Of particular importance here was the handling of raw experimental data and the extraction of the compression phases of the tableting processes. Within this context, a program was to be created that automatically applies the pre-processing and standard in-die analysis methods to the data from the compression phases. More in-depth examinations aimed to demonstrate the suitability of the database for subsequent steps. This includes the investigation on the correlation between compression descriptors of starting materials and their blends.

- The final aim of this work was to calculate properties of tablets by exclusively utilizing the in-die data from compression cycles. To achieve this, methods had to be found in order to collect, organize, reduce, and evaluate the data. Subsequently, a model should be derived that, by exploiting the serial information in the data, could be used to calculate the tensile strength of the produced tablets.

1.6 Outline of the thesis

This work consists of four publications in peer reviewed scientific journals. In the second chapter, a new method for the evaluation of stationary time series is introduced and it's use-case demonstrated. The new analytical procedure serves to quantify the chaotic portion in data and provides the opportunity to characterize such data. Additionally, the method works with the data of the internal dosing unit of gravimetric bulk solid dispensers, which is why it can theoretically be employed in-line. The application of this technique is demonstrated on the example of bulk solid dosing data of two different materials. Through this exemplary application, it is shown that the data from external scales, used to catch the dosed material, may display strong artifacts that tremendously complicate the evaluation of bulk solid dosing processes.

The further chapters of this work deal with the non-stationary data series of tableting processes. The third chapter describes the examination of a constructed database, which should serve as a foundation for later investigations. Here, the compression of 12 raw materials at various machine settings is investigated in detail using multivariate decomposition. Beside the demonstration of the reliability of the data, the automatic pre-processing and evaluation of the data is the focus of this study within this thesis. Methods are applied that enable reproducible and precise detection of the typical data patterns of the compression phase of the tableting process. Understanding the influence of tableting parameters and material properties on the data enables a targeted expansion of the database. This is explained in chapter four. This chapter investigates binary mixtures of pharmaceutical excipients for tableting. The influence of the component ratios on different parameters could be clarified, and several trends are described. This work is further incorporated by the additional expansion of the data in chapter 5. This chapter also deals with the extension of the data by including ternary and quaternary mixtures, as well as the construction of an artificial neural network. The architecture of the network is described in detail, and the influence of individual structures is explained. An overview of the training of the model, as well as the prediction of the tensile strength of the tablets in the database, is provided. It is shown that the prediction accuracy of the network is uniformly precise and correct across different tablet masses and materials. Furthermore, a new method has been developed and explained that allows determining the compression profiles of different substances using a single compression cycle per material. The successful application of the novel method across a wide and realistic range of compression parameters is demonstrated. The final section of the work contains the discussion about the conducted studies and the outline for experiments and future improvements.

2 Classification of deviations in pharmaceutical continuous bulk solid dosing

Pretext

The evaluation of fluctuations of pharmaceutical processes in the steady state is not only of scientific interest but also a focus for the understanding of critical parameters and the control of production lines. Since the dosing of powders and bulk solids usually presents the first process unit in continuous production of oral solid dosage forms, the analysis of the available data of dosing processes is of highest importance for the understanding of the later processes. The consideration of the flow properties of the dosed materials, the complex and chaotic processes in flowing bulk solids, as well as externally caused fluctuations play a role in the evaluation of such data. The uniformity of the material flow into the downstream processes like tableting influences the homogeneity of the products. The investigation in the improvement of continuous dosing processes were therefore focused in the first chapter for this work. Result of the studies about powder dosing was an analytical procedure, utilizing the frequency domain of time series, to evaluate the chaotic character of processes. On the example of continuous material feeding the benefit of novel perspectives on the data was demonstrated and lead to a deeper understanding for properties of pharmaceutical process related data series. This publication concluded the work on data series of a stationary characteristic. After the completion of the work, the focus was shifted from powder dosing processes to non-stationary compression profiles, which display entirely different properties in various aspects. These studies will be addressed in the next three chapters.

The following research paper was published in volume 412 of “Powder Technology” in the year 2022 with the number 118003. The first author developed the idea to characterize the behavior of fluctuations in dosing processes. Julian Quodbach supported with ideas to the design of collection vessels and the comparison with the external balance. The experimental investigations were conducted by Stefan Klinken. He developed the Software and performed the evaluation of the data. He developed the evaluation method from scratch. The first draft and the main part of the final manuscript was written by Stefan Klinken. Julian Quodbach supported with reading and editing of the draft.

author / co-author	idea [%]	study design [%]	experimental [%]	evaluation [%]	manuscript [%]
Stefan Klinken	90	80	100	100	75
Julian Quodbach	10	20	0	0	25

Sums of amplitudes analysis – A new non-parametric classification method for time series deviation evaluation in pharmaceutical processes

Stefan Klinken and Julian Quodbach

Institute of Pharmaceutics and Biopharmaceutics,

Heinrich Heine University, Duesseldorf

Powder Technol. 2022, doi.org/10.1016/j.powtec.2022.118003

Abstract

The evaluation of various types of data and the understanding of their properties is the backbone of scientific work. Time series data are abundant in the investigation of manufacturing processes. Since they result from measurements of parameters over time, time series have distinctive statistical characteristics. Focused on bulk solid dosing, the authors illustrate the potential for misinterpretations if these characteristics are ignored in data evaluation. The sums of amplitudes analysis (SAA) is developed, based on discrete Fourier transformations as a novel non-parametric statistical method for description of time series. Subsequently, the authors use the SAA to investigate bulk solid dosing data. In this process, artifications of data of external catch scales, commonly used for the evaluation of bulk solid dosing, are found, and discussed.

3 Acquisition and multivariate decomposition of data of pharmaceutical tableting processes

Pretext

This publication was created as part of a collaboration with a research project of the topic of "Mechanical deformation behavior of materials during compression." In this project 12 starting materials were tableted using a compaction simulator. The materials were selected to range over a range of chemical entities, including cellulose and its derivatives, α-lactose, anhydrous dibasic calcium phosphate, carrageenan, and starches. Furthermore, various grades of lactose, anhydrous dibasic calcium phosphate and corn starch were included in the study. The different types possess varying particle size distributions, and/or particle shapes. Corn starch was examined in both a native and pre-gelatinized grade. Various settings of the tableting machine were investigated including different tablet weights, two different punch tip geometries and two compression speeds. The aim of the study was the investigation of correlations and collinearities of compressibility and tableting parameters including for example the energies of the force-displacement, the tabletability, the coefficients of the equations according to Heckel, Kawakita or Kuentz-Leuenberger or the solid fraction of the tablets. To handle the database of more than 3000 tableting cycles, an algorithmic computational approach was developed in context of this study. The self-developed computer program EvTab was used for the preprocessing of the raw data, as well as for the in-die based evaluation procedures. This algorithmic approach serves not only as the solution in the handling of this large database, but also as the basement for later developed methods. This evaluation procedures are addressed in Chapter 5. The start point of each compression process was analyzed using an iterative change point detection algorithm. The information about the location and dispersion of the compression pressure of the upper punch during the ejection phase served as the calibration set for this change point detection at each data series. From the maximum of the compression pressure each compression profile was screened in reverse order until the compression pressure was lower than the mode of calibration plus one standard deviation. Following this approach, the beginning of each compression cycle was detected without manual adjustments with sufficient precision. Evaluations for the different in-die analyses were carried out using algorithms with minimal need for manual adjustments. The linear portions of the plots according to Heckel, Walker, Kawakita and Johanson were accurately captured through an iterative process of linear regression models. Within this process the portion of the linear regression was varied within predefined limits. The limits were based on the respective compressibility descriptors.

The study demonstrates that the scattering of the different measures is considerably low in the triplicates of each factor-level combination. Therefore, the applicability of the developed algorithms is demonstrated over a broad range of data settings. It is shown that the slopes of the tabletability profiles have collinear components with the ab parameter of the Kawakita equation. Negative correlations of the tabletability were shown with C^{-1} of the Kuentz-Leuenberger equation and K of the Johanson compressibility model. This indicates that the course and shape of the of the compression profiles contain information about the tablet properties like the tensile strength. The different materials form clusters in the data space which are almost perfect to separate in 2 orthogonal dimensions. This indicates that the differences within compression profiles seems to be significant for the clear separation of the materials and their tablet properties. As the axis along which the materials can be distinguished is collinear with the loading of the slope of the tabletability profile in the PCs one and two, the prediction of the tensile strength based on the compression profiles of the substances appeared to be feasible. This target was focused as the aim of subsequent studies in context of this thesis.

As expected, punch pressure had the most significant impact on the variance in the data, while tablet compression speed, punch geometry, and tablet mass had a less pronounced influence. As the highest data density was within the settings of 8 mm punches this punch geometry was selected for following

studies. The optimal compression speed was found with 3 mm s^{-1} as higher compression speeds correlate with the occurrence of fluctuations in the data. In context of this thesis, the study provided the access to a well described and sufficient large database of compression data. The wide range of materials, compressions pressures, punch geometries and compressions speeds enabled the selection of the most reliable settings for the following studies. The subsequent expansion of the database, described in the following chapters, thus utilized these optimal settings. The developed algorithms showed their applicability within the scope of this published investigation. They were partially used in subsequent studies.

The following research paper was published in volume 637 of "International Journal of Pharmaceutics" in the year 2023 with the number 122890. The first author Sabrina Berkenkemper performed the experiments in lab according to the designed study. The idea for this paper and the study design was developed mainly by Stefan Klinken and Sabrina Berkenkemper in equal parts. Peter Kleinebudde assisted with ideas for the study design and handling of the data. The software utilized for the whole project was written in Python code by Stefan Klinken. This includes main parts of the evaluation of the data. Sabrina Berkenkemper performed manual evaluations and sorting of the data. She wrote main parts of the manuscript assisted by Stefan Klinken. Within the reviewing process, Stefan Klinken conducted additional computer assisted investigations on the behavior of Suns and Kuentz-Leuenberger constants to explain the exclusion of data. Peter Kleinebudde supported with reviewing and editing of the manuscript.

author / co-author	idea [%]	study design [%]	experimental [%]	evaluation [%]	manuscript [%]
Sabrina Berkenkemper	40	45	100	45	65
Stefan Klinken	40	45	0	55	25
Peter Kleinebudde	20	10	0	0	10

Multivariate data analysis to evaluate commonly used compression descriptors

Sabrina Berkenkemper, Stefan Klinken, Peter Kleinebudde

Institute of Pharmaceutics and Biopharmaceutics,

Heinrich Heine University, Duesseldorf

Int.J.Pharm. 2023, doi.org/10.1016/j.ijpharm.2023.122890

Abstract

Numerous studies elucidated material behavior based on compression analyses. Especially compressibility, compactibility and tabletability were in the focus of these investigations. In the present study, a comprehensive multivariate data analysis was performed using principal component analysis method. Twelve pharmaceutically used excipients were selected for direct compression tableting and subsequent evaluation of several compression analyses. Material properties, tablet properties, tableting parameters and parameters from compression analyses were used as input variables. The materials could successfully be grouped using principal component analysis. Of the tableting parameters, the compression pressure showed the greatest influence on the results. The tabletability was found to be the most important compression analysis in the material characterization. Compressibility and compactibility only played a minor role in the evaluation. Some important insights have been gained for a deeper understanding of the tableting process using the multivariate approach to evaluate the variety of compression data.

4 Investigations on the compression of binary mixtures

Pretext

The applicability of the algorithms to preprocess and evaluate the data of compression processes was shown in Chapter 3 on a data base of pure pharmaceutical excipients. Hereby, the procedures were successfully shown within various settings of the tableting process including different punch speeds, tablet masses and punch geometries. In the current chapter the database was expanded to binary mixtures of four different materials. In the study, hydroxypropyl cellulose and microcrystalline cellulose were included as plastic substances, while dibasic calcium phosphate and lactose were chosen as brittle materials. The four binary mixtures were described in terms of compressibility, tabletability and compactability. Through varying parameters of the three properties and comparing compression profiles, it was demonstrated that the different mixtures were clearly distinguishable based on the compressibility parameters. Additionally, the different parameters of the Kawakita, Heckel, Kuentz-Leuenberger and Walker equations showed different dependencies to the change of the composition of the mixtures. This clearly indicates that none of them would be sufficient to solely describe the shape of the compression profiles. Based on this observation the focus was on the application of non-parametric approaches for the description of the shape of compression profiles. Moreover, the primary aim of the study was achieved by identifying parameters that were linearly dependent on the mixing ratios of the substances.

For the current work, the expansion of the database through the study was of particular importance as the application of neural networks requires the access to sufficiently large databases. Additionally, it was demonstrated that the algorithm-based evaluation and processing of data is also applicable to new unknown data series. Again, the spread of the six measurements at every factor-level combination was demonstrated to be considerable low. Due to the subsequent renewed demonstration of the distinguishability of substances based on tableting parameters and compression profiles, the next step involved the use of machine learning based regression models to calculate tablet properties from the compression profiles. This will be addressed in the subsequent chapter, during which the underlying database will additionally undergo a final expansion.

