Pharmaceutical application of multivariate modelling techniques: a review on the manufacturing of tablets

Guolin Shi, † Longfei Lin, † Yuling Liu, Gongsen Chen, Yuting Luo, Yanqiu Wu and Hui Li*

The tablet manufacturing process is a complex system, especially in continuous manufacturing (CM). It includes multiple unit operations, such as mixing, granulation, and tableting. In tablet manufacturing, critical quality attributes are influenced by multiple factorial relationships between material properties, process variables, and interactions. Moreover, the variation in raw material attributes and manufacturing processes is an inherent characteristic and seriously affects the quality of pharmaceutical products. To deepen our understanding of the tablet manufacturing process, multivariable modeling techniques can replace univariate analysis to investigate tablet manufacturing. In this review, the roles of the most prominent multivariate modeling techniques in the tablet manufacturing process are discussed. The review mainly focuses on applying multivariate modeling techniques to process understanding, optimization, process monitoring, and process control within multiple unit operations. To minimize the errors in the process of modeling, good modeling practice (GMoP) was introduced into the pharmaceutical process. Furthermore, current progress in the continuous manufacturing of tablets and the role of multivariate modeling techniques in continuous manufacturing are introduced. In this review, information is provided to both researchers and manufacturers to improve tablet quality.

1. Introduction

Tablets are the prevalent dosage forms, which have many advantages, including high-precision dosing, good physical and chemical stability, manufacturing efficiency, and low costs.¹ In the global pharmaceutical market, it has been speculated that tablets are the preferred dosage form and tablet sales are expected to exceed US$500 billion by the end of 2027.²

Due to the enormous market share of tablets, higher requirements have been put forward to improve product quality and supervision. As mentioned by Yu and Kopcha, there is still a long way to go to improve the quality of pharmaceutical products.³ One of the major challenges is determining the relationship between raw material attributes, process conditions, and critical quality attributes.⁴ The tablet manufacturing process involves powder blending, granulation, tablet compression, and coating operation units. This process makes it difficult for univariate methods to comprehensively study the tablet manufacturing process.

Mathematical models, which could provide a scientific understanding of the manufacturing process and predict the state of the pharmaceutical system, play an important role in the tablet manufacturing process. According to an investigational report by Kourti and Davis in 2012,⁵ the application of mathematical modeling can achieve lots of benefits, including improving product quality, and enhancing product and process understanding. For mechanism models, the discrete element method (DEM)⁶ ⁷ and finite element method (FEM)⁸ ⁹ have been used less often than the multivariable modeling technique in tablet manufacturing, due to the complexity of raw material properties and the continuum duality of particles.¹⁰ For empirical models, multivariate modeling techniques are one of the most commonly used statistical models. In the pharmaceutical field, with the application of process analytical technology (PAT) and with the emergence of continuous manufacturing (CM), multivariate modeling techniques are feasible options for mathematically describing the pharmaceutical process.¹¹ For example, regression models and latent variable models (LVMs) can extract useful information related to product quality from a large amount of data.

In this paper, different multivariate model techniques are introduced for process understanding, optimization, quality monitoring, and control in the powder blending process, granulation, tableting and the coating process. To comprehensively investigate the roles of multivariate modeling techniques in the tablet manufacturing process, relevant studies published in the past 15–20 years (mostly within 15 years) were searched on the Web of Science, by using the following search
terms: QbD, empirical model, statistics model, pharmaceutical tablet, blending, granulation, coating, continuous manufacture, near-infrared (NIR), Raman, chemometrics, process model, multivariate data analysis, and partial least squares (PLS). The review is organized as follows: Section 2 briefly exhibits the theoretical background of commonly-used multivariate modeling techniques. Section 3 describes good modeling practice (GMoP) in the pharmaceutical process. Sections 4 and 5 provide an overview of multivariate modeling techniques in the basic unit operations and continuous manufacturing of tablets.

2. Brief theoretical background of commonly-used multivariate modeling techniques

In the pharmaceutical process, multivariate methods can be used to maximize speed and minimize costs. Multivariate analysis (MVA) models, also called data-driven models, are usually available for analyzing data obtained from the pharmaceutical process. In this review, we aimed to illustrate the applications of multivariate modeling techniques within the tablet manufacturing process, and, therefore, little effort has been expended on the theoretical aspects of the multivariate modeling techniques. Brief explanations of the theoretical background are given below.

2.1. Brief theory of simple regression

Various regression models can be used to investigate the relationships between independent and response variables. Multiple linear regression (MLR) and polynomial regression are often used to analyze data from the design of experiment (DoE). These models quantify the relationship between several explanatory variables and one or more response variables.

The MLR model is fitted by minimizing the sum of squared residuals based on least squares. According to the principle, the regression coefficient can be calculated with eqn (1)

$$B = (X^TX)^{-1}X^TY$$

In eqn (1), $Y$, $X$, $B$, and $X^T$ represent the response variable matrix, independent variables, regression coefficient matrix and transpose matrix of the independent variables, respectively. However, the inverse of $X^TX$ no longer exists due to collinearity between the independent variables. Except for cases where $X$ variables are controlled in designed experimentation, measured data in pharmaceutical applications are typically multivariate and collinear and MLR cannot be used. As a special linear regression model, polynomial regression describes the nonlinear relationship between independent and dependent variables in pharmaceutical manufacturing.

2.2. Basic theoretical concepts in latent variable modeling

The latent variable model (LVM) is a statistical model that can be used to process large amounts of correlated data. LVMs were used for experimental data as early as the 1970s. By extracting latent variables (LVs), an LVM projects original high-dimensional data into a low-dimensional space, also called the latent space. The new LVs still retain as much original information as possible. Next, in the latent space, the relationships among samples, independent variables, and response variables can be analyzed. Fig. 1 displays a straightforward interpretation of LVMs. For a detailed introduction to LVMs, refer to related studies.

In the last decade, principal component analysis (PCA), partial least squares (PLS) and orthogonal partial least squares (OPLS) were commonly used in the pharmaceutical manufacturing process. However, PCA and PLS models are not suitable for dealing with more complex datasets including more than two matrices. Multi-block PLS (MB-PLS), L-shape PLS and joint-Y PLS (JY-PLS) can analyze complex datasets, and quantify the relationship between input matrices and output matrices.

2.3. Basic theoretical concepts about artificial neural networks

The artificial neural network (ANN) model is a type of machine learning tool, which attempts to imitate the neural structures of the biological brain. In general, ANN models can be roughly divided into two types: static ANN and dynamic ANN. Although many neural networks have been continually invented, all ANN models can be described by neurons, also called processing elements (PE), learning algorithms, and connection formulas. In the neuronal system, a PE receives one or more signals, and input data is multiplied by its weight. Finally, an activation
function can generate an output.\textsuperscript{23} For the detailed principle of ANN, refer to previous publications.\textsuperscript{23–25} Fig. 2 shows a feedforward multi-layer perceptron (MLP) network, which is the most frequently and successfully used ANN model in pharmaceutical processes.\textsuperscript{21} MLP can develop a non-linear mapping between independent variables and response variables by using many parallel neurons.

However, building an ANN model involves lots of parameters and activation functions. Hence, it is not easy to obtain any insight into the structure of approximate functions. In other words, because of intermediate equations, the relationship between the input and output variables is not a direct path.\textsuperscript{26} Consequently, it is difficult to gain the mathematical equation between independent and response variables.\textsuperscript{21} Although the ANN model is a black-box model, it is a powerful tool to solve non-linear problems and multi-response systems. One of its main advantages is that it does not require a rule-based experimental design. Moreover, ANN can use historical data or incomplete data to map functions. Because the ANN model can be used to efficiently develop different response surfaces, it plays a crucial role in drug delivery and pharmaceutical processes.\textsuperscript{22}

3. Brief introduction to good modeling practice in the pharmaceutical process

To minimize errors in the process of modeling, or in analysis and subsequent applications,\textsuperscript{27} it is necessary to implement GMoP to build multivariate models.

Fig. 2 Structure of a multilayer artificial neural network in pharmaceutical manufacturing.

Fig. 3 Model development framework of pharmaceutical process.
According to the practice guidelines for building a model, the key components of GMoP should include model purpose, model evaluation, and performance measures.\textsuperscript{28,29} To standardize the modeling process in the GMoP, a model development framework is proposed (shown in Fig. \ref{fig:framework}). In the pharmaceutical process, the first step in developing the model process is to put forward the issues according to the actual situation and the modeling purpose. In the first step, it is important to choose an appropriate modeling technique. For the second step, one needs to acquire data according to the DoE or historical data in the manufacturing process. Subsequently, the data should be organized and preprocessed before model development. In this stage, it is suggested that exploratory analysis of the data should be implemented to select variables. Next, the model can be built, and the performance of the model can be evaluated. In this step, especially the development of multivariate models, the interpretability of the model needs to be considered. Finally, after validation, the developed model can be used in the actual manufacturing process.

In the GMoP, model verification and validation are essential parts of the modeling process before using the model for any pharmaceutical unit operation, and are directly related to the accuracy and applicability of the model. Consequently, it is necessary to perform model evaluation. For the model validation of the ANN model and LVMs, chemometric indicators, such as correlation coefficient and root mean square error (RMSE), are in general used to evaluate the model. To evaluate an MLR or a polynomial model, in addition to the above-mentioned indexes, the model should be further studied for statistical analysis, such as analysis of variance (ANOVA). The premise for further application of these developed models is that the $P$-value of the model is less than 0.05, and the $P$-value of the lack of fit is greater than 0.05. When the model meets these requirements, the relationship between response variables and independent variables is significant, and the equation is well fitted.

Model validation does not indicate model verification. Nevertheless, in the actual modeling process, model validation is always combined with model verification. The purpose of model validation is to prove that the model recreates the behavior of the system with high fidelity to meet the analysis goal. The commonly-used validation method compares the predictive model results to experimental data by using the internal or external data set. According to ICH Points to Consider (R2) guidelines,\textsuperscript{48} the model can be divided into three categories: high impact models, medium impact models, and low impact models. The developed model should be validated at different levels based on the category used in practice. For a high impact model, validation by the internal data set is often inadequate. It is necessary to use the external data set to validate model performance. In addition, data sets used for validation should take the expected variability into account in a future pharmaceutical process. Model accuracy, especially for statistically-based models, should be validated by uncertainty analysis.\textsuperscript{33,34} However, a low impact model may not implement rigorous verification. Finally, according to the pharmaceutical quality system of the company, the validated model can be used in the manufacturing process, and should be continuously updated and maintained.

4. Application of multivariate modeling techniques in tablet manufacturing

The tablet manufacturing process generally includes three pathways: \textit{i.e.}, direct compaction (DC), dry granulation, and wet granulation. This section will summarize and discuss the application of multivariate modeling techniques in the tablet manufacturing process.

4.1. Pharmaceutical process understanding and optimization based on multivariate modeling techniques

The main role of understanding the pharmaceutical process is identifying and controlling critical sources of variation that affect product quality. The objective is to understand how raw material properties and process variables can affect quality attributes. Based on process knowledge, pharmaceutical processes can be optimized.

4.1.1. Understanding and optimizing the blend process.

The blending unit operation is one of the basic processes for preparing solid dosage forms. In addition, it is a critical unit operation for ensuring that pharmaceutical powder mixing is homogeneous. Only homogeneous powders can ensure the content uniformity (CU) of active pharmaceutical ingredients (APIs).\textsuperscript{32,34} Blending uniformity (BU) is a critical quality attribute (CQA) of the intermediate and final product because it relates significantly to drug quality, safety, and therapeutic efficacy. However, during the blending process, BU is strongly impacted by several factors, including raw material attributes, mixing equipment, process parameters and environmental conditions (Fig. \ref{fig:bu}).\textsuperscript{35-41}

To explore the relationship between the above factors and BU, the effects of raw material properties and process parameters on the mixing process can be explored by multivariate modeling techniques. Table \ref{table:models} shows the application of multivariate models in the powder mixing process. It can be seen from Table \ref{table:models} that multiple regression models and LVMs were widely used. DoE-integrated multivariate modeling techniques were often utilized to quantify the impact of various factors on the mixing process, which helps to identify CMAs (critical material attributes) or CPPs (critical process parameters). Based on established models, a design space can be developed to control the mixing process, and this is beneficial for guiding and optimizing the mixing unit operation.