The following research paper was published in volume 424 of “Powder Technology” in the year 2023 with the number 118571. The idea to investigate the compression behavior of mixtures was initialized by Stefan Klinken and Sabrina Berkenkemper in equal parts. Peter Kleinebudde provided additional ideas and suggestions on the topic. The lattice of the design was set up by Stefan Klinken while Sabrina Berkenkemper planed the practical investigation including e.g. machine settings, investigation procedures. Main parts of the analysis and manuscript work were carried out by Sabrina Berkenkemper. Stefan Klinken supported in the computational based evaluation of the results and the discussion of the results in the manuscript. Peter Kleinebudde supported with reviewing and editing of the manuscript.

author / co-author	idea [%]	study design [%]	experimental [%]	evaluation [%]	manuscript [%]
Sabrina Berkenkemper	40	60	100	60	70
Stefan Klinken	40	30	0	20	10
Peter Kleinebudde	20	10	0	20	20

Investigating compressibility descriptors for binary mixtures of different deformation behavior

Sabrina Berkenkemper, Stefan Klinken, Peter Kleinebudde

Institute of Pharmaceutics and Biopharmaceutics,

Heinrich Heine University, Duesseldorf

Powder Technol. 2023, doi.org/10.1016/j.powtec.2023.118571

Abstract

One of the most important dosage forms remains tablets. Predicting the tableting behavior of formulations is a major goal in the development of those dosage forms. Resources and consequently costs can be saved if time-consuming experiments for empirical formulation development are eliminated. Predicting the properties of a mixture from the properties of the materials it is composed of has been studied many times. To simplify this approach, binary mixtures were tableted in this work. Both pure materials and mixtures in different compositions were characterized with respect to their compressibility and tablet properties. The aim was to investigate different compressibility descriptors for their suitability to predict the tableting behavior of binary mixtures. For this, linear mixing rule for weight and volume fractions was applied to predict the properties of the mixture from the corresponding starting materials. Compressibility analysis, Walker equation and rearrangement index of Kawakita analysis showed promising results. A linear relationship between the compression parameter and the composition of the binary mixture was found for the three parameters. These descriptors can be used as a foundation to describe more complicated mixtures in the future.

5 Prediction of tensile strength of tablets using LSTM networks on compression profiles

Pretext

The insights gained in the chapters 3 and 4 about compression profiles and their utilization for predicting properties of the powder mixture or the produced tablet were fundamental for the application in the following study. The scope of this publication was to evaluate patterns in the compression profiles and using them to predict the tensile strength of tablets produced from different excipients and mixtures. A particular challenge was dealing with compression profiles of different temporal lengths and different degrees of compaction. The LSTM architecture was utilized to design a recurrent neural network to model the correlation between the shape of compression profiles and the tensile strength of the tablets. A masking method for missing values was implemented to handle the different length of the data series, where the missing data points were end-padded with a value of zero. This enabled the future processing of the model via graphics processing units. Based on a consolidated database of more than 1200 tableting cycles, the model was trained and tested. The data was utilized from portions of the databases obtained in the chapters 3 and 4. Data of ternary and quaternary mixtures as well as data of two more pure components were additionally investigated in this study. After demonstrating sufficient prediction of the tensile strength of individual tablets, a method was developed that allows for the prediction of the tabletability profile of a substance over a wide range of compression pressures. This method was demonstrated on the example of 12 starting materials. The procedure allows in range of the interpolation of the data base material independent predictions of the tensile strength of tablets. Only data from a single compression process per material were required to calculate the complete tabletability profile of the respective substance. The model presented is, to the author's best knowledge, the first model that uses methods of natural language processing to describe the shape of compression profiles. Utilizing an LSTM-based neural network enables remarkably accurate predictions of tensile strength for tablets, covering a broader range of materials than ever before in published literature. The investigation suggests that, due to the inherent information, compression profiles could be used to calculate the tensile strength of tablets made from materials unknown to the model. At this time, the presented model stands alone in the literature for its suitability in in-die determination of tensile strength, exclusively utilizing data from compression pressure, punch movement, and process time. The demonstrated method underscores the potential advantages of considering characteristics of data series via machine learning approaches when analyzing data of pharmaceutical manufacturing processes.

The following research paper was published in volume 645 of “International Journal of Pharmaceutics” in the year 2023 with the number 123280. This work was mainly carried out by Stefan Klinken. Jill de Bisshop performed the compression experiments and analyzed the tablets. The idea to predict the strength of tablets from the shape of compression profiles was developed by Stefan Klinken following experiments to describe the course of such curves via MVDA. He independently developed the design of the study. Both the underlying software and the evaluation were carried out by Stefan Klinken based on a variety of Python packages. The written elaboration was also undertaken by him. Jill de Bisshop read the manuscript but had no influence on its content.

author / co-author	idea [%]	study design [%]	experimental [%]	evaluation [%]	manuscript [%]
Jill De Bisshop	0	0	100	0	0
Stefan Klinken	100	100	0	100	100

Prediction of the tensile strength of tablets using LSTM networks on compression profiles

Jill de Bisshop and Stefan Klinken

Institute of Pharmaceutics and Biopharmaceutics,

Heinrich Heine University, Duesseldorf

Int.J.Pharm. 2023, doi.org/10.1016/j.ijpharm.2023.123280

Abstract

This publication's objective was to predict the tensile strength of tablets using an analysis of process data comprising compression pressure, sampling timestamps, and punch positions. A recurrent neural network, specifically designed with Long Short-Term Memory layers, was utilized to accommodate the time-series characteristics of the data. A dataset from 344 tablet compression cycles was employed for model training, after which the model demonstrated a predictive ability with a coefficient of determination of 0.954 on test data from 804 tableting cycles. The foundational database incorporated data from both pure substances and mixtures consisting of up to four components compressed at various compression pressures and with three different tablet masses. Interestingly, the prediction errors did not exhibit any significant correlation with specific materials, mixtures, maximum compression pressures, or tablet weights. With the aid of the model, it was possible to calculate the entire tabletability profile of twelve substances from just a single compression process each. Models of this nature bear promising potential for future application in the research and development of formulations as well as in production processes to predict tensile strength.

6 Discussion and outlook

6.1 The future of data analysis in bulk solid dosing

The sum of amplitudes analysis introduced in chapter 2 can serve as an example for the examination of stationary time series which exhibit a high level of noise. Within this publication, it was demonstrated that commonly used methods for characterizing powder dosing have the disadvantage that measurement artifacts are likely to occur. These artefacts are prominent in the application of the most utilized qualification method for powder dosing devices. The use of an external catch balance to measure the dosed powder over time produces data which not only reflects the fluctuations in the dosing but also slipping effects of the powder cone on the balance. The presented analytical procedure was shown to be capable of the reproducible identification of this artefacts in the data. Nonetheless, the impact of this study on the industrial application of screw-based bulk solid feeders is limited.

In this regard, multiple factors have to be named which reduce the impact of the presented study. Notably, one of the less influential factors in this publication is the limited range of investigated materials. Furthermore, a comprehensive exploration of the impact of diverse process parameters was not included in the investigation. However, the foremost determinant limiting factor appears to be the question of whether fluctuations in powder dispensers, a topic frequently discussed in the literature at few-seconds intervals, possess meaningful industrial implications. Examinations concerning the blending behavior of continuous powder mixers with respect to input fluctuations reveal that the residence time distribution of the material is the filter function of the fluctuations of input [187]. Theoretical investigations suggest that the consideration of the residence time distribution is crucial in the assessment of fluctuations of the powder dosing in all kind of pharmaceutical manufacturing processes [188]. This convolution of input and filter function results in a smoothed scattering of the output. In numerous pharmaceutical processes involving bulk solid dosing, the median residence times typically span from 10 seconds to several minutes, often accompanied by distribution widths exceeding 30 seconds [23,189–192]. Assuming that input fluctuations are likewise filtered by the residence time in these processes, fluctuations within the frequency range of a few seconds are likely negligible. Nevertheless, it is crucial to acknowledge that potential effects, particularly in processes characterized by extremely short residence times, such as twin-screw wet granulation, or in scenarios featuring pronounced variations in powder dosing, cannot be conclusively dismissed. Consequently, conducting measurements for this purpose may represent a prospective avenue for subsequent research in this thesis. In particular, Fourier spectrum-based studies of various materials must be questioned based on these findings [43]. It is unclear whether the detected fluctuations are due to fluctuations in the powder dosing or caused by slippage processes of the powder cone on the external balance. Differences in the quality of dosing processes may thus be merely differences in the measurements, representing data artifacts.

These controversies clearly show that a measurement method is necessary which can be used in-line and is not subject to material-dependent artifacting. Microwave resonance seems to be a suitable technology for this purpose. It is surprising that, despite the first publication about this topic [193], no further studies on this topic have been conducted. Besides the possibility of in-line implementation, this technology promises another advantage. As described in the publication underlying chapter 2, the current assessment of bulk solid dosing using internal or external balances results in derivative time series. The weight is measured, and the dosing rate is calculated as the temporal derivative of the data. This entails a limitation for the sample frequency. As the measurement intervals become shorter, the denominator of the ratio in the calculation of the derivative also decreases. For very short measurement intervals, the denominator approaches zero, resulting in an increase of the calculated fluctuations. Therefore, measurements at different sample frequencies are only hardly comparable. Microwave technology measures the density in the sensor field, which is directly a measure of the quantity of the falling powder over time and thus not mathematically dependent on the measurement interval. The

application of this technology for external quantification of powder feeders or for in-line monitoring of dosing processes appears to be very promising. It is already shown that microwave resonance technology is able to predict masses of tablets in a range of 355 - 380 mg in an high accuracy with a coefficient of determination of 0.97 in an industrial relevant setting [194]. This finding indicates that reproducible and precise measurements of small mass variations in the sensor field are achievable with this technology. However, the signal of the microwave resonance sensor is dependent on the surrounding temperature [195] and on the humidity of the air [196]. These parameters would need to be additionally recorded and calibrated in order to calculate a compensation term for the mass flow rate. The use of resonance microwave sensors would allow monitoring the dosing with very high sampling rates. Higher sampling rates offer the possibility to precisely work out actual differences between the dosing of different materials. The presented methodology of SAA for data analysis can then be one part in the toolbox to gain an accurate understanding of the properties of the collected data.

The SAA is part of a multitude of metrics for the sequence of time series, representing parameters for the entropy of the time series [197–199]. Particularly, comparing the SAA with the permutation entropy, introduced by [111], or with the Fourier entropy [200] would be of highest interest. Future work could reveal the advantages and disadvantages of the methods in the application to data from pharmaceutical processes.

6.2 Advancements in the regression of compression profiles

The evaluation of compression processes presented in this work has made a decisive advance in information extraction from compression profiles in comparison to the former literature. As described in the introduction of this thesis, the description of the course of compression profiles in the domains of pressure and distance between punches is frequently conducted using least-square regression-based models. The resulting compressibility models do not account the temporal dimension of the compression processes. Within chapter 3 the punch speed was identified as a less influential parameter on the compression profiles and the properties of the tablets. Even through the findings of Tye et al. [71] suggest the presence of an effect of the tableting speed on the tabletability, the compactability of the materials was not influenced. Similar findings are given by Haware et al. [201]. In contrast, investigations on the compression of paracetamol indicate that the speed of the compression process influences the elastic energy and therefore the capping tendency of the produced tablets [202,203]. The consideration of the temporal component of compression processes therefore seems to be relevant for some materials. Consequently, the neglect of this aspect can be seen as a minor disadvantage of the described bivariate models.

The advantage of the parametric regression models lies within their simplicity. They typically consist of only a few coefficients. However, this advantage is also the major weakness of the models. The coefficients are frequently not orthogonal, as demonstrated in chapter 3 [204]. The presented correlation matrix indicates that coefficients within the same model often correlate. In combination with the simplicity of the models, this reduces the flexibility of the regression to describe the local structure within the data. Historically, this can be justified since most measurements were typically performed using out-of-die methodology. An out-of-die investigation under a considerable effort, only results in a few points describing the compression profiles. Consequently, models with a higher count of coefficients were impractical due to the limited degrees of freedom. With the widespread application of PAT for measuring punch pressure and punch positions, high-resolution compression curves are now available. Therefore, the only remaining justification for using models with fewer coefficients is their ease of interpretation. However, considering the variety of the processes which can take place during the compression of bulk solids [78], it seems that models with only a few coefficients may be too simplistic.

It must be mentioned that the describability of the curves by the compressibility regression models is dependent on the material. The data on Heckel, Kawakita, and compressibility analysis according to Johanson in chapter 4 clearly show that some curves can be described over large ranges by the linear models. Curves of other materials show significant deviations from linear trajectories, and thus a poorer describability by the models [205]. Similar observations can also be made in the supplementary materials of chapter 3. Here, some curves show good describability by the Kuentz-Leuenberger equation, while others exhibit significant deviations [204].

Another insight extracted from chapter 3 is that the regression results of the different models for the compressibility of substances can contain orthogonal information. For this reason, not just one of the models can be used to describe the same information about the compressibility of the substances. As an example, the slope of the Walker equation and the reciprocal of the intercept of the Kawakita equation can be mentioned. In respect to the described variance of around 60 % in the two-dimensional biplot of the study, it can be assumed that further orthogonal information of the compression parameters exists. The results of a similar study by Dai et al. [206] with a explained variance of approximately 55 % gives a comparable conclusion. For some parameters this seems to be addressable to the different portions of the data on which the models are applicated. As already mentioned in the introduction of this thesis, the portion of the data which is described differs between the different models. Therefore, it is not surprising that the information obtained from the parameters of the models is partially orthogonal. The present work utilizes algorithmic detection of the portions for the different models and increases thereby the reproducibility of the evaluation methods.

The consideration of the disadvantages of the mentioned models resulted in the application of a recurrent neural network in this presented work. With this, it was possible to describe the compression profiles of the substances. Additionally, the information from the compression profiles could be utilized predicting tensile strength of the resulting tablets. The introduced LSTM model consists of around 40,000 trainable coefficients. Beside the detection of the first contact of the upper punch with the powder, no further change point detection algorithm was needed for the pretreatment of the data. The recurrent architecture of the model enables a precise regression of the time series data and the exploitation of the underlying autocorrelation within the data. To the authors best knowledge, no other published study describes a precise regression of the compressibility and the tabletability over a broad range of materials. The methodology presented here is most comparable to the study by Lee et al [90]. They used a cross-decomposition method via projection to latent structures by partial least square regression to calculate tensile strength based on interpolated compression profiles. The interpolation of the curves in this study is possible because the x-data have an almost identical range. This directly prevents the application of the method to substances with significantly different compression behaviors. Also, a comparison of different masses or multiple punch speeds is not feasible with this model. They compared two different models. One model was interpolated over time, while the other model interpolated the force respective to the displacement of the punches. Both models achieved no noteworthy predictive quality, with a coefficient of determination ranging between 0.4 and 0.8. Additionally, no cross-validation of the results was conducted, which makes the statistical significance of the outcomes questionable.