4.1.2. Understanding and optimizing the granulation process.

Granulation, as a process of powder agglomeration and size enlargement, is a key pharmaceutical unit operation. Prior to tablet compression, granulation can significantly improve the physical properties of the powder, including flowability and density.\textsuperscript{49-52} To obtain optimal characteristics of granules, statistical or mechanistic models have been used to study the physical properties of granules produced by various granulation methods. In this section, the application of multivariate modeling techniques to commonly-used granulation methods is summarized, such as high shear wet granulation (HSWG), roller compaction and twin-screw granulation processes.

HSWG is an important granulation method because of its inherent advantages. It includes several operational steps, such
as mixing and wet mass. Therefore, this process is complex, and influenced by many factors, including agitator speed, massing time, liquid addition rate, and interactions between them. Based on the QbD approach, the CMAs and CPPs can be identified in the HSWG process. In order to comprehensively understand and study the HSWG process, multivariate modeling techniques were applied combined with QbD or DoE to explore the impact of the CMAs and CPPs on granule

**Table 1** The applications of multivariate methods in the mixing process

| Models               | Blender type          | Application                                                                                   | References |
|----------------------|-----------------------|---------------------------------------------------------------------------------------------|------------|
| MLR                  | KG-5 blender          | The model was used to correlate the critical formulation and CPPs with the response variables. A design space for the powder mixing process was built | 42         |
| MLR                  | Continuous mixer      | The relationship between residence time and total particle length was explored                | 43         |
| Polynomial           | V-blender             | The effects of CPPs on CQAs were quantified to identify their relationship, and the design space was established | 44         |
| Quadratic model      | Cone shape tank       | Based on the regression model, the optimal mixing conditions, including the impeller speed and eccentricity, were found | 45         |
| PLS                  | Twin-screw blender    | In twin-screw blend feeding, the relationship between blend material properties and feeding capacity was developed | 46         |
| PLS                  | Square cone mixer     | The relationship between raw material variability and mixing time was quantified. The CMAs affecting the mixing process were identified according to VIP | 47         |
| PLS and MLR          | Continuous mixer      | The effect of material properties on the mean residence time was studied. The relationship between bulk density and mean residence time at different flow rates was determined | 48         |
characteristics.\textsuperscript{24} Table 2 summarizes the application of multivariate modeling techniques in HSWG. For example, Zhang et al.\textsuperscript{25} used QbD principles to improve their understanding of the HSWG process. Firstly, the CPPs were screened through a Plackett–Burman experimental design. Then, based on the Box–Behnken experimental design, a multivariate analysis (MVA) model was developed to investigate the relationship between granule size and CPPs. Finally, a design space was developed to optimize the HSWG process. Han et al.\textsuperscript{26} applied risk assessment to determine three CPPs in HSWG based on prior knowledge. In this study, the effects of agitator speed, spray rate, and massing time on the CQAs of granules were quantified by the MVA model.

For dry granulation, roller compaction is the most commonly-used type of dry granulation in the manufacturing of tablets. The roller compaction process does not use any liquid binders, which are available for water or heat-sensitive pharmaceutical materials.\textsuperscript{59} In the roller compaction process, mixed powders are first fed from the hopper to the compaction region by a screw. Then, the powder in the compaction region is pressed by rotating rollers to form a compacted ribbon.\textsuperscript{69,70} Finally, the ribbon is transported to a size reduction region, where the ribbon is milled to the desired particle size.

The traditional trial and error approach is very time-consuming and costly for developing or optimizing process parameters. Therefore, several mathematical models were established, including one-dimensional (1D), 2D, and 3D models to study the roller compaction process.\textsuperscript{76–79} However, due to the diversity of pharmaceutical powders and the complexity of manufacturing solid dosage forms, these mathematical models are only suitable for particular powders and roller compactors,\textsuperscript{74} making it challenging to build a “one size fits all” model.\textsuperscript{73}

The ribbon properties can significantly affect the attributes of granules, such as flowability, granule size distribution (GSD), and compressibility.\textsuperscript{52,76–78} They may influence the tablet compression process. Therefore, it is necessary to develop multivariate models to investigate the roller compaction process. In many studies, the MVA model was developed to study the complex roller compaction process.\textsuperscript{79–81} For example, Yu et al.\textsuperscript{74} developed a PLS model to explore the effect of raw material properties on ribbon properties using 81 pharmaceutical powders. Next, the CMAs and CPPs of roller compaction were identified based on the VIP of the PLS model. Finally, feasible material properties and operating regions were found by developing a multi-objective design space. The applications of the multivariate modeling techniques in the roller compaction process are presented in Table 3.

With the background of continuous manufacturing (CM) in the pharmaceutical industry, twin-screw granulation (TSG) is an alternative technique for transforming a batch pharmaceutical process to continuous manufacturing.\textsuperscript{82} It can granulate powders to granules in a short time and efficiently mix API and excipient.\textsuperscript{88} TSG includes three main components, respectively conveying elements, kneading blocks and comb mixer elements. To enhance the manufacturing efficiency of TSG, it is important to comprehensively study the impact of formulation and process parameters on the granule attributes. Many studies have reported on the relationship between CMAs, CPPs and CQAs.\textsuperscript{73–90} For example, Keleb et al.\textsuperscript{99} found that granulation yield decreased when increasing the screw speed. For the application of multivariate modeling techniques in TSG, Stauffer et al.\textsuperscript{100} developed an MLR model to investigate the relationships between raw material variability, process variables, and CQAs, such as GSD and granule friability. In their study, PCA was used to analyze the batch-to-batch variability of physical properties. According to the process variables in TSG, Ismail et al.\textsuperscript{101} built an ANN model to predict median residence time distribution based on the screw speed and L/S ratio in the process of twin-screw granulation.\textsuperscript{101}

In addition, a few studies have reported the effect of mill process parameters on the granule properties in drying granulation. Most of these studies focus on the properties of the ribbon. Consequently, a comprehensive study of all the granulation processes combined with multivariable modeling techniques is warranted.

4.1.3. Process understanding and optimization of tablet compaction. In general, the tablet compression process includes three consecutive steps: powder or granule die filling first, then compression, and finally ejection of the tablet.\textsuperscript{102} In this process, tablet formation may involve rearrangement and elastic and plastic deformation, thereby forming interparticulate bonds.\textsuperscript{103} In brief, the mechanism of forming tablets is more complicated, leading to a complex tablet compaction process. Even a small change in the compression process has a serious impact.\textsuperscript{104} The multivariate modeling technique is a useful tool to understand the tableting process.

4.1.3.1 Die filling in tablet compaction. The weight variation between tablets, content uniformity, tensile strength, disintegration time, friability, and dissolution were defined as CQAs.\textsuperscript{105} However, these CQAs mainly depend on die filling during tablet compaction. Non-uniformity in die filling results directly in variation in the weight and content uniformity of the tablets and may critically affect the compression force.\textsuperscript{106} Hence, die filling is a critical process for obtaining the required quality of the tablet. Although a DEM could explain underlying physical phenomena by first principles\textsuperscript{107} and is used in the die filling process,\textsuperscript{108–111} the computational costs were high. At present, multivariate modeling techniques still play a key role in investigating die filling. Studies performed at Ghent University used multivariate modeling techniques to systematically study the die filling process of high-speed rotary tablet compactors.\textsuperscript{112,113} In this study, a PCA model was used to analyze powder physical properties or classify relevant data obtained by DoE experiments. A PLS model was developed to investigate the relationship between attributes of blend powder, process conditions, and responses of product and process. This revealed that powder flowability, density, turret speed, paddle speed, feed frame, etc. significantly affect the die filling performance. According to the developed model, the process conditions can be optimized to meet target CQAs. In addition, multivariate models combined with a DEM were applied to study the die filling process. The main function of the DEM is to simulate the die filling process and to visualize the movement of particles in the die. Multivariate models, such as PLS and polynomial
models, were used to investigate the effects of potential CMAs, potential CPPs, and their interaction on tablet weight variation, segregation, tablet mass, etc.\textsuperscript{110,114}

4.1.3.2 Tableting process. The tablet manufacturing process can be regarded as a whole system. The effects of powder properties, mixing parameters, granulation parameters, and tableting parameters on CQAs can be comprehensively investigated. Researchers at Wyeth Pharmaceuticals studied the effects of granulation process parameters, granule properties and tableting parameters on dissolution behavior.\textsuperscript{115} In this study, a PLS model was used to quantify the relationship between these variables and to identify the CMAs and CPPs. Sun \textit{et al.}\textsuperscript{116} used a D-optimal experimental design to study tablet quality. Tablets were prepared using various raw materials under different process parameters of HSWG and compression. In their study, a MBPLS model was developed to quantify the relationship between raw material properties, various process conditions, tensile strength, and tablet disintegration time. The critical pharmaceutical unit operation, CMAs and CPPs were identified by the MBPLS model. Finally, according to the multiple linear

| Models                  | Application                                                                                              | References |
|-------------------------|----------------------------------------------------------------------------------------------------------|------------|
| Polynomial regression   | The effects of operational parameters, such as impeller speed, dosing speed, chopper speed and wet massing time on granule size were quantified | 57         |
|                         | The effects of process parameters, including granulation time, impeller, and formulation variables, on packing coefficient and strength of granules were investigated | 58         |
|                         | The relationship between granulation variables and the specific energy of the granules was determined      | 59         |
|                         | The best-fit equation was used to accurately predict the Carr’s index for granules under different formulation factors | 60         |
| Polynomial, MLR         | The impact of formulation variables on granule properties like flowability and size was assessed, which was beneficial for selecting the desired formulation | 61         |
|                         | Combined with DoE, models which correlated the process parameters with granule properties, were developed. This provided the basis for adjusting process parameters according to the product quality attributes | 62         |
|                         | Using DoE techniques, the effects of amount of water and massing time on the key quality attributes of granules were investigated | 63         |
| PLS                     | The relationship between impeller speed and total power spectral densities (TPSDs) was developed. The research demonstrated that audible acoustic emissions could monitor process changes in real time | 64         |
| Gene expression programing model | Impeller power can be predicted according to the impeller diameter, impeller speed, the percentage of the liquid and mean torque | 65         |
| PCA, MLR                | The relationship between process variables on granule hardness and Carr’s index was developed. Based on the PCA model, it was shown that there was a strong correlation between the impeller speed and wet massing time with the granule attributes | 66         |
| Polynomial, MLR, PLS, ANNs | Based on various MVA models, the relationship between three process parameters and CQAs of granules such as mean size and flowability was quantified | 67         |
| PLS, MBPLS, OPLS        | Various MVA models were developed to investigate the effects of HSWG process variables and granule properties on tablet quality | 68         |
regression model, a multi-objective design space was established to control tablet quality. Table 4 shows the many multivariate modeling techniques for studying the tableting process.

In an LVM, such as a PCA or PLS model, multivariate modeling techniques are mainly applied in the tableting compression process. The MBPLS algorithm analyzes the multiple block data from various unit operations in tablet manufacturing, which is beneficial for investigating the tablet manufacturing process.

### 4.1.4. Coating process.
Coating of tablets is an important pharmaceutical process. In general, during the coating process, tablets are repeatedly exposed to a spray that contains several solutes and solvents. The coating process is repeated until the desired coating quality and/or uniformity are achieved. The purpose of coating includes masking the taste or odor to improve compliance, thereby obtaining a better appearance and easier distinguishability, changing drug release behavior, and adding a second API in the coat. The pharmaceutical coating process needs to ensure that the tablets are coated uniformly. To ensure uniformity of the coating, the variability of the coating is mainly controlled by the thickness of inter- and intra-coating. Although coating is a widespread pharmaceutical process and has been used for many decades, there are still some serious issues, such as a lack of comprehensive understanding of the coating process.

To solve the above issues and promote a more in-depth understanding, the multivariate modeling technique can be used to study the coating process. Rege et al. applied the Plackett–Burman experimental design to investigate the relationship between manufacturing parameters and final product quality. In their study, the effects of process parameters on the CQAs were analyzed by an MLR model.
pressure, pan speed, and coating time were identified as CPPs affecting content uniformity. For the film coating process, Tanabe et al.\textsuperscript{133} generated a PLS model according to the data derived from the production process. Moreover, Cahyadi et al.\textsuperscript{134} used a multivariate model combined with DoE to optimize process parameters in the quasi-continuous tablet coating process. Table 5 shows other studies involved in the application of multivariate modeling techniques in the coating process.