Considering that the compression profiles of different substances sometimes show clear differences, it seems surprising that only a few studies deal with the prediction of tablet properties from these profiles. Characteristics of compression curves can intuitively be attributed to specific tabletability profiles. This becomes apparent from the examples of substances with very high tabletability, such as microcrystalline cellulose or hydroxypropyl cellulose, in comparison to substances with lower tabletability like carrageenan or lactose. These examples are highlighted in the data of chapter 3 and 4 of this work [204,205]. Another indication of the relationships between compressibility parameters and the tabletability of substances is provided by the correlations shown in the study in chapter 3. Here, for example, a collinearity between the reciprocal intercept of the Kawakita equation and the tabletability could be demonstrated. However, a mathematical regression of such relationships is challenging without suitable, straightforward parametric methods. Furthermore, a major demand in the field of

research on pharmaceutical compressions is to employ models with a physical basis. However, the study in chapter 5 clearly demonstrates that a method which takes into account the complexity of the processes occurring during compression surpasses the methods previously published.

The demonstrated model can be understood as a prototype. It certainly represents the first model that could enable a material-independent in-die prediction of the tensile strength of tablets. In contrast to methods using ultrasound, near infrared, or other spectrometric measurements, the presented model is not calibrated for a specific composition of substances. Furthermore, it is not limited to a particular mixture system, such as linear or geometric models based on pure substances for calculating binary mixtures. A further advancement in predicting tabletability is the demonstrated curve striping of the compression profiles. Since the descending part of the compression curve is not considered, the LSTM-based model is able to calculate any tensile strength below the measured maximum compression pressure. This makes this method suitable for predicting the current tensile strength of the tablet in real-time during the compression of a material. This is of enormous value in the development of pharmaceutical formulations.

The application of artificial intelligence in scientific data analysis faces primarily two major critiques. First, there is the complaint about the black box nature of the models. Often, the predictions made by the models are difficult to verify, and it is usually unclear how each part of the input contributes to the output. The second critique concerns the complexity of the models and the associated need for large amounts of data.

In a comparison of the work presented here and models from the literature, the difference in the size of the data used is clear. For instance, the model by Lee et al. [90] utilizes 108 data series for training the model. Another study for tensile strength calculation supported by cross-decomposition by Casian et al. [207] used 60 data points for calibration. Jin et al. [86] even utilized only 32 data points in their cross-decomposition. In contrast, the model presented in this work utilizes 344 data points, which is between three and ten times the data volume of comparable studies. However, the diversity of data in this work is much greater than in the models known from the literature. Additionally, it must be emphasized again that the model presented here, unlike other comparable methods, appears to allow for meaningful extrapolations between materials and does not require material-dependent calibration. The study by Djuris et al. [87] suggests that for less complex neural networks for predicting the tensile strength of tablets, significantly fewer data points are required for successfully training of the data. The conclusion from the comparison within the literature is that the fear of an enormous experimental effort for the training of artificial intelligence in the realm of compression analysis seems to be unfounded.

Aside from the required amount of data, the second point of criticism the black box nature of neural networks has to be addressed. The introduced curve-stripping method of compression profiles made it possible to precisely estimate the tabletability profiles of substances from just a single compression process. The question of the influence of different portions of curves can thus be easily answered. The entirety of the curve up to a certain point contributes to the predicted strength. This does not only include the pure values but also the order and correlations of the values. The exact influence of each data point can therefore be precisely explained for this model, with the influence of each point clearly depending on every data point before and after it. This consideration significantly reduces the black-box character of the model.

However, the database of underlying materials is severely limited. Beside limitations in the compression profiles and the punch geometries the database lacks for mixtures that correspond to real formulations. The inclusion of APIs or excipients with different deformation behavior, as well as the investigation of real formulations, may potentially require an increase in the complexity of the model. In this context it will be crucial to strike a balance between the diversity and size of the underlying database and the complexity of the model to avoid overfitting of the training data. Possible extensions of the model include the use of attention mechanisms such as multi-head attention layers [208]. The comparison of the recurrent structure with one-dimensional convolutional networks would be of high

interest to enable further improvements of the model. One-dimensional temporal convolutional networks employ an architecture distinct from recurrent networks to analyze autocorrelation in time series data. This architecture demonstrated an improved performance on time series data in comparison to LSTM based networks [209].

Further extensions of the demonstrated methodology include the use of data from rotary tablet presses. Rotary tablet presses represent production scaled tableting machines and are, therefore, highly relevant for the transfer of the method's use from formulation development to industry relevant process monitoring. In this context, three main limitations arise. Firstly, the speeds of compression processes in rotary tablet machines are much faster. The punch speeds can easily exceed 300 mm s^{-1} [65]. The second limitation is based on the lesser instrumentation of these machines with PAT. While most modern rotary tablet presses are equipped with force measurement instrumentation, the measurement of the punch displacement is less common. When data on punch displacement are available, they are usually results of simulations rather than measured values. Additionally, the recording frequency is often lower than the 10 kHz used in this work. Usually, an encoder is used to trigger data collection at fixed angles of the rotor [66]. This results in a constant number of data points for each compression process. Consequently, relative to the force-time profile of the data, the density of data points decreases as the machine runs slower. However, the reduced number of datapoints is likely not a significant hurdle. The data shown in chapter 5 were reduced by a decade, corresponding to a sample frequency of 1 kHz. The absence of punch displacement data presents a more substantial problem. Future studies have to investigate whether simulations are sufficient to adequately replace the data of the displacement of the punches.

7 Summary

The advent of new technologies and trends in the development and manufacturing of pharmaceutical dosage forms has led to increased availability of data to describe the processes and properties of produced goods. Values from individual measurements do not only provide relevant information about processes, but the correlation of values within and between data series also contributes to the overall information contained in the data. The applicated methodology generally falls under the term of time series analysis. In this current work the focus was laid on the subprocesses in the direct compression of bulk solids into tablets. Beside the compression process the powder dosing of the bulk solid materials was investigated. In this context a novel method for the qualification of powder dosing units was developed. The new procedure allows for the systematic analysis of deviations in stationary time series of the feed rate over time. The determination between chaotic events and systematic deviation from running mean of the time series was done based on permutations and subsequent transformation in frequency domain of the data. During this procedure artefacts within commonly used data were shown and discussed.

In the context of the compression step of materials in direct compression a database was created. It was analyzed regarding various features including measures of the compressibility, tabletability and compactability of the materials, as well as changes of these properties in mixtures of excipients. In two separated studies the diversity and reliability of the database was demonstrated successfully. In the first investigation about pure pharmaceutical excipients different compression speeds, geometries of the tooling and target tablet masses were analyzed. It was shown within this study that the accuracy of the compression speed of the compaction simulator was heterogeneous over the range of the analyzed range. It was concluded that the punch speed of 3 mm s^{-1} seemed to be most applicable. Therefore, in the future expansions of the database this punch speed was targeted. The multivariate approach revealed that the tensile strength of tablets seemed to be correlated with certain shape of the compression profiles of the substances. Mainly the ab parameter in the Kawakita equation was found to highly correlate with the hardness of the produced tablets. The prediction of the tensile strength of tablets based on their compression profiles was therefore set as the target of this project. As the evaluation of pure components is of academic interest but of less practical impact an expansion of data to binary mixtures was done. It was shown that small changes of the portions within the mixtures can results in sharp changes of compression descriptors or tablet properties, while others were of higher correlation with the change of the composition. The differences in the behavior within the descriptors of the compressibility over the change of the mixtures composition indicated that the utilization of multiples descriptors of the compression profiles would be best for the prediction of the tensile strength. The inclusion of ternary and quaternary systems in following studies further increased the diversity of the database towards settings comparable to industrial application. In the context of database management, the development of algorithmic data handling approaches was the key challenge in all these projects.

The final study presented in this work finalized the application of time series analysis in data of processes, related to direct compression. Based on the tools in natural language processing a LSTM layer based neural network was designed to analyze the curve shape of the compression profiles in the database. After a detailed descripted training procedure, the model was used to calculate the tensile strength of more than 800 tablets of various composition with a coefficient of determination 0.954. The detailed evaluation of residuals revealed no systematic correlation of errors in the model's prediction and properties of the materials investigated. The power of the model was demonstrated in the development of a procedure to calculate the tabletability profile of substances or mixtures based on a single compression process. Hereby an iterative curve stripping procedure was utilized for the generation of sub curves of the single compression profile. The applicability of the method was shown on 12 of the pure materials.

This thesis demonstrated the application of methodology of the field of statistics and data science in the evaluation of data of pharmaceutical processes in the production of tablets. It was shown that the

proper data handling and database maintenance as well as the assurance of the interoperability of the data, combined with suitable analyzing tools yields promising results on pharmaceutical data. The highlight of this work certainly is the material independent prediction of the tensile strength of tablets based on in process data. This method provides the first tool described in the literature to investigate the tabletability and compactability of materials solely based on inline data from the tableting process.

The understanding of bulk solid dosing processes is still increasing and the study within this work demonstrates that external calibration of dosing devices can lead in severe artefacts in the data. One focus of future research in this field must therefore be on the development of PAT based systems improve the monitoring and control of bulk solid dosing processes.

As the benefit from machine learning increases with the increase in the amount of data, the expansion of the database is a substantial part of future studies. The aim of this work will be to include various settings and machine parameters ranging from punch geometries overcompression speed to various compression profiles. The investigation of industrially relevant mixtures of bulk solids will improve the applicability of the model. Improving the architecture of the model and increasing its complexity, including addition of layers, convolutional structures or multi-head attention, may potentially demonstrate further improvement of the predictability.

8 Zusammenfassung

Neue Technologien und Trends in der Entwicklung und Herstellung von Arzneiformen führen zu einer erhöhten Verfügbarkeit von Daten, welche den Prozess oder die Produkteigenschaften beschreiben. Dabei enthalten nicht nur die einzelnen aufgenommenen Werte Information. Die Abfolge von Einzelwerten kann aufgrund der Abhängigkeiten zwischen den Werten zusätzliche Information beinalten. Besonders die chronologische Abfolge von Werten ist von Wichtigkeit in der Betrachtung pharmazeutischer Prozesse. Diese besondere Form der Datenreihen wird als Zeitreihen bezeichnet. Die Wichtigkeit in ihrer korrekten Auswertung wird durch die Zunahme von kontinuierlichen Herstellungsverfahren in der pharmazeutischen Industrie mehr und mehr bedeutsam für die pharmazeutische Forschung. In der Auswertung der Daten kann durch Berücksichtigung von Auto- und Kreuzkorrelation die Beschreibung von Prozessen verbessert werden. Zusammengefasst finden sich diese Methoden meist unter dem Begriff der Zeitreihenanalyse. Von besonderem Fokus für die gezeigte Arbeit waren Prozesse im Kontext der Herstellung von Tabletten durch Direkttablettierung. Dabei wurde vor allem die kontinuierliche Dosierung und die Verpressung von Haufwerken fokussiert. Am Beispiel der kontinuierlichen Pulverdosierung wurde eine neue Methode zur Charakterisierung von Datenreihen vorgestellt. Mithilfe dieser ist es möglich die Systematik gemessener Prozessschwankungen zu ermitteln und zu vergleichen. Die Methode basiert auf einer Permutationsanalyse und anschließender Fouriertransformation der gemessenen Datenreihe. Die Summe der Amplituden der Frequenzdomäne der Daten und der Permutationen wurde im Verhältnis betrachtet. Es konnte gezeigt und diskutiert werden, dass dieses Verhältnis Informationen über den Anteil der chaotischen Schwankungen an den Prozessvariationen hat. Systematische Variationen, etwa gleichförmige Änderungen oder sinusoidale Schwankungen, können dann quantifiziert werden. Diese Methode bietet eine neue Möglichkeit in der Beschreibung von Prozessdaten. Durch die Anwendung der entwickelten Methode konnten Artefakte, die durch die externe Vermessung von Pulverdosiervorgängen entstanden, beschrieben werden. Das Auffangen von Pulver auf einer externen Waage resultiert in der Bildung eines Pulverkegels. Dabei kann abrutschendes Material an diesem Kegel zu detektierten Schwankungen in den Daten der Waage führen. Es konnte gezeigt werden, dass zwei unterschiedliche Materialien, ähnlich guter Fließeigenschaften sehr unterschiedliche Artefakte erzeugen. Diese Artefakte sind von Interesse, da die externe Untersuchung von Pulverdosiergeräten weit verbreitet ist. Die neue Evaluation der Daten ist ein Beispiel für die Nutzbarmachung von Korrelationsinformationen in stetigen Datenreihen im pharmazeutischen Kontext.