In summary, the above-mentioned tables (from Tables 1–5) show that PLS, MLR, and polynomial models are the most widely used in the pharmaceutical process, while OPLS, MBPLS, and ANN models only have a few applications. However, manufacturing tablets that includes several operation units may separate the process variables into specific blocks. MBPLS and OPLS can be applied to investigate the effects of each block of data on the response variables.\textsuperscript{142} The most prominent shortcoming of these methods in the process analysis is that the prediction blocks are independent. In the actual manufacturing process, the various process units are usually highly correlated; therefore, the multivariate modeling technique should consider these relationships. PLS-path modeling (PLS-PM) is beneficial for understanding the inner mechanism of the manufacturing process by incorporating process knowledge into predictive modeling.\textsuperscript{143} However, PLS-PM is rarely used in the tablet manufacturing process. An ANN model is a black-box model. However, in a recent study,\textsuperscript{144} an ANN model combined with specialized software identified CMAs and CPPs. This demonstrated that an ANN model is not an absolute “black-box” model

### Table 4: The applications of multivariate models in the tableting process

| Model         | Application                                                                 | References |
|---------------|------------------------------------------------------------------------------|------------|
| Polynomial model | The relationship between raw material properties and tablet tensile strength was quantified | 84         |
| PCA           | The model was used to evaluate the relationship between various powder properties for direct compaction | 117        |
| PCA, PLS      | The PCA model can investigate the relationship between various powder and compression properties. The PLS model was developed to identify key effecting factors, and to quantify the relationship between those factors and tablet tensile strength | 118        |
|               | The purpose of developing the PLS model was to study the effects of input variables, such as properties of raw materials or intermediates, and process conditions on tablet dissolution | 119        |
|               | The PCA model can reduce the dimensionality of the original data and explore the relationship between different physical properties. The PLS model was developed to investigate the effects of raw material attributes, tableting process variables and compression behavior indices on the tablet quality attributes | 120        |
| PCA, MBPLS    | According to the PCA model, the variability in powder and granules was analyzed. MBPLS was used to identify critical factors and critical process units. It can quantify the relationship between process variables and tablet dissolution | 121        |
| PCA, MLR      | The linear model was used to determine the effect of excipient properties on tablet attributes, and the PCA model was built to correlate the properties of filler and binder with tablet properties | 122        |
| PLS           | The purpose of developing the PLS model was to explore the impact of process variables on the mass flow rate per unit orifice area in the die filling | 123        |
| MLR           | Based on the model, the effects of roller compaction conditions and milling process variables on the attributes of granule and tablet were studied | 124        |
| ANN           | The relationships between material physical properties, process parameters and tablet quality attributes were determined | 125        |
and that ANN can be used as a valuable tool to understand and optimize the pharmaceutical process. Liu et al. developed a controllable and readable ANN model (CR-ANN), also called a white-box model. The CR-ANN model can improve the readability of the internal structure; therefore, it is easy to overcome the shortcomings of a traditional ANN model. CR-ANN can be further widely used in process understanding and optimization. In particular, the mechanism model combined with multivariate modeling techniques should be developed for the pharmaceutical process, which will be conducive to comprehensively gaining process knowledge. Thus, in the future, it will be necessary to expand the application of this modeling technique in the manufacturing process.

### 4.2. Pharmaceutical process monitoring based on multivariate modeling techniques

To ensure consistency of product quality and stability of therapeutics, it is necessary to monitor, analyze, and control each unit operation. With American Food and Drug Administration (FDA) guidance on PAT, most scholars and the pharmaceutical industry applied the new process analysis methods to research the pharmaceutical process combined with multivariate modeling techniques. Most of these studies used multivariate modeling techniques to establish a multivariate calibration model; then the calibration model was implemented to predict the quality of the product in real time.

#### 4.2.1. Monitoring blending processes

To effectively detect the CQAs and gain a deeper understanding of the change in dynamics of the mixing process, near-infrared (NIR) instruments are widely used PAT tools to monitor the blending process. The application of Raman spectroscopy in a study of the mixing process has also been reported. However, in order to obtain meaningful results, spectral data must be transformed by chemometrics. Multivariate modeling approaches, which play an important role in translating spectral signals to BU, can be applied to monitor the mixing process.

For qualitative methods, PCA, as one LVM, is one of the most preeminent approaches that provides reliable results. PCA has been used in many studies to detect and monitor mixing homogeneity. For PCA, one method plotted the score graph of the first principal component (PC1) as a function of mixing time. It assumes that mixed uniformity is achieved when the

| Models          | Application                                                                                                                                                                                                 | References |
|-----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| MLR, polynomial | The MLR model identifies critical factors affecting CQAs. The polynomial model quantifies the relationship between process variables and CQA in the coating process. Based on the model, the optimal region for process variables was defined. According to a full factorial design, MLR models were developed to correlate coating process conditions with response variables such as coating uniformity and surface roughness. Based on the polynomial model, the effects of critical process variables on coating uniformity were investigated at lab and pilot scales. | 135        |
| Polynomial      | Using a full factorial design, polynomial models were developed to correlate coating process conditions with response variables such as coating uniformity and surface roughness. According to DoE, based on the polynomial model, the effects of critical process variables on coating uniformity were investigated at lab and pilot scales. Based on the central composite design, the polynomial model was developed to study the relationship between key process parameters and CQAs such as weight gain and surface roughness. | 137, 138, 139 |
| Multivariate model | According to a central composite - face-centered response surface design, the relationship between five process parameters and CQAs like tablet appearance were correlated by a multivariate model. The optimal process variables could be determined based on the model. | 140        |
| Quadratic polynomial | A multivariate model, which links four operating variables such as spray rate, rotation speed of pan, and spray temperature to the weight variability index, was established. | 141        |
score level is stable and close to a point.\textsuperscript{156,157} In another method, the first two principal components of homogeneous powders (the target mixture) were analyzed. If the powders (the test mixture) approach or overlap the target mixture cluster in the score plot, this would indicate that the test mixture had achieved a state of uniformity.\textsuperscript{157,158} For the PLS model, the

**Table 6** Pharmaceutical applications of multivariate modeling techniques in monitoring the powder mixing process

| Models | Characterization methods | Application | References |
|--------|--------------------------|-------------|------------|
| PCR    | NIR                      | Detection of blending homogeneity | 156, 158, 161 and 162 |
|        | NIR-CI                   | Detection of blending end-point | 163 |
|        | NIR-CI                   | Determination of mixture homogeneity | 164 |
|        | NIR                      | Confirming the end-point of blending | 165 |
|        | Raman                    | Confirming the end-point of blending | 165 |
|        | NIR                      | The endpoint of the start-up phase | 166 |
|        | NIR                      | Process spectral data | 167 |
|        | NIR                      | Determining the blend uniformity | 146 |
| PLS    | NIR                      | Detection of blending end-point | 147, 158, 159, 165, 168 and 169 |
|        | NIR-CI                   | Verifying the NIRS analyzer response and assessing homogeneity | 167 |
|        | NIR                      | Drug concentration | 169 |
|        | NIR                      | Confirmation of blend uniformity | 170 |
|        | NIR                      | Monitoring the drug level at the outlet of the continuous blender | 166 |
|        | NIR                      | Measurement of contents | 171 and 172 |
|        | NIR                      | Assessment of powder blend uniformity | 146 and 173-176 |
|        | NIR                      | Evaluation of degree of homogeneity | 177 |
| PCA-ANN| NIR                      | Measurement of blend uniformity | 175 |
| ANN    | Image analysis           | Prediction of mixing time | 178 |
| ANN    | NIR                      | Determining the blend uniformity | 146 |

**Table 7** Applications of multivariate models in monitoring or detecting the granulation process

| Model             | Application                                                                 | References |
|-------------------|-----------------------------------------------------------------------------|------------|
| PCA, PLS          | The PCA model was used to detect changes in ribbon density qualitatively. However, the PLS model can quantitatively monitor ribbon density during the roller compaction process. The envelope density and moisture content of the ribbon were monitored. The PCA model can be used to analyze textural descriptors of the ribbon. The density can be predicted by the PLS model. The PCA model can obtain the variation between various ribbons, and the PLS model can detect the porosity of the ribbon. The water contents in the wet mass can be measured by in-line NIR based on the PLS model. | 187 |
| PLS               | Granule properties like granule size can be detected in real time according to the PLS model. It can be used to monitor the hardness of the ribbon online. The content uniformity and ribbon properties, like moisture content or relative density, could be determined in real time. The key ribbon properties, like relative density and tensile strength, could be predicted by the model. The model was used to predict the solid fraction of ribbon. The attributes of flakes, such as tensile strength, Young's modulus, and relative density, can be monitored in real time. | 192 |
| PLS, ANN          | Based on the models, the APIs in traditional Chinese medicine (TCM) granules can be detected. | 198 |
content of constituents can be predicted during the mixing process. Firstly, it is necessary to develop an off-line PLS calibration model. Then, ingredient contents could be predicted in real time during the mixing process. When the deviation of component concentration reaches a pre-defined limit within a specific time, it can be considered that the test mixture had reached a state of mixed uniformity. Another multivariate model is the ANN algorithm. An ANN model detects blend uniformity in a similar manner to a PLS model. Compared with a PLS model, the performance of an ANN model is excellent. However, in practice, considering the complexity and time-consuming nature of developing an ANN model, ANN models were applied relatively less often. This is shown in Table 6.

4.2.2. Monitoring the granulation process. The multivariate modeling technique is not only used to understand the granulation process, but also to monitor the granule properties or content of API based on PAT. For example, Khorasani et al. applied a PCA model to map the ribbon porosity distribution based on near-infrared chemical imaging (NIR-CI) technology. The API concentrations in the ribbon were quantified by a PLS model. In many studies, NIR-CI or NIR were used, as shown in Table 7. NIR or microwaves can measure key properties of the ribbon online during the roller compaction process in real time. For instance, Gupta et al. used a PLS calibration model to simultaneously predict relative density, moisture content, and tensile strength in real time. Gupta et al. used microwave online real-time monitoring density, API content and moisture content of ribbon based on ANN or PLS models. In this study, the performance of the different multivariate models was compared. The prediction accuracy of microwave and NIR sensors was also compared. In addition, the multivariate modeling technique has been widely used to predict or detect CQAs or API contents in the twin screw granulation process.

4.2.3. Monitoring or prediction of tablet quality. The pharmacopoeia stipulates that items to be inspected for tablet quality include content uniformity, hardness, dissolution, and disintegration time. Content uniformity is a critical quality parameter, which is directly related to drug efficacy. The distribution of API in the tablet, which may affect texture and dissolution, is another CQA. The disintegration time affects the release of the drug in vivo. Mechanical properties, such as fragility and hardness, affect the transport and storage of the tablets. Hence, non-destructive and quick detection of the quality of tablets is an important task. For example, Blanco et al. used NIR spectroscopy to simultaneously detect the chemical composition and physical properties of tablets. In many reports, the component content in the tablet can be predicted by multivariate modeling techniques in real-time monitoring of the tabletting process. In many studies, real-time release test (RTRT) models were developed, and aimed to monitor or predict the CQAs of the product during the process, without requiring lengthy off-line assays. For example, based on a PLS or MLR model, the content uniformity, hardness, disintegration time, friability, and dissolution of tablets can be tested in the manufacturing process and can significantly reduce the release time. The applications of multivariate modeling techniques combined with the PAT tool to predict or monitor tablet CQAs are shown in Table 8.

4.2.4. Monitoring or prediction of the tablet coating process. Based on the multivariate model, there are many reports of using the PAT tool to monitor the coating process.