In der weiteren Arbeit werden Datenreihen untersucht, die keinen stetigen Charakter besitzen. Im Kontext der Herstellung von Tabletten wurden in drei aufeinander aufbauenden Studien Kompressionsprofile von Reinstoffen und physikalischen Mischungen analysiert. In einer ersten Studie wurden 12 verschiedene Substanzen mittels eines Kompaktionssimulators verpresst. Es wurden Substanzen in die Studie inkludiert, welche stark unterschiedliche Kompressionseigenschaften aufweisen. Neben sehr plastisch verformbaren Substanzen wie Hydroxypropylcellulose oder mikrokristalliner Cellulose oder stark elastischem Carrageenan finden sich auch spröde Stoffe wie Lactose und dibasisches Calciumphosphat in der Datenbank. Um auch Einflüsse unterschiedlicher Qualitäten chemisch gleicher Materialien zu untersuchen, wurden vier unterschiedliche dibasische Calciumphosphate und drei Typen von Lactose in die Datenbank aufgenommen. Um die optimalen Parameter für die weiteren Untersuchungen zu finden, wurden mehrere Einstellungen untersucht. Dabei wurden Tablettenmassen von 150, 200 und 250 mg, Tablettiergeschwindigkeiten zwischen 3, 30 und 300 mm s^{-1} und 8 sowie 11,26 mm Stempeldurchmesser eingeschlossen. Um die Eigenschaften der eingeschlossenen Daten genauer zu beschreiben, wurde in der in dieser Arbeit vorgestellten Studie eine multivariate Dekomposition unterschiedlicher Deskriptoren von unter anderem der Kompressibilität, Tablettierbarkeit und Kompaktierbarkeit der Materialien vorgenommen. Es konnte gezeigt werden, dass viele Kompressionsdeskriptoren ähnliche Anteile von Informationen liefern. Als Nebenergebnis der Studie konnten langsame Tablettiergeschwindigkeiten von 3 mm s^{-1} als geeignet für die optimale Differenzierung der Substanzen identifiziert werden, da höhere Geschwindigkeiten zu starken Fehlern der Stempelbewegung des Kompaktionssimulators führten. Die Tablettenmasse wurde als weniger

wichtiger Einflussfaktor auf die Kompressionsparameter bestimmt. Von besonderem Interesse für diese Arbeit war die Erkenntnis, dass der Kompressibilitätsparameter ab der Kawakitagleichung Kollinearität mit der Tablettiebarkeit der Substanzen aufweist. Eine Verbindung des Kompressionsprofils mit der Festigkeit der entstehenden Tabletten wurde aus diesem Grund vermutet.

In einer weiteren Studie im Rahmen dieser Arbeit wurden die durchgeführten Untersuchungen auf binäre Mischungen ausgeweitet. Hierbei wurden zwei plastische und zwei sprödbrüchige Substanzen inkludiert. Einer der Hauptaspekte der publizierten Studie lag auf der Anwendung linearer Gesetzmäßigkeiten für Kompressionsparameter. Zudem konnte gezeigt werden, dass die Kompressionsprofile unterschiedlicher Mischungen eindeutige Unterschiede aufweisen. Auch wurde demonstriert, dass die verschiedenen Kompressionsdeskriptoren in unterschiedlicher Weise von der Zusammensetzung der Mischungen abhängig sind. Die Berücksichtigung der Parameter von nur einer der in der Literatur viel beschriebenen Gleichungen, etwa nach Heckel oder Kawakita schien nicht ausreichend zu sein, um die Kompressionsprofile zu charakterisieren. Die nichtparametrische Beschreibung der Profile der Kompressionsvorgänge wurde aus diesem Grund als wichtiger Aspekt in der Konstruktion eines Modells für die Druckfestigkeit einbezogen. Die Untersuchungen wurden bei einer Geschwindigkeit von 3 mm s^{-1} und Stempeln von 8 mm Durchmesser durchgeführt. Die erhaltenen Ergebnisse dienten als Grundlage für folgende Studien zu komplexeren Mischungen von bis zu vier Komponenten.

Der Transfer der gesammelten Erfahrungen zur Datenvorbearbeitung von Kompressionsdaten und Kompressionsprofilen wurde in der letzten Studie dieser Arbeit zusammengetragen. Im Rahmen dieser Studie wurden 14 Einzelkomponenten sowie binäre, ternäre und quaternäre Mischungen untersucht. Das Ziel der Untersuchung war die Generierung eines Modells zur Vorhersage der Tablettenfestigkeit anhand der Form der Kompressionsprofile. Für diesen Zweck wurde ein rekurrentes neuronales Netzwerk erstellt. Dieses ist, aufgrund der eingesetzten Schichtstruktur geeignet einen hohen Teil der Gesamtinformation aus Datenreihen nutzbar zu machen. Um Fehler im Training zu vermeiden, wurden spezielle Architekturen, wie Dropout oder die leaky-ReLu Funktion eingesetzt. Als Kernelement dienten Long-Short-Term-Memory-Schichten. Mithilfe des neuronalen Netzwerks und einem Trainingsdatensatz von 344 Kompressionszyklen konnte die Bruchfestigkeiten eines Testdatensatzes von mehr als 800 Tabletten mit einem Regressionskoeffizienten von 0,954 vorhersagt werden. Anhand einer ausführlichen Diskussion der Residuen der Regression konnte gezeigt werden, dass Parameter, wie etwa die Materialart, Tablettenmasse oder der maximale Kompressionsdruck, keinen Einfluss auf die Richtigkeit des Vorhersageergebnises zu haben scheinen. Basierend auf diesem Modell wurde eine neue Methode zur in-die Bestimmung der Tablettierbarkeit und Kompaktierbarkeit von Materialien entwickelt. Für diese muss ein Material bis zu einem vorher definierten Druck verpresst werden. Durch Abschälen des Kompressionsprofils können dann Subkurven generiert werden, welche anschließend, nach einem Datenvorbereitungsschritt, als Input ins neuronalen Netzwerk dienen. Am Beispiel von 12 der untersuchten Materialien konnte die Anwendbarkeit der Methode erfolgreich demonstriert werden. Die erhaltenen Tablettierbarkeitsplots zeigten eine hohe Richtigkeit der Vorhersage über eine breite Spanne von Kompressionsdrücken. Zwischen 50 MPa bis zu teilweise 600 MPa konnten sehr gute Interpolationen erreicht werden. Für die meisten Materialien waren Extrapolationen niedrigerer Kompressionsdrücke nicht möglich. Durch die angewendeten Methoden wurden Grundlagen für die in-line Messung der Bruchfestigkeit von Tabletten gelegt, auch wenn die Applikation in laufenden Prozessen sicherlich weitere Arbeit erfordern wird.

Zukünftige Forschungen, welche sich an dieser Arbeit orientieren, sollten zwei Kernaspekte aufgreifen. Zum einen ist die externe Kalibrierung von Dosierdaten fehleranfällig und naturgemäß für ein Prozessmonitoring ungeeignet. Aus diesem Grund ist die Erforschung von PAT im Bereich der Druchflussratenmessung von fallenden Haufwerken ein essenzieller Fokus für die pharmazeutische Herstellung.

Im Bereich der Kompressionsanalyse von Pulvern sollten die Datenbanken des neuronalen Netzwerks ausgebaut werden. Hierzu sind zum einen unterschiedliche Maschinentypen, Stempelgeometrien und

Geschwindigkeiten von großer Relevanz. Auf der anderen Seite ist die Untersuchung industrierelevanter Pulvermischungen zukünftig zu leisten. Darunter fällt die Implementierung von Daten von Arzneistoffen, Schmiermitteln, Zerfallshilfsmitteln, sowie von vorprozessierten Zwischenprodukten wie Granulaten.

9 List of References

1. Plum, K., (2005). Continuous processing in the pharmaceutical industry: Changing the mind set. Chem. Eng. Res. Des., 83, 730–738.
2. Myerson, A. S., Krumme, M., Nasr, M., Hayden, T. and Braatz, R. D., (2015). Control systems engineering in continuous pharmaceutical manufacturing May 20–21, 2014 continuous manufacturing symposium. J. Pharm. Sci., 104, 832–839.
3. Lee, S. L., O'Connor, T. F., Yang, X., Cruz, C. N., Chatterjee, S., Madurawe, R. D., Moore, C. M. V., Yu, L. X. and Woodcock, J., (2015). Modernizing pharmaceutical manufacturing: From batch to continuous production. J. Pharm. Innov., 10, 191–199.
4. Throckmorton, D. Statement to the House of Representatives: Examining drug shortages and recent efforts to address them. Washington, February 10, 2014.
5. 21st Century Cures Act: H.R.34, 114th Cong., (2016).
6. Guideline Q13 on continuous manufacturing of drug substances and drug products, ICH, (2022).
7. Guideline Q8 (R2) on pharmaceutical development, ICH, (2006).
8. Yu, L. X., (2008). Pharmaceutical quality by design: Product and process development, understanding, and control. Pharm. Res., 25, 781–791.
9. PAT: A framework for innovative pharmaceutical development, manufacturing, and quality assurance, US FDA, (2004).
10. Shumway, R. H. and Stoffer, D. S., (2017). Time series analysis and its applications: With R Examples, 4th ed. 2017, ISBN 9783319524528.
11. Azcarate, S. M., Ríos-Reina, R., Amigo, J. M. and Goicoechea, H. C., (2021). Data handling in data fusion: Methodologies and applications. TrAC, Trends Anal. Chem., 143, 116355.
12. Ferrero, E., Brachat, S., Jenkins, J. L., Marc, P., Skewes-Cox, P., Altshuler, R. C., Gubser Keller, C., Kauffmann, A., Sassaman, E. K., Laramie, J. M., Schoeberl, B., Borowsky, M. L. and Stiefl, N., (2020). Ten simple rules to power drug discovery with data science. PLoS Comput. Biol., 16, e1008126.
13. Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G. J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M., Ronneberger, O., Bates, R., Kohl, S. A. A., Potapenko, A., Ballard, A. J., Romera-Paredes, B., Nikolov, S., Jain, R., Clancy, E., Reiman, D., Petersen, S., Senior, A. W., Kavukcuoglu, K., Birney, E., Kohli, P., Jumper, J. and Hassabis, D., (2021). Highly accurate protein structure prediction for the human proteome. Nature, 596, 590–596.
14. Meng, X.-Y., Zhang, H.-X., Mezei, M. and Cui, M., (2011). Molecular docking: A powerful approach for structure-based drug discovery. Curr. Comput.-Aided Drug Des., 7, 146–157.
15. Lo, Y.-C., Rensi, S. E., Torng, W. and Altman, R. B., (2018). Machine learning in chemoinformatics and drug discovery. Drug Discov. Today, 23, 1538–1546.
16. Žagar, J. and Mihelič, J., (2022). Big data collection in pharmaceutical manufacturing and its use for product quality predictions. Sci. Data, 9, 99.
17. Casian, T., Nagy, B., Kovács, B., Galata, D. L., Hirsch, E. and Farkas, A., (2022). Challenges and opportunities of implementing data fusion in process analytical technology—A Review. Molecules, 27, 4846.
18. Steinwandter, V. and Herwig, C., (2019). Provable data integrity in the pharmaceutical industry based on version control systems and the blockchain. PDA J. Pharm. Sci. Technol., 73, 373–390.

19. Damiati, S. A., (2020). Digital Pharmaceutical Sciences. AAPS PharmSciTech, 21, 206.

20. GlobalData Plc, (2022). Contract Pharmaceutical Dose Manufacturing Industry: Composition, Size, Market Share and Outlook – 2022 Edition. Available online: https://www.globaldata.com/store/report/contract-pharmaceutical-dose-manufacturing-industry-analysis/ (accessed on 11 January 2023).

21. Fact.MR, (2022). Oral Solid Dosage Pharmaceutical Market. Available online: https://www.factmr.com/report/oral-solid-dosage-pharmaceutical-market (accessed on 11 January 2023).

22. Schulze, D., (2008). Powders and bulk solids: Behavior, characterization, storage and flow, 1st ed., ISBN 978-3-540-73767-4.

23. Meier, R., Thommes, M., Rasenack, N., Moll, K.-P., Krumme, M. and Kleinebudde, P., (2016). Granule size distributions after twin-screw granulation - Do not forget the feeding systems. Eur. J. Pharm. Biopharm., 106, 59–69.

24. Muehlenfeld, C. and Thommes, M., (2014). Small-scale twin-screw extrusion - Evaluation of continuous split feeding. J. Pharm. Pharmacol., 66, 1667–1676.

25. Harting, J. and Kleinebudde, P., (2018). Development of an in-line Raman spectroscopic method for continuous API quantification during twin-screw wet granulation. Eur. J. Pharm. Biopharm., 125, 169–181.

26. Zimmermann, M. and Thommes, M., (2021). Residence time and mixing capacity of a rotary tablet press feed frame. Drug. Dev. Ind. Pharm., 47, 790–798.

27. Oka, S. and Muzzio, F. J., (2022). Continuous powder mixing and lubrication. How to design and implement powder-to-tablet continuous manufacturing systems; Academic Press, pp 59–92, ISBN 978-0-12-813479-5.

28. Karttunen, A.-P., Poms, J., Sacher, S., Sparén, A., Ruiz Samblás, C., Fransson, M., Martin De Juan, L., Remmelgas, J., Wikström, H., Hsiao, W.-K., Folestad, S., Korhonen, O., Abrahmsén-Alami, S. and Tajarobi, P., (2020). Robustness of a continuous direct compression line against disturbances in feeding. Int. J. Pharm., 574, 118882.

29. Engisch, W. E. and Muzzio, F. J., (2015). Loss-in-weight feeding trials case study: Pharmaceutical formulation. J. Pharm. Innov., 10, 56–75.

30. Mendez, R., Muzzio, F. and Velazquez, C., (2010). Study of the effects of feed frames on powder blend properties during the filling of tablet press dies. Powder Technol., 200, 105–116.

31. Peeters, E., Beer, T. de, Vervaet, C. and Remon, J.-P., (2015). Reduction of tablet weight variability by optimizing paddle speed in the forced feeder of a high-speed rotary tablet press. Drug. Dev. Ind. Pharm., 41, 530–539.

32. Ward, H. W., Blackwood, D. O., Polizzi, M. and Clarke, H., (2013). Monitoring blend potency in a tablet press feed frame using near infrared spectroscopy. J. Pharm. Biomed. Anal., 80, 18–23.

33. Ketterhagen, W. R., (2015). Simulation of powder flow in a lab-scale tablet press feed frame: Effects of design and operating parameters on measures of tablet quality. Powder Technol., 275, 361–374.

34. Blackshields, C. A. and Crean, A. M., (2018). Continuous powder feeding for pharmaceutical solid dosage form manufacture: A short review. Pharm. Dev. Technol., 6, 554–560.

35. Wang, H., Wu, L., Zhang, T., Chen, R. and Zhang, L., (2018). Continuous micro-feeding of fine cohesive powders actuated by pulse inertia force and acoustic radiation force in ultrasonic standing wave field. Int. J. Pharm., 545, 153–162.

36. Hou, P., Besenhard, M. O., Halbert, G., Naftaly, M. and Markl, D., (2023). Development and implementation of a pneumatic micro-feeder for poorly-flowing solid pharmaceutical materials. Int. J. Pharm., 635, 122691.

37. Fathollahi, S., Kruisz, J., Sacher, S., Rehrl, J., Escotet-Espinoza, M. S., DiNunzio, J., Glasser, B. J. and Khinast, J. G., (2021). Development of a controlled continuous low-dose feeding process. AAPS PharmSciTech, 22, 247.

38. Sacher, S., Fathollahi, S. and Khinast, J. G., (2022). Comparative study of a novel micro-feeder and loss-in-weight feeders. J. Pharm. Innov., 17, 1205–1214.

39. Engisch, W. E. and Muzzio, F. J., (2015). Feedrate deviations caused by hopper refill of loss-in-weight feeders. Powder Technol., 283, 389–400.

40. Crowley, M. E. and Crean, A. M., (2015). Quality by design in an evolving manufacturing sector. EJPPS, 20, 63–69.

41. Dosing accuracy of continuous scales, NAMUR, (2006).

42. Santos, B., Carmo, F., Schlindwein, W., Muirhead, G., Rodrigues, C., Cabral, L., Westrup, J. and Pitt, K., (2018). Pharmaceutical excipients properties and screw feeder performance in continuous processing lines: A quality by design (QbD) approach. Drug. Dev. Ind. Pharm., 44, 2089–2097.

43. Engisch, W. E. and Muzzio, F. J., (2012). Method for characterization of loss-in-weight feeder equipment. Powder Technol., 228, 395–403.

44. Johnson, B. J., Sen, M., Hanson, J., García-Muñoz, S. and Sahinidis, N. V., (2022). Stochastic analysis and modeling of pharmaceutical screw feeder mass flow rates. Int. J. Pharm., 621, 121776.

45. Madarász, L., Köte, Á., Gyürkés, M., Farkas, A., Hambalkó, B., Pataki, H., Fülöp, G., Marosi, G., Lengyel, L., Casian, T., Csorba, K. and Nagy, Z. K., (2020). Videometric mass flow control: A new method for real-time measurement and feedback control of powder micro-feeding based on image analysis. Int. J. Pharm., 580, 119223.

46. Kotzé, R., Wiklund, J. and Haldenwang, R., (2016). Application of ultrasound Doppler technique for in-line rheological characterization and flow visualization of concentrated suspensions. Can. J. Chem. Eng., 94, 1066–1075.

47. Kazys, R., Sliteris, R., Mazeika, L., van den Abeele, L., Nielsen, P. and Snellings, R., (2021). Ultrasonic monitoring of variations in dust concentration in a powder classifier. Powder Technol., 381, 392–400.

48. Bekaert, B., van Snick, B., Pandelaere, K., Dhondt, J., Di Pretoro, G., Beer, T. de, Vervaet, C. and Vanhoorne, V., (2022). In-depth analysis of the long-term processability of materials during continuous feeding. Int. J. Pharm., 614, 121454.

49. Liu, Z., Muzzio, F. J. and Callegari, G., (2023). Powder property change after passing through a feeder: The effect of electrostatics on powder flow. Powder Technol., 425, 118532.

50. Gyürkés, M., Madarász, L., Köte, Á., Domokos, A., Mészáros, D., Beke, Á. K., Nagy, B., Marosi, G., Pataki, H., Nagy, Z. K. and Farkas, A., (2020). Process design of continuous powder blending using residence time distribution and feeding models. Pharmaceutics, 12, 1–20.

51. European Pharmacopoeia (Ph. Eur.): 11th Edition,, European Pharmacopoeia Commission, (2023).

52. Fu, Y., Yang, S., Jeong, S. H., Kimura, S. and Park, K., (2004). Orally fast disintegrating tablets: Developments, technologies, taste-masking and clinical studies. Crit. Rev. Ther. Drug Carr. Syst., 21, 433–475.

53. Norman, J., Madurawe, R. D., Moore, C. M. V., Khan, M. A. and Khairuzzaman, A., (2017). A new chapter in pharmaceutical manufacturing: 3D-printed drug products. Adv. Drug Deliv. Rev., 108, 39–50.

54. Wang, J., Wen, H. and Desai, D., (2010). Lubrication in tablet formulations. Eur. J. Pharm. Biopharm., 75, 1–15.

55. Jivraj, M., Martini, L. G. and Thomson, C. M., (2000). An overview of the different excipients useful for the direct compression of tablets. Pharm. Sci. Technol. Today, 3, 58–63.

56. McCormick, D., (2005). Evolutions in direct compression. Pharm. Technol., 29, 52–62.

57. Hiestand, E. N., Wells, J. E., Peot, C. B. and Ochs J. F., (1977). Physical processes of tableting. J. Pharm. Sci., 66, 510–519.

58. David, S. T. and Augsburger, L. L., (1977). Plastic flow during compression of directly compressible fillers and its effect on tablet strength. J. Pharm. Sci., 66, 155–159.

59. Rees, J. E. and Rue, P. J., (1978). Time-dependent deformation of some direct compression excipients. J. Pharm. Pharmacol., 30, 601–607.

60. Patel, S., Kaushal, A. M. and Bansal, A. K., (2006). Compression physics in the formulation development of tablets. Crit. Rev. Ther. Drug Carr. Syst., 23, 1–65.

61. Metin, C. and Keith, M., (1989). Use of a compaction simulator system in tabletting research. Drug. Dev. Ind. Pharm., 15, 759–800.

62. Korsch AG). STYL'One Evo brochure. Available online: https://www.korsch.com/fileadmin/files/Product-Images/Evo/Broschuere_STYL_one-Evo-DE_01.pdf (accessed on 16 August 2023).

63. Lura, A., Elezaj, V., Kokott, M., Fischer, B. and Breitkreutz, J., (2021). Transfer and scale-up of the manufacturing of orodispersible mini-tablets from a compaction simulator to an industrial rotary tablet press. Int. J. Pharm., 602, 120636.

64. Führer, C., (1963). Über den Druckverlauf bei der Tablettierung: Untersuchungen über die Verdichtungscharakteristik an Rundläufertablettenpressen. Pharmazeutische Technologie, 11, 674–676.

65. Lura, V., Klinken, S. and Breitkreutz, J., (2023). Challenges in the transfer and scale-up of mini-tableting: Case study with losartan potassium. Eur. J. Pharm. Biopharm., 192, 161–173.

66. Bauer-Brandl, A. and Ritschel, W., (2022). Die Tablette: Handbuch der Entwicklung, Herstellung und Qualitätssicherung, 4th ed.

67. Fell, J. T. and Newton, J. M., (1970). Determination of tablet strength by the diametral-compression test. J. Pharm. Sci., 59, 688–691.

68. Hertz, H., (1882). Über die Berührung fester elastischer Körper. J. Reine Angew. Math., 92, 156–171.

69. Kokott, M., Lura, A., Breitkreutz, J. and Wiedey, R., (2021). Evaluation of two novel co-processed excipients for direct compression of orodispersible tablets and mini-tablets. Eur. J. Pharm. Biopharm., 168, 122–130.

70. Kokott, M., Klinken, S., Breitkreutz, J. and Wiedey, R., (2023). Downstream processing of amorphous solid dispersions into orodispersible tablets. Int. J. Pharm., 631, 122493.

71. Tye, C. K., Sun, C. C. and Amidon, G. E., (2005). Evaluation of the effects of tableting speed on the relationships between compaction pressure, tablet tensile strength, and tablet solid fraction. J. Pharm. Sci., 94, 465–472.

72. Sun, C. C. and Himmelspach, M. W., (2006). Reduced tabletability of roller compacted granules as a result of granule size enlargement. J. Pharm. Sci., 95, 200–206.

73. Sun, C. C., (2011). Decoding powder tabletability: Roles of particle adhesion and plasticity. J. Adhes. Sci. Technol., 25, 483–499.

74. Shang, C., Sinka, I. C. and Pan, J., (2013). Modelling of the break force of tablets under diametrical compression. Int. J. Pharm., 445, 99–107.

75. Shang, C., Sinka, I. C., Jayaraman, B. and Pan, J., (2013). Break force and tensile strength relationships for curved faced tablets subject to diametrical compression. Int. J. Pharm., 442, 57–64.

76. Pitt, K. G. and Heasley, M. G., (2013). Determination of the tensile strength of elongated tablets. Powder Technol., 238, 169–175.

77. Busignies, V., Leclerc, B., Porion, P., Evesque, P., Couarraze, G. and Tchoreloff, P., (2006). Investigation and modelling approach of the mechanical properties of compacts made with binary mixtures of pharmaceutical excipients. Eur. J. Pharm. Biopharm., 64, 51–65.

78. Patel, S. and Bansal, A. K., (2011). Prediction of mechanical properties of compacted binary mixtures containing high-dose poorly compressible drug. Int. J. Pharm., 403, 109–114.

79. Reynolds, G. K., Campbell, J. I. and Roberts, R. J., (2017). A compressibility based model for predicting the tensile strength of directly compressed pharmaceutical powder mixtures. Int. J. Pharm., 531, 215–224.

80. Wu, C.-Y., Best, S. M., Bentham, A. C., Hancock, B. C. and Bonfield, W., (2005). A simple predictive model for the tensile strength of binary tablets. Eur. J. Pharm. Sci., 25, 331–336.

81. Wu, C.-Y., Best, S. M., Bentham, A. C., Hancock, B. C. and Bonfield, W., (2006). Predicting the tensile strength of compacted multi-component mixtures of pharmaceutical powders. Pharm. Res., 23, 1898–1905.

82. Michrafy, A., Michrafy, M., Kadiri, M. S. and Dodds, J. A., (2007). Predictions of tensile strength of binary tablets using linear and power law mixing rules. Int. J. Pharm., 333, 118–126.

83. Nassar, J., Williams, B., Davies, C., Lief, K. and Elkes, R., (2021). Lubrication empirical model to predict tensile strength of directly compressed powder blends. Int. J. Pharm., 592, 119980.

84. Puckhaber, D., Finke, J. H., David, S., Serratoni, M., Zafar, U., John, E., Juhnke, M. and Kwade, A., (2022). Prediction of the impact of lubrication on tablet compactibility. Int. J. Pharm., 617, 121557.

85. Haware, R. V., Tho, I. and Bauer-Brandl, A., (2009). Multivariate analysis of relationships between material properties, process parameters and tablet tensile strength for alpha-lactose monohydrates. Eur. J. Pharm. Biopharm., 73, 424–431.

86. Jin, C., Zhao, L., Feng, Y., Hong, Y., Shen, L. and Lin, X., (2022). Simultaneous modeling prediction of three key quality attributes of tablets by powder physical properties. Int. J. Pharm., 628, 122344.

87. Djuris, J., Cirin-Varadjan, S., Aleksic, I., Djuris, M., Cvijic, S. and Ibric, S., (2021). Application of machine-learning algorithms for better understanding of tableting properties of lactose co-processed with lipid excipients. Pharmaceutics, 13, 663.

88. Hayashi, Y., Nakano, Y., Marumo, Y., Kumada, S., Okada, K. and Onuki, Y., (2021). Application of machine learning to a material library for modeling of relationships between material properties and tablet properties. Int. J. Pharm., 609, 121158.

89. Sovány, T., Papós, K., Kása, P., Ilič, I., Srčič, S. and Pintye-Hódi, K., (2013). Application of physicochemical properties and process parameters in the development of a neural network model for prediction of tablet characteristics. AAPS PharmSciTech, 14, 511–516.

90. Lee, W. B., Widjaja, E., Heng, P. W. S. and Chan, L. W., (2021). Effect of excipient particle size distribution variability on compact tensile strength; and its in-line prediction by force-displacement and force-time profiling. Eur. J. Pharm. Sci., 159, 105703.

91. Paul, S., Tajarobi, P., Boissier, C. and Sun, C. C., (2019). Tableting performance of various mannitol and lactose grades assessed by compaction simulation and chemometrical analysis. Int. J. Pharm., 566, 24–31.

92. Razavi, S. M., Callegari, G., Drazer, G. and Cuitiño, A. M., (2016). Toward predicting tensile strength of pharmaceutical tablets by ultrasound measurement in continuous manufacturing. Int. J. Pharm., 507, 83–89.

93. Igne, B., Anderson, C. A. and Drennen, J. K., (2011). Radial tensile strength prediction of relaxing and relaxed compacts by near-infrared chemical imaging. Int. J. Pharm., 418, 297–303.

94. Halenius, A., Lakio, S., Antikainen, O., Hatara, J. and Yliruusi, J., (2014). Fast tablet tensile strength prediction based on non-invasive analytics. AAPS PharmSciTech, 15, 781–791.

95. Juban, A., Nouguier-Lehon, C., Briancon, S., Hoc, T. and Puel, F., (2015). Predictive model for tensile strength of pharmaceutical tablets based on local hardness measurements. Int. J. Pharm., 490, 438–445.

96. Joiris, E., Di Martino, P., Berneron, C., Guyot-Hermann, A.-M. and Guyot, J.-C., (1998). Compression behavior of orthorhombic paracetamol. Pharm. Res., 15, 1122–1130.

97. Heckel R. W., (1961). An analysis of powder compaction phenomenon. Trans. Am. Inst. Min. Metall., 221, 1001–1008, 671.

98. Kawakita, K. and Tsutsumi, Y., (1965). An empirical equation of state for powder compression. Jpn. J. Appl. Phys., 4, 56 - 63.