| Model   | Tool          | Application                          | References          |
|---------|---------------|--------------------------------------|---------------------|
| MLR     | NIR           | Detection of tablet hardness          | 199 and 206         |
| PCA     | NIR           | Measurement of drug content           | 207                 |
| PCR     | NIR           | Measurement of drug content           | 208                 |
| PLS     | NIR           | Detection of tablet hardness          | 209-212             |
| PLS     | NIR           | Measurement of tablet weight variation| 212                 |
| PLS     | NIR           | Prediction of tablet porosity         | 209                 |
| PLS     | NIR           | Detection of tablet hardness          | 199, 206, 207, 213 and 214 |
| PLS     | NIR           | Prediction of tablet porosity         | 213                 |
| PLS     | NIR           | Measurement of tensile strength       | 215-217             |
| NIR     | NIR           | Measurement of moisture               | 216                 |
| NIR     | NIR-CI        | Prediction of dissolution behavior    | 199 and 219-227     |
| NIR     | Raman         | Prediction of dissolution behavior    | 228                 |
| NIR     | NIR-CI        | Measurement of drug content           | 224                 |
| NIR     | Raman         | Measurement of drug content           | 199, 201, 203, 214, 216 and 229-249 |
| NIR     | NIR           | Measurement of drug content           | 250 and 251         |
| NIR     | NIR-CI        | Measurement of drug content           | 235, 248, 252 and 253 |
| NIR     | Raman         | Measurement of drug content           | 241                 |
| NIR     | NIR           | Measurement of drug content           | 218                 |
| NIR     | NIR           | Measurement of disintegration time    | 254                 |
| NIR     | NIR           | Measurement of tablet weight          | 237                 |
| NIR     | NIR           | Prediction of dissolution behavior    | 224                 |
| NIR     | NIR           | Measurement of drug content           | 224                 |
| NIR     | NIR           | Detection of tablet hardness          | 214, 229 and 255    |
| NIR     | NIR           | Detection of tablet hardness          | 214                 |
the most commonly used PAT tool, NIR can measure and monitor any CQAs during the coating process. Romer et al. developed an off-line PLS model to in-line predict the coating thickness in the film process. Lee et al. used the following equipment (Fig. 5) to in-line monitor the fluid bed coating process based on a PCA model and a PLS model. There are many studies on NIR testing of the coating process based on a multivariate model, as shown in Table 9, which shows that NIR combined with the multivariate modeling technique is widely applied in the pharmaceutical coating process.

As another commonly used PAT tool, Raman spectroscopy is also used in the pharmaceutical coating process. Hagrasy et al. first demonstrated that real-time monitoring of the coating process could be achieved by using in-line Raman spectroscopy. Muller et al. systematically investigated the application of Raman spectroscopy in the coating process. In their study, a calibration PLS model was built to link Raman spectral data to diprophylline contents in the coating layer. According to the PLS model, the diprophylline contents could be determined in real time. The developed method was validated in accordance with ICH guidelines. This proved the feasibility of in-line monitoring by Raman spectroscopy of the coating process combined with a multivariate modeling technique. In the following study, Raman spectroscopy could measure other CQAs, such as thickness based on the PLS model. Table 9 summarizes the applications of NIR and Raman spectroscopy in the coating process, and shows that PLS models are widely used in the coating process.

In summary, in pharmaceutical process monitoring, a PLS model is one of the standard tools in the arsenal of making predictions based on the large amounts of data collected during production processes. A PLS model can extract the most predictive and correlative information from a large number of spectral and high-dimensional data. Consequently, a PLS model is powerful in monitoring and predicting the tablet manufacturing process. However, the disadvantage is that the model handles all metrics without hierarchy or conditions. Therefore, a PLS model is limited in understanding the pharmaceutical process that produces variation in the measurement data. Therefore, much of the process knowledge is still untouched, because information about the pharmaceutical process is not comprehensively included in predictive modeling. The performance of an ANN model was proven to be superior to conventional multivariate modeling techniques. For example, according to a study by Nagy et al., it was found that the prediction error provided by an ANN model was lower than that of a PLS model in detecting dissolution behavior. However, ANN models in pharmaceutical process monitoring are still used relatively less often, which is in agreement with the findings presented in a previous report. For an ANN model, since the spectra contain thousands of variables, directly inputting them as independent variables into the ANN model will greatly increase the complexity of the model and significantly reduce the efficiency of modeling. This may be one of the reasons why ANN model are used less often in pharmaceutical process monitoring. Dimensionality reduction of the spectral data is an effective way to improve the efficiency of ANN modeling, while maintaining a strong prediction performance. In the pharmaceutical process, PCA is always used to reduce the dimensionality of the data. Hence, an ANN model combined with PCA is utilized to monitor CQAs in the tablet manufacturing process.

### 4.3 Pharmaceutical process control based on multivariate modeling techniques

Pharmaceutical product quality is multivariate in nature, and all measured properties must meet requirements simultaneously. Process control depends on an understanding of the

---

**Table 9** The applications of multivariate models in monitoring or detecting for the coating process

| Model | Application                                  | Analysis tool | References |
|-------|----------------------------------------------|---------------|------------|
| PCA   | Clustering of spectral data                  | NIR           | 257 and 262|
| PCA   | Visualization of the coating process         | NIR           | 263        |
| PLS   | Prediction of coating thickness              | NIR           | 256, 257, 262 and 263 |
|       | End-point detection of a coating             | NIR           | 264        |
|       | Prediction of weight gain                    | NIR           | 262 and 263|
|       | Prediction of moisture content               | NIR           | 263, 265 and 266 |
|       | Measurement of curing degree                | NIR           | 267        |
|       | Prediction of API content                    | Raman         | 259, 260 and 268–270|
|       | Prediction of weight gain                    | Raman         | 259        |
|       | Prediction of coating thickness              | Raman         | 271        |
|       | Prediction of moisture content               | Raman         | 266        |
pharmaceutical process and the accumulation of relevant knowledge. Multivariate modeling techniques, especially LVMs, are important tools for pharmaceutical process control.

LVMs can be used to analyze the historical data from the pharmaceutical process, and thus corresponding feedback or feedforward control strategies could be developed. For example, combined with the NIR tool, Hattori\textsuperscript{274} developed a robust feed-forward control model based on a PLS model in the tabletting process. The process parameters predicted by the control model could produce tablets that met the required quality. Westerhuis \textit{et al.}\textsuperscript{275} established a step-by-step control method for the wet granulation and tableting process based on PLS and MBPLS. First, the properties of the granules were determined by the formulation variables and granulation process parameters. Then, the tableting process parameters were adjusted according to the properties of the granules. Consequently, this control

| Model                        | Application                                                                                                                                                                                                 | References |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| MSPC                         | Based on the PCA model, the MSPC tool can be developed to control the granulation and drying process. After the deviation has been corrected, the process system can return to a stable state. An MSPC tool was built based on PLS and PCA models. The control chart was used to monitor humidity and temperature in the granulation process and to detect process abnormalities. | 279        |
| PLS                          | According to the feed-forward PLS model, it can be used to determine the process conditions based on the preceding conditions of unit operation. According to the PLS model, the feed-forward process can construct a control plan to determine the optimal process conditions. Then a release profile meeting the required quality was obtained. The feed-forward model for tablet compaction was developed based on NIR. Based on this model, the process parameters that meet the final product quality can be predicted. Based on NIR, the PLS model can measure the bulk density in real time. The measured signals can be used for forward feed control to ensure small density variation. | 254 280 281 282 283 |
| Latent projection model      | According to the latent projection model, the variations in raw material and from batch-to-batch can be controlled by adjusting some process parameters. | 284        |

Fig. 6  Overview of the control strategy combined with risk management tools.
scheme for manufacturing tablets can adjust process conditions to better meet the specifications. Based on a PLS model, Garcia et al.\textsuperscript{276} developed a feed-forward control method to deal with the variation in initial raw material quality in high-speed shear wet granulation. A comprehensive design space was developed, considering a complex network of relationships between process conditions and product quality. Muteki et al.\textsuperscript{277} developed a feed-forward control method for the process of dry granulation.

| Model | Application | References |
|-------|-------------|------------|
| PLS, PCA | In the continuous powder blending and tableting process, the PCA model, as an exploratory data analysis tool, was used to explore the effects of experimental variables on PAT spectra. The PLS model was applied to predict the CQAs of the tablet. | 300 |
|       | In the CDC manufacturing process, the PCA model was applied to identify possible outliers or abnormalities. Transmission NIR spectroscopy, combined with the PLS model, was used to measure blending uniformity and detect tablet content uniformity. | 301 |
|       | In the CM line, the PLS model was performed to detect CQAs. The multivariate analysis model could be applied for process monitoring. The PCA model was performed to analyze spectral data. The PLS model, which links spectral data to response variables of interest, was established. The developed multivariate models can be integrated into the online prediction tool. | 302 |
|       | Based on the PCA and PLS model, multivariate monitoring charts could monitor various units in the continuous manufacturing process. | 303 |
| PLS/PCA, multiblock PCA/PLS | Various latent models were used to describe and monitor the time variables in the continuous twin-screw granulation and drying process. It can detect and diagnose deviations in the continuous manufacturing process. | 304 |
| PCA, MLR | In the RTRT of the tablet, the multivariate model was developed to predict dissolution profiles in a CDC system. The NIR data were processed by the PCA model. Based on the NIR spectra, MLR could be applied to predict tablet dissolution behavior. | 305 |
| PLS | Based on the calibration model, the powder density in a continuous line can be predicted. The PLS model was developed to predict blending powder bulk density in the CDC manufacturing process based on the NIR data. Using the NIR tool, the developed off-line PLS calibration model could monitor the continuous pharmaceutical manufacturing process's API concentration. NIR spectroscopy as a PAT tool was used to measure API content combined with the PLS model. | 306 |
|       | Using the NIR tool, the developed off-line PLS calibration model could monitor the continuous pharmaceutical manufacturing process's API concentration. NIR spectroscopy as a PAT tool was used to measure API content combined with the PLS model. | 307 |
| MLR | MLR models were built to explore the relationship between process conditions and response variables, such as flowability, ejection force, and tablet strength. Based on the model, a design space was developed for high-dose tablets in CM. | 308 |
| PCA | The PCA model can extract concentration-related information from NIR spectral data. | 309 |

---

Table 11: The applications of multivariate modeling techniques in continuous tablet manufacturing

© 2021 The Author(s). Published by the Royal Society of Chemistry.

RSC Adv., 2021, 11, 8323-8345 | 8337
granulation and tableting. For different batches of raw materials, this method can simulate and calculate the best operating parameters, and reduce the differences in tablet quality.

In addition, based on the multivariate modeling technique, the multivariate statistical process control (MSPC) tool deals with large sets of data derived from pharmaceutical processes. It is an effective tool for process control. For example, Burggraeve et al.278 developed a real-time online monitoring method for fluidized bed granulation based on LVMs. This method monitors the quality of particles in real time, and products that do not meet the expected quality requirements can be identified. Table 10 shows the applications of these models in manufacturing tablets.

In the tablet manufacturing process, multivariate modeling techniques play an important role in process control. The core of using the multivariate model technique for process quality control is model inversion. According to the quality goals for the final product, optimal process operating conditions can be solved by the developed model, which can accurately determine the specific plan and autonomously control it. Consequently, the risk of manufacturing a product outside the acceptance criteria would be decreased. Based on the hierarchical three-layer control design proposed by Su et al.,285 the above-mentioned control strategies are layer 1 control approaches. At this control level, an efficient feedback or feed-forward control method is used to diminish the effect of variation that may propagate in downstream unit operations.286 Another feature of layer 1 control is that automatic adjustment usually spans several pharmaceutical operation units. Real-time monitoring of CQAs is one of the key components in layer 1 control. However, this remains challenging due to the complexity of PAT tool calibration and model validation in real time.287 In particular, PAT sensor positions, sampling problems, and contamination can cause measurement drift and deviations, which can affect the accuracy of real-time process data. Therefore, it is necessary to systematically study the PAT tool network design to develop approaches for maintaining reliable measurements of CQA. In the future, layer 2 control design should be developed in the tablet manufacturing process based on dynamic models. At this control level, many advanced process control techniques, such as predictive models and real-time optimization, can be integrated to achieve multi-input multi-output control. Even smart manufacturing (SM)288,289 can be developed for the pharmaceutical process, which can rapidly and in real time optimize manufacturing pharmaceutical products and the dynamic response to the demand for pharmaceutical product quality.290 Combined with risk management tools (Fig. 6), the consistency of the quality of the final product can be ensured.

5. Application of multivariate modeling techniques in a continuous manufacturing line

For manufacturing tablets, continuous manufacturing (CM), as an approach for modernizing the production system, can produce pharmaceutical products of higher quality.291 Because CM has many advantages, such as flexibility, speeding up the supply chain, lower risk of stock-outs, and reduced space and investment costs,292,293 it has been widely used. Nowadays, seven drug products from four different pharmaceutical companies, Vertex, Johnson & Johnson, Eli Lilly and Pfizer, have been approved by different regulatory agencies for continuous manufacturing.294 Drug products from CM have brought significant economic benefits to these four pharmaceutical companies. A study from Transparency Market Research forecast that from 2016 to 2025, the market values of pharmaceutical continuous manufacturing technology might increase from US$1.74 billion to US$3.69 billion.295 Therefore, the CM of pharmaceutical products has vast market prospects.

Using the quality by control (QbC) approach,296 multivariate modeling techniques are a useful tool for studying continuous manufacturing. For example, as an integrated continuous tablet manufacturing line, the PROMIS-line was developed at the University of Eastern Finland.296 Furthermore, in the process of powder