99. Johanson, J. R., (1965). A rolling theory for granular solids. J. Appl. Mech., 32, 842–848.

100. Kuentz, M. and Leuenberger, H., (1999). Pressure susceptibility of polymer tablets as a critical property: A modified heckel equation. J. Pharm. Sci., 88, 174–179.

101. Sun, C. C., (2004). A novel method for deriving true density of pharmaceutical solids including hydrates and water-containing powders. J. Pharm. Sci., 93, 646–653.

102. Ryshkewitch, E., (1953). Compression strength of porous sintered alumina and zirconia. J. Am. Ceram. Soc., 36, 65–68.

103. Vreeman, G. and Sun, C. C., (2022). A powder tabletability equation. Powder Technol., 408, 117709.

104. Anbalagan, P., Sarkar, S., Liew, C. V. and Heng, P. W. S., (2017). Influence of the punch head design on the physical quality of tablets produced in a rotary press. J. Pharm. Sci., 106, 356–365.

105. Akande, O. F., Ford, James, L., Rowe, P. H. and Rubinstein Michael H., (1998). The effects of lag-time and dwell-time on the compaction properties of 1:1 paracetamol/microcrystalline cellulose tablets prepared by pre-compression and main compression. J. Pharm. Pharmacol., 50, 19–28.

106. Sinka, I. C., Motazedian, F., Cocks, A. and Pitt, K. G., (2009). The effect of processing parameters on pharmaceutical tablet properties. Powder Technol., 2, 276–284.

107. Adolfsson, A. and Nyström, C., (1996). Tablet strength, porosity, elasticity and solid state structure of tablets compressed at high loads. Int. J. Pharm., 132, 95–106.

108. Paul, S. and Sun, C. C., (2017). Gaining insight into tablet capping tendency from compaction simulation. Int. J. Pharm., 524, 111–120.

109. Persson, A.-S., Pazesh, S. and Alderborn, G., (2022). Tabletability and compactibility of α-lactose monohydrate powders of different particle size. I. Experimental comparison. Pharm. Dev. Technol., 27, 319–330.

110. Steendam, R. and Lerk, C. F., (1998). Poly(DL-lactic acid) as a direct compression excipient in controlled release tablets: Part I. Compaction behaviour and release characteristics of poly(DL-lactic acid) matrix tablets. Int. J. Pharm., 175, 33–46.

111. Bandt, C. and Pompe, B., (2002). Permutation entropy: A natural complexity measure for time series. Physical Review Letters, 17, 1741021–1741024.

112. Nichols, T. and Hayasaka, S., (2003). Controlling the familywise error rate in functional neuroimaging: A comparative review. Statistical Methods in Medical Research, 5, 419–446.

113. Zanin, M., Zunino, L., Rosso, O. A. and Papo, D., (2012). Permutation entropy and its main biomedical and econophysics applications: A review. Entropy, 8, 1553–1577.

114. Palani, S., (2022). Signals and systems: A Practical Approach, 2nd ed. 2022, ISBN 9783030757427.

115. Barth, A., (2007). Infrared spectroscopy of proteins. Biochimica et Biophysica Acta (BBA) - Bioenergetics, 9, 1073–1101.

116. Harris, F. J., (1978). On the use of windows for harmonic analysis with the discrete Fourier transform. Proc. IEEE, 66, 51–83.

117. Donoho, D. L., (2006). Compressed sensing. IEEE Trans. Inform. Theory, 4, 1289–1306.

118. Candes, E. J., Romberg, J. and Tao, T., (2006). Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inform. Theory, 52, 489–509.

119. Rioul, O. and Vetterli, M., (1991). Wavelets and signal processing. IEEE Signal Process. Mag., 8, 14–38.

120. Wong, M. W., (2011). Discrete Fourier analysis, ISBN 9783034801157.

121. Wartewig, S. and Neubert, R. H. H., (2005). Pharmaceutical applications of mid-IR and Raman spectroscopy. Adv. Drug Deliv. Rev., 57, 1144–1170.

122. Nyquist, H., (1924). Certain factors affecting telegraph speed. Bell Labs Tech. J., 3, 324–346.

123. Aminikhanghahi, S. and Cook, D. J., (2017). A survey of methods for time series change point detection. KAIS, 51, 339–367.

124. Rajalahti, T. and Kvalheim, O. M., (2011). Multivariate data analysis in pharmaceutics: A tutorial review. Int. J. Pharm., 417, 280–290.

125. Pearson, K., (1901). On lines and planes of closest fit to systems of points in space. Lond. Edinb. Dublin philos. mag. j. sci., 2, 559–572.

126. Izenman, A., (2013). Modern multivariate statistical techniques: Regression, classification, and manifold learning, 2nd ed., ISBN 978-1493938322.

127. Podani, J., Kalapos, T., Barta, B. and Schmera, D., (2021). Principal component analysis of incomplete data – A simple solution to an old problem. Ecol. Inform., 61, 101235.

128. Horvitz, E. and Mulligan, D., (2015). Machine learning: Trends, perspectives, and prospects. Sci., 349, 253–255.

129. Kaufmann, E., Bauersfeld, L., Loquercio, A., Müller, M., Koltun, V. and Scaramuzza, D., (2023). Champion-level drone racing using deep reinforcement learning. Nature, 620, 982–987.

130. Almalioglu, Y., Turan, M., Trigoni, N. and Markham, A., (2022). Deep learning-based robust positioning for all-weather autonomous driving. Nat. Mach. Intell., 4, 749–760.

131. Ekins, S., (2016). The next era: Deep learning in pharmaceutical research. Pharm. Res., 33, 2594–2603.

132. Mazur, H., Erbrich, L. and Quodbach, J., (2023). Investigations into the use of machine learning to predict drug dosage form design to obtain desired release profiles for 3D printed oral medicines. Pharm. Dev. Technol., 28, 219–231.

133. Munir, N., Nugent, M., Whitaker, D. and McAfee, M., (2021). Machine learning for process monitoring and control of hot-melt extrusion: Current state of the art and future directions. Pharmaceutics, 13, 1432.

134. Singhal, K., Azizi, S., Tu, T., Mahdavi, S. S., Wei, J., Chung, H. W., Scales, N., Tanwani, A., Cole-Lewis, H., Pfohl, S., Payne, P., Seneviratne, M., Gamble, P., Kelly, C., Babiker, A., Schärli, N., Chowdhery, A., Mansfield, P., Demner-Fushman, D., Agüera Y Arcas, B., Webster, D., Corrado, G. S., Matias, Y., Chou, K., Gottweis, J., Tomasev, N., Liu, Y., Rajkomar, A., Barral, J., Semturs, C., Karthikesalingam, A. and Natarajan, V., (2023). Large language models encode clinical knowledge. Nature, 620, 172–180.

135. Madani, A., Krause, B., Greene, E. R., Subramanian, S., Mohr, B. P., Holton, J. M., Olmos, J. L., Xiong, C., Sun, Z. Z., Socher, R., Fraser, J. S. and Naik, N., (2023). Large language models generate functional protein sequences across diverse families. Nat. Biotechnol., 41, 1099–1106.

136. Raissi, M., Perdikaris, P. and Karniadakis, G. E., (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys., 378, 686–707.

137. Sliwoski, G., Kothiwale, S., Meiler, J. and Lowe, E. W., (2014). Computational methods in drug discovery. Pharmacol. Rev., 66, 334–395.

138. Li, B., Wei, Z., Wu, J., Yu, S., Zhang, T., Zhu, C., Zheng, D., Guo, W., Zhao, C. and Zhang, J., (2023). Machine learning-enabled globally guaranteed evolutionary computation. Nat. Mach. Intell., 5, 457–467.

139. Elkhashap, A. and Abel, D., (2022). Reduction and observer design for a grey-box model in continuous pharmaceutical manufacturing. IFAC-PapersOnLine, 55, 26–32.

140. van Hauwermeiren, D., Stock, M., Beer, T. de and Nopens, I., (2020). Predicting pharmaceutical particle size distributions using kernel mean embedding. Pharmaceutics, 12, 12030271.

141. Unnikrishnan, S., Donovan, J., Macpherson, R. and Tormey, D., (2020). Machine learning for automated quality evaluation in pharmaceutical manufacturing of emulsions. J. Pharm. Innov., 15, 392–403.

142. Zaki, M. J., (2020). Data mining and machine learning: Fundamental concepts and algorithms, ISBN 9781108473989.

143. James, G., Witten, D., Hastie, T., Tibshirani, R. and Taylor, J., (2017). An introduction to statistical learning, 8th ed., ISBN 9781461471370.

144. Minsky, M. and Papert, S. A., (1969). Perceptrons: An introduction to computational geometry, 1st ed., ISBN 9780262343930.

145. Krizhevsky, A., Sutskever, I. and Hinton, G. E., (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, pp 1097–1105.

146. Kůrková, V., (1992). Kolmogorov's theorem and multilayer neural networks. Neural Netw., 3, 501–506.

147. Derry, A., Krzywinski, M. and Altman, N., (2023). Convolutional neural networks. Nat. Methods, 20, 1269–1270.

148. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., (2014). Generative Adversarial Nets. Advances in Neural Information Processing Systems, pp 2672–2680.

149. Rumelhart, D. E. and McClelland, J. L., (1987). Parallel distributed processing: Explorations in the microstructure of cognition: Foundations, ISBN 9780262291408.

150. Che, Z., Purushotham, S., Cho, K., Sontag, D. and Liu, Y., (2018). Recurrent neural networks for multivariate time series with missing values. Sci. Rep., 8, 6085.

151. Pires, D., Blundell, T. L. and Ascher, D. B., (2015). pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. J. Med. Chem., 58, 4066–4072.

152. Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M. and Zhao, S., (2019). Applications of machine learning in drug discovery and development. Nat. Rev. Drug Discov., 18, 463–477.

153. Chen, H., Engkvist, O., Wang, Y., Olivecrona, M. and Blaschke, T., (2018). The rise of deep learning in drug discovery. Drug Discov. Today, 23, 1241–1250.

154. Yang, Y., Ye, Z., Su, Y., Zhao, Q., Li, X. and Ouyang, D., (2019). Deep learning for in vitro prediction of pharmaceutical formulations. Acta Pharm. Sin. B., 9, 177–185.

155. Ibrić, S., Djuriš, J., Parojčić, J. and Djurić, Z., (2012). Artificial neural networks in evaluation and optimization of modified release solid dosage forms. Pharmaceutics, 4, 531–550.

156. Han, R., Yang, Y., Li, X. and Ouyang, D., (2018). Predicting oral disintegrating tablet formulations by neural network techniques. Asian J. Pharm. Sci., 13, 336–342.

157. O'Reilly, C. S., Elbadawi, M., Desai, N., Gaisford, S., Basit, A. W. and Orlu, M., (2021). Machine learning and machine vision accelerate 3d printed orodispersible film development. Pharmaceutics, 13, 13122187.

158. Elbadawi, M., Muñiz Castro, B., Gavins, F., Ong, J. J., Gaisford, S., Pérez, G., Basit, A. W., Cabalar, P. and Goyanes, A., (2020). M3DISEEN: A novel machine learning approach for predicting the 3D printability of medicines. Int. J. Pharm., 590, 119837.

159. Muñiz Castro, B., Elbadawi, M., Ong, J. J., Pollard, T., Song, Z., Gaisford, S., Pérez, G., Basit, A. W., Cabalar, P. and Goyanes, A., (2021). Machine learning predicts 3D printing performance of over 900 drug delivery systems. J. Control. Release, 337, 530–545.

160. Elbadawi, M., McCoubrey, L. E., Gavins, F., Ong, J. J., Goyanes, A., Gaisford, S. and Basit, A. W., (2021). Harnessing artificial intelligence for the next generation of 3D printed medicines. Adv. Drug Deliv. Rev., 175, 113805.

161. He, Y., Ye, Z., Liu, X., Wei, Z., Qiu, F., Li, H.-F., Zheng, Y. and Ouyang, D., (2020). Can machine learning predict drug nanocrystals? J. Control. Release, 322, 274–285.

162. Noorain, L., Nguyen, V., Kim, H.-W. and Nguyen, L., (2023). A machine learning approach for PLGA nanoparticles in antiviral drug delivery. Pharmaceutics, 15, 15020495.

163. Jin, I. S., Yoon, M. S., Park, C.-W., Hong, J. T., Chung, Y. B., Kim, J.-S. and Shin, D. H., (2020). Replacement techniques to reduce animal experiments in drug and nanoparticle development. J. Pharm. Investig., 50, 327–335.

164. Bennett-Lenane, H., Griffin, B. T. and O'Shea, J. P., (2022). Machine learning methods for prediction of food effects on bioavailability: A comparison of support vector machines and artificial neural networks. Eur. J. Pharm. Sci., 168, 106018.

165. Paul, S., Baranwal, Y. and Tseng, Y.-C., (2021). An insight into predictive parameters of tablet capping by machine learning and multivariate tools. Int. J. Pharm., 599, 120439.

166. Akseli, I., Xie, J., Schultz, L., Ladyzhynsky, N., Bramante, T., He, X., Deanne, R., Horspool, K. R. and Schwabe, R., (2017). A practical framework toward prediction of breaking force and disintegration of tablet formulations using machine learning tools. J. Pharm. Sci., 106, 234–247.

167. Hansen, L. K. and Salamon, P., (1990). Neural Network Ensembles. IEEE Trans. Pattern Anal. Mach. Intell., 10, 993–1001.

168. Bottou, L., (2010). Large-scale machine learning with stochastic gradient descent. Proceedings of COMPSTAT'2010: 19th International Conference on Computational Statistics; Lechevallier, Y., Saporta, G., Eds., pp 177–186.

169. Hagan, M. T. and Menhaj, M. B., (1994). Training feedforward networks with the Marquardt algorithm. IEEE Trans. Neural Netw., 5, 989–993.

170. Fukushima, K., (1969). Visual feature extraction by a multilayered network of analog threshold elements. IEEE Transactions on Systems Science and Cybernetics, 4, 322–333.

171. Lu, L., Shin, Y., Su, Y. and Karniadakis, G. E., (2020). Dying ReLU and initialization: Theory and numerical examples. Commun. Comput. Phys., 28, 1671–1706.

172. Xu, J.; Li, Z.; Du, B.; Zhang, M.; Liu, J., Eds., (2020. Reluplex made more practical: Leaky ReLU. Proceedings - IEEE Symposium on Computers and Communications.

173. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R., (2014). Dropout: A simple way to prevent neural networks from overtting. J. Mach. Learn. Res., 15, 1929–1958.

174. Kingma, D.P.; Ba, J.L., Eds., (2015. Adam: A method for stochastic optimization. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.

175. F. Kamalov, L. Smail and I. Gurrib, (2020). Stock price forecast with deep learning. International Conference on Decision Aid Sciences and Application (DASA), pp 1098–1102.

176. Kempaiah, B. U., Mampilli, R. J. and Goutham, K. S., (2022). A deep learning approach for speed bump and pothole detection using sensor data. Lecture Notes in Electrical Engineering, pp 73–85.

177. Manohar, B. and Das, R., (2022). Artificial neural networks for the prediction of monkeypox outbreak. Tropical Medicine and Infectious Disease, 7, 424.

178. Pascanu, R., Mikolov, T. and Bengio, Y., (2013). On the difficulty of training recurrent neural networks. 30th International Conference on Machine Learning, pp 2347–2355.

179. Hochreiter, S., (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. Int. J. Uncertain. Fuzziness Knowl.-Based Syst., 6, 107–116.

180. Hochreiter, S. and Schmidhuber, J., (1997). Long short-term Memory. Neural Comput., 9, 1735–1780.

181. Fischer, T. and Krauss, C., (2018). Deep learning with long short-term memory networks for financial market predictions. Eur. J. Oper. Res., 270, 654–669.

182. Ma, X., Tao, Z., Wang, Y., Yu, H. and Wang, Y., (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transp. Res. Part C Emerg. Technol., 54, 187–197.

183. Giambattista, P., Heikki, H. and Tuomas, V., (2016). Recurrent neural networks for polyphonic sound event detection in real life recordings. IEEE International Conference on Acoustics, Speech, and Signal Processing, pp 6440–6444, ISBN 9781479999880.

184. Zhang, X.-Y., Yin, F., Zhang, Y.-M., Liu, C.-L. and Bengio, Y., (2018). Drawing and recognizing chinese characters with recurrent neural network. IEEE Trans. Pattern Anal. Mach. Intell., 40, 849–862.

185. Sundermeyer, M., Schlüter, R. and Ney, H., (2012). LSTM neural networks for language modeling. Interspeech, pp 194–197.

186. Tang, A., (2023). Machine learning for pharmacokinetic/pharmacodynamic modeling. J. Pharm. Sci., 112, 1460–1475.

187. Yijie, G., Muzzio, F. and Ierapetritou, M. G., (2011). Characterization of feeder effects on continuous solid mixing using fourier series analysis. AIChE J., 57, 1144–1153.

188. Engisch, W. and Muzzio, F., (2016). Using residence time distributions (RTDs) to address the traceability of raw materials in continuous pharmaceutical manufacturing. J. Pharm. Innov., 1, 64–81.

189. Kiricenko, K. and Kleinebudde, P., (2023). Drying behavior of a horizontal vibrated fluidized bed dryer for continuous manufacturing. Pharm. Dev. Technol., 5, 440–451.

190. Mélo, T. J. A. and Canevarolo, S. V., (2002). An optical device to measure in-line residence time distribution curves during extrusion. Polym Eng Sci, 1, 170–181.

191. Reitz, E., Podhaisky, H., Ely, D. and Thommes, M., (2013). Residence time modeling of hot melt extrusion processes. Eur. J. Pharm. Biopharm., 3, Part B, 1200–1205.

192. Furukawa, R., Singh, R. and Ierapetritou, M., (2020). Effect of material properties on the residence time distribution (RTD) of a tablet press feed frame. Int. J. Pharm., 119961.

193. Meier, R., Harting, J., Happel, J. and Kleinebudde, P., (2017). Implementation of microwave sensors in continuous powder feeding – A novel tool to bridge refill phases. Pharm. Ind., 79, 576–582.

194. Brouckaert, D., Vandenbroucke, F., Chauchard, F., Dollinger, M., Roggo, Y., Pellegatti, L. and Krumme, M., (2022). Towards real-time release of pharmaceutical tablets: 100% in-line control via near-infrared spatially resolved spectroscopy and 3D microwave resonance technology. J. Pharm. Biomed. Anal., 209, 114491.

195. Malmberg, C. G. and Maryott, A. A., (1956). Dielectric constant of water from 0 to 100 C. Journal of research of the National Bureau of Standards, 56, 1–8.

196. Hasegawa, S. and Stokesberry, D. P., (1975). Automatic digital microwave hygrometer. Rev. Sci. Instrum., 7, 867–873.

197. Shannon, C. E., (1948). A mathematical theory of communication. Bell Labs Tech. J., 27, 379–423.

198. Pincus, S. M., (1991). Approximate entropy as a measure of system complexity. Proc. Natl. Acad. Sci. USA, 88, 2297–2301.

199. Renyi, A., (1961). On measures of entropy and information. Berkeley Symp. on Math. Statist. and Prob., 547–561.

200. Powell, G. E. and Percival, I. C., (1979). A spectral entropy method for distinguishing regular and irregular motion of Hamiltonian systems. J. Phys. A: Math. Gen., 12, 2053–2071.

201. Haware, R. V., Tho, I. and Bauer-Brandl, A., (2009). Application of multivariate methods to compression behavior evaluation of directly compressible materials. Eur. J. Pharm. Biopharm., 1, 148–155.

202. Garekani, H. A., Ford, J. L., Rubinstein, M. H. and Rajabi-Siahboomi, A. R., (2001). Effect of compression force, compression speed, and particle size on the compression properties of paracetamol. Drug. Dev. Ind. Pharm., 27, 935–942.

203. Garr, J. and Rubinstein, M. H., (1991). An investigation into the capping of paracetamol at increasing speeds of compression. Int. J. Pharm., 72, 117–122.

204. Berkenkemper, S., Klinken, S. and Kleinebudde, P., (2023). Multivariate data analysis to evaluate commonly used compression descriptors. Int. J. Pharm., 637, 122890.

205. Berkenkemper, S., Klinken, S. and Kleinebudde, P., (2023). Investigating compressibility descriptors for binary mixtures of different deformation behavior. Powder Technol., 424, 118571.

206. Dai, S., Xu, B., Zhang, Z., Yu, J., Wang, F., Shi, X. and Qiao, Y., (2019). A compression behavior classification system of pharmaceutical powders for accelerating direct compression tablet formulation design. Int. J. Pharm., 572, 118742.

207. Casian, T., Nagy, B., Lazurca, C., Marcu, V., Tőkés, E. O., Kelemen, É. K., Zöldi, K., Oprean, R., Nagy, Z. K., Tomuta, I. and Kovács, B., (2023). Development of a PAT platform for the prediction of granule tableting properties. Int. J. Pharm., 648, 123610.

208. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. and Polosukhin, I., (2017). Attention is all you need. Advances in Neural Information Processing Systems. NIPS, pp 5999–6009, ISBN 9781510860964.

209. Lea, C., Vidal, R., Reiter, A. and Hager, G. D., (2016). Temporal convolutional networks: A unified approach to action segmentation. Computer Vision – ECCV 2016 Workshops, pp 47–54, ISBN 978-3-319-49409-8.

10 Appendix

10.1 Sums of amplitudes code snippet

10.1.1 Sums of amplitudes calculation

```
def FibonacciNumberCalculation(X):
    sqrt_of_five = np.sqrt(5)
    alpha = (1 + sqrt_of_five) / 2
    beta = (1 - sqrt_of_five) / 2
    fibonacci_numbers = np.rint((alpha ** X) - (beta ** X)) / sqrt_of_five
    return np.abs(np.around(fibonacci_numbers, decimals=0)).astype(int)


def SumsOfAmplitudesAnalysisStepPermutation(Time, Y_Values):
    MinimumSpectraLength = 10
    RandomCurvesIteration = 25
    sample_rate = 1 / Time.diff().mean()
    seeds = FibonacciNumberCalculation(np.arange(2, (RandomCurvesIteration + 1)))
    Y_Values_normalized = Y_Values - np.mean(Y_Values)
    window_sizes = np.arange((MinimumSpectraLength - 1), np.size(Time, 0))
    last_value_in_SAA = np.array(Time.iloc[(MinimumSpectraLength - 1):])
    SAA_permutation = np.empty([np.size(window_sizes), np.size(seeds)])
    SAA_data = np.empty([np.size(window_sizes)])
    for i in trange(len(seeds)):
        seed = seeds[i]
        index1 = np.where(seeds == seed)[0]
        np.random.seed(seed)
        for j in window_sizes:
            index2 = np.where(window_sizes == j)[0]
            permuted_Y_Values = np.random.permutation(Y_Values_normalized[:j])
            amplitudes_permutation = 2.0 / j * np.abs(fft(permuted_Y_Values)[0:j // 2])
            amplitudes_data = 2.0 / j * np.abs(fft(np.array(Y_Values_normalized)[:j])[0:j // 2])
            SAA_permutation[index2, index1] = amplitudes_permutation.sum()
            SAA_data[index2] = amplitudes_data.sum()
    SAA_permutation = SAA_permutation.mean(axis=1)
    SAA_data = SAA_data / sample_rate
    SAA_permutation = SAA_permutation / sample_rate
    SAR = SAA_data / SAA_permutation
    SAA_output = pd.DataFrame({'window size': window_sizes + 1,
                               'Process time / s': last_value_in_SAA,
```

```
                'data SAA': SAA_data,
                'permutation SAA': SAA_permutation,
                'SAR': SAR})
    return SAA_output
```

10.2 EvTab 2.0 code snippets

10.2.1 Change point detection for the compression phase in EvTab code

```
def separate_compression_cycle(X):
    noise_start = X.shape[0] - next(x for x, val in enumerate(X.iloc[::-1, 1]) if val < 0)
    noise_mean = np.nanmean(X.iloc[noise_start:, 3])
    noise_std = np.nanstd(X.iloc[noise_start:, 3], ddof=1)
    noise = noise_mean + noise_std
    index_maxmimum = X.iloc[:, 3].idxmax()
    start_intern = index_maxmimum - next(x for x, val in enumerate(X.iloc[index_maxmimum::-1, 3]) if val < noise)
    stop_intern = index_maxmimum + next(x for x, val in enumerate(X.iloc[index_maxmimum:, 1]) if val >= 0)
    return X.iloc[start_intern:stop_intern, :]
```

10.2.2 Algorithm for the Heckel equation regression

```
def HeckelCalculation():
    def CheckErrorsInDensity(X1):
        if (X1 > true_density).any():
            list_indexes = [x for x, val in enumerate(X1) if val - true_density >= 0]
            in_die_density.drop(list_indexes, axis=0, inplace=True)
            pressure.drop(list_indexes, axis=0, inplace=True)
            return True
        else:
            return False

    X.reset_index(inplace=True, drop=True)
    pressure = X['Upper punch force']
    in_die_density = float(df_test.iloc[i_intern, -1]) / (
            X['Distance between punches'] * square_of_die_radius * np.pi)
    sigmoid_master_intern = CheckErrorsInDensity(in_die_density)
    deg_densification = in_die_density / true_density
    ln_rec_epsilon = np.log(1 / (1 - deg_densification))
    ln_rec_epsilon = ln_rec_epsilon.loc[~(pressure < 1)]
    pressure = pressure.loc[~(pressure < 1)]
    return pd.DataFrame({'Pressure': pressure,
                         'ln(1/e)': ln_rec_epsilon}), sigmoid_master_intern

def HeckelFit():
    df_data_heckel = df_data_intern.copy()
    df_data_heckel = df_data_heckel.iloc[:df_data_heckel.dropna().iloc[:, 0].idxmax(), :]
    df_data_heckel.reset_index(inplace=True, drop=True)
```

```
    interpolation_steps = 10000
    interpolation = interp1d(df_data_heckel.iloc[:, 0],
                             df_data_heckel.iloc[:, 1],
                             kind='linear')
    x_uniform = np.linspace(df_data_heckel.iloc[:, 0].min(),
                            df_data_heckel.iloc[:, 0].max(),
                            interpolation_steps)
    df_data_heckel = pd.DataFrame({'x': x_uniform,
                                   'y': interpolation(x_uniform)}).to_numpy()


    window_length = int(interpolation_steps // 3)
    if (window_length % 2) == 0:
        window_length += 1
    df_data_heckel = np.append(df_data_heckel,
                               scipy.signal.savgol_filter(df_data_heckel[:, 1],
                                                          window_length=window_length,
                                                          polyorder=2)[:, None],
                               axis=1)
    first_derivative = np.round(np.diff(df_data_heckel[:, 2], n=1) / np.diff(df_data_heckel[:, 0], n=1), 3)
    mode_first_derivative = scipy.stats.mode(first_derivative)[0]
    index_start = next(x for x, val in enumerate(first_derivative) if val == mode_first_derivative)
    if index_start == 0:
        index_start = 1
    index_end = first_derivative.shape[0] - next(x for x, val in enumerate(first_derivative[::-1]) if val == mode_first_derivative)
    if index_end == first_derivative.shape[0] - 1:
        index_end = first_derivative.shape[0] - 2
    Reg_K = linregress(df_data_heckel[index_start:index_end, 0],
                       df_data_heckel[index_start:index_end, 1])
    return Reg_K, df_data_heckel[index_start:index_end, 0]
```

10.2.3 Algorithm for force-displacement diagrams

```
def CalculateForces():
    pressure_upper = X['Upper punch force']
    force_upper_inner = pressure_upper * square_of_die_radius * np.pi
    pressure_lower = X['Lower punch force']
    force_lower_inner = pressure_lower * square_of_die_radius * np.pi


    return force_upper_inner, force_lower_inner
def DetectStart():
```

```
    return next(x for x, val in enumerate(X['Upper punch force']) if val > 1)
def DetectEnd(force_intern):
    try:
        return next(x for x, val in enumerate(force_intern) if val <= 0)
    except StopIteration:
        return force_intern.shape[0] - 1
def E_1_Integration(X_int, Y_int):
    integral = 0
    for i_int in range(X_int.shape[0]):
        if i_int == 0:
            integral += Y_int[i_int] * X_int[i_int] / 2
        else:
            integral += (Y_int[i_int] + Y_int[i_int - 1]) * (X_int[i_int] - X_int[i_int - 1]) / 2
    return integral
def E_2_Integration(X_int, Y_int):
    integral = 0
    for i_int in range(1, X_int.shape[0]):
        integral += (Y_int[i_int] + Y_int[i_int - 1]) * (X_int[i_int] - X_int[i_int - 1]) / 2
    return abs(integral)
def CalculateEnergies(force, dist):
    force.reset_index(inplace=True, drop=True)
    dist.reset_index(inplace=True, drop=True)
    compression_force = force[:dist.idxmax()]
    compression_distance = dist[:dist.idxmax()]
    elastic_force = force[dist.idxmax():]
    elastic_distance = dist[dist.idxmax():]
    elastic_force.reset_index(inplace=True, drop=True)
    elastic_distance.reset_index(inplace=True, drop=True)
    end = DetectEnd(elastic_force)
    elastic_force = elastic_force[:end]
    elastic_distance = elastic_distance[:end]
    E_ges = force.max() * dist[force.idxmax()] / 2
    E_3 = E_2_Integration(elastic_distance, elastic_force)
    E_2 = E_1_Integration(compression_distance, compression_force) - E_3
    E_1 = E_ges - E_3 - E_2
    df_result_intern = pd.concat([compression_distance, compression_force], axis=1)
    df_result_intern = pd.concat([df_result_intern, pd.concat([elastic_distance, elastic_force], axis=1)], axis=0)
    return [E_1, E_2, E_3], df_result_intern
```

10.2.4 Algorithm for the Walker equation regression

```
def WalkerFit(x_main, y_main):
    interpolation_steps = 10000
    interpolation = interp1d(x_main,
                             y_main,
                             kind='linear')
    x = np.linspace(np.nanmin(x_main),
                    np.nanmax(x_main),
                    interpolation_steps)
    y = interpolation(x)
    window_length = int(y.shape[0] // 3)
    if (window_length % 2) == 0:
        window_length += 1
    y_smoothed = scipy.signal.savgol_filter(y,
                                            window_length=window_length,
                                            polyorder=3)
    if np.nanmax(x) <= 1 or list_materials[selector] == 'HPC':
        start_intern = np.where(x > 0)[0][0]
    elif np.nanmax(x) > 1:
        start_intern = np.where(x > 1)[0][0]
    stop_intern = np.argmax(x)
    x = x[start_intern:stop_intern]
    y_smoothed = y_smoothed[start_intern:stop_intern]
    y = y[start_intern:stop_intern]
    first_derivative = np.round(np.diff(y_smoothed, n=1) / np.diff(x, n=1), 1)
    mode_first_derivative = scipy.stats.mode(first_derivative)[0]
    index_start = next(x for x, val in enumerate(first_derivative) if val <= mode_first_derivative)
    if index_start == 0:
        index_start = 1
    index_end = first_derivative.shape[0] - next(x for x, val in enumerate(first_derivative[::-1]) if val <= mode_first_derivative)
    if index_end == first_derivative.shape[0] - 1:
        index_end = first_derivative.shape[0] - 2
    Reg_K = linregress(x[index_start:index_end],
                       y[index_start:index_end])
    return Reg_K, x[index_start:index_end]
```

10.2.5 Algorithm for the Kawakita equation regression

```
def FittingLoop(pressure, P_C):
    pressure_limit_1 = 10
    interpolation = interp1d(pressure,
```

```
                    P_C,
                    kind='linear')
pressure = np.linspace(pressure.min(),
                    pressure.max(),
                    250)
P_C = interpolation(pressure)
try:
    start = next(x for x, val in enumerate(pressure) if val >= pressure_limit_1)
    end = np.argmax(pressure)
except:
    start = next(x for x, val in enumerate(pressure) if val >= 1)
    end = np.argmax(pressure)

while True:
    X1 = pressure[start:end]
    Y1 = P_C[start:end]
    n = end - start
    r_sq = 1 - ((n - 1) / (n - 2)) * (1 - linregress(X1, Y1).rvalue ** 2)
    if X1.shape[0] <= 11:
        return X1, Y1
    if r_sq >= 0.999:
        return X1, Y1
    list_optimum = [r_sq]
    try:
        start_1 = start + 1
        if start_1 < end:
            X1 = pressure[start_1:end]
            Y1 = P_C[start_1:end]
            n = end - start_1
            list_optimum.append(((n - 1) / (n - 2)) / (1 - linregress(X1, Y1).rvalue ** 2))
        else:
            list_optimum.append(0)
    except IndexError:
        list_optimum.append(0)
    try:
        start_1 = start - 1
        if start_1 >= 0:
            X1 = pressure[start_1:end]
            Y1 = P_C[start_1:end]
            n = end - start_1
```

```
            list_optimum.append(((n - 1) / (n - 2)) / (1 - linregress(X1, Y1).rvalue ** 2))
        else:
            list_optimum.append(0)
    except IndexError:
        list_optimum.append(0)
    try:
        end_1 = end + 1
        if end_1 <= X1.shape[0]:
            X1 = pressure[start:end_1]
            Y1 = P_C[start:end_1]
            n = end_1 - start
            list_optimum.append(((n - 1) / (n - 2)) / (1 - linregress(X1, Y1).rvalue ** 2))
        else:
            list_optimum.append(0)
    except IndexError:
        list_optimum.append(0)
    try:
        end_1 = end - 1
        if end_1 > start:
            X1 = pressure[start:end_1]
            Y1 = P_C[start:end_1]
            n = end_1 - start
            list_optimum.append(((n - 1) / (n - 2)) / (1 - linregress(X1, Y1).rvalue ** 2))
        else:
            list_optimum.append(0)
    except IndexError:
        list_optimum.append(0)

    if list_optimum[0] == max(list_optimum):
        X1 = pressure[start:end]
        Y1 = P_C[start:end]
        return X1, Y1
    if list_optimum[1] == max(list_optimum):
        start += 1
    if list_optimum[2] == max(list_optimum):
        start -= 1
    if list_optimum[3] == max(list_optimum):
        end += 1
    if list_optimum[4] == max(list_optimum):
        end -= 1
```

10.3 LSTM-based neural network code snippets

10.3.1 Change point detection for the compression phase in LSTM code

```
def separate_compression(X):
    mode_pressure = sc.stats.mode(X.iloc[:, 3].round(1))[0][0]
    X.iloc[:, 0] *= 1e-3
    x_argmax = np.argmax(X.iloc[:, 3])
    idx_start = np.where(X.iloc[:x_argmax, 3].round(1) <= mode_pressure)[0][-1]
    idx_end = np.where(X.iloc[x_argmax:, 3].round(1) <= mode_pressure)[0][0] + x_argmax
    return X.iloc[idx_start:idx_end, :]
```

10.3.2 LSTM model architecture

```
def BuildTensorflowModel(list_inputs, list_outputs):
    input_layer_compression = tf.keras.layers.Input(shape=list_inputs[0].shape[1:])
    masked_compression = tf.keras.layers.Masking(mask_value=0.0)(input_layer_compression)
    x = tf.keras.layers.LSTM(units=64,  activation='tanh', recurrent_activation='sigmoid')(masked_compression)
    x = tf.keras.layers.Dense(256, activation=keras.layers.LeakyReLU(alpha=0.01))(x)
    x = tf.keras.layers.Dropout(0.2)(x)
    x = tf.keras.layers.Dense(32, activation=keras.layers.LeakyReLU(alpha=0.01))(x)
    output = tf.keras.layers.Dense(1, activation=None)(x)
    return tf.keras.Model(inputs=input_layer_compression, outputs=output)
```

10.3.3 LSTM model training

```
X_train, X_test, y_train, y_test = train_test_split(data_compression, info['Tensile Strength'], train_size=0.3, shuffle=True)
tf_model = BuildTensorflowModel([X_train], 1)
tf_model.compile(loss='mean_squared_error', optimizer='Adam', metrics='accuracy')
tf_model.fit(x=X_train, y=y_train, epochs=100, batch_size=20, verbose=2, use_multiprocessing=True, validation_split=0.2)
```

Eidesstattliche Versicherung

Ich versichere an Eides statt, dass die vorliegende Dissertation von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität in Düsseldorf“ verfasst worden ist.

_____

Stefan Klinken

Danksagung

Ein besonderer Wegweiser in jeder Promotionsarbeit ist sicherlich der Doktorvater oder die Doktormutter. Jörg, ich danke dir, dass du mir die Möglichkeit gegeben hast, an diesem Institut zu forschen, neue Einsichten zu erlangen und meine Dissertationsschrift zu verfassen. Ich habe sehr viel von dir gelernt und freue mich darauf, das auch in der Zukunft noch zu dürfen. Deinem Vertrauen in mich, in Bezug auf die Ermöglichung des DFG-Projektes, und deiner Geduld während meines Schreibens der Thesis, danke ich besonders. Mein Dank gilt Herrn Professor Kleinebudde für die Verfassung des zweiten Gutachtens zu dieser Arbeit. Ich danke ihm zudem für die vielen Denkanstöße in Bezug auf meine Arbeiten und meine statistischen Ausflüge. Sowohl Jörg als auch Herrn Kleinebudde möchte ich für die Entzündung meines wissenschaftlichen Feuers danken.

Julian, obwohl wir sicherlich einige Reibungspunkte hatten, möchte ich dir danken, dass du mir in meinem Drang, Neues zu tun, keine zu engen Grenzen gesetzt hast und meine Entfaltungen, auch die absolut sinnlosen Ansätze, in gewisser stoischer Gelassenheit betrachtet hast. Auch von dir und durch dich konnte ich sehr viel lernen.

Ich danke der Kooperation mit Sabrina beim Aufbau der Datenbank. Die gemeinsame Arbeit hat mir sehr darin geholfen, meinen Forschungsschwerpunkt zu finden. Unser beider Dissertationen haben hierdurch enorm profitiert.

Verbunden in tiefer Freundschaft danke ich Valentinë und Marcel. Ihr habt die natürlichen Sorgen der Promotion und des Lebens gemildert und in vielem, arbeitsbezogen und privat, in unvergleichbarer Weise Unterstützung geleistet. Ich danke euch sehr und werde dies alles stets in besonderer Erinnerung halten. Leonie, dir danke ich in Liebe. Deine Hilfe auf den letzten Metern, deine Geduld, deine Zusprache, alles, was unser ist und was noch kommen wird, erfüllt mich mit großer Dankbarkeit. Du hast zu dieser Arbeit beigetragen, und auch du wirst die deine mit besonderem Erfolg beenden. Ard, deiner Freundschaft, absoluten Aufrichtigkeit, Geradlinigkeit und der Hilfe in vielen Dingen danke ich sehr. Du warst sicherlich ein Vorbild in einigen Aspekten meiner Promotion. Philipp und Olga haben meinen Dank für die Freundschaft, welche uns verbindet. Olga, man kann sagen, dass durch dich und die E-Tongue meine erste Berührung mit multivariater Analyse zustande kam. Danke auch dafür. Raphael trug viel zum Knospen meiner Faszination für Daten bei. Laura möchte ich danken für die kurzen, aber heiteren Zeiträume an unserem Institut und auf den gemeinsamen Dienstreisen.

Ich danke meinen Freunden und Kaligatoren Chris, Frede, Kim, Micha und Sixt. Niemals vergesse ich die sich wiegende Krone der Buche über unseren Köpfen und die Zeit, die uns immer verbinden wird. Habt Dank für jedes Licht in schwieriger Stunde.

Michael, ich freue mich in jedem kommenden Semester auf unsere fünf gemeinsamen Plaudertermine mit Publikum. Der kurzen Zeit mit Björn möchte ich danken. Auch wenn diese vorliegende Arbeit nicht von dir beeinflusst ist, so leistest du mir doch viel Unterstützung in spektroskopischen und weiteren Aspekten.

Ich bin dankbar, als Doktorandensprecher des Institutes in einer schwierigen Phase tätig gewesen zu sein. Diese Erfahrung erteilte mir Lektionen, die ich nicht missen möchte. Ebenso erfüllt es mich mit Freude, dass ich die Möglichkeit hatte und haben, werde Studierenden meine wissenschaftliche Passion näherzubringen. Allen Projektpartnern danke ich für die Kooperationen, den erfolgreichen und den gescheiterten. Ich freue mich, vielen Kommilitonen Unterstützung leisten gekonnt zu haben und über die Früchte, die geerntet wurden. Arne, mein Dank gebührt dir, auch wenn ich nicht immer mit deiner bedingungslosen, rücksichtsvollen Art d'accord ging. Herausheben möchte ich Jill, welche mit besonderem Interesse ein Thema, welches andere in die Flucht geschlagen hätte, auf sich genommen hat. Zudem möchte ich meinen WPPlern und Nazgol danken.

Meine Eltern sowie Christina, Cäcilia und Monika haben mir mehr als nur die grundlegenden Werkzeuge des Lebens mit auf den Weg gegeben. Ohne diese hätte ich ihn nicht gehen können. Mein Dank und meine Liebe sind bei euch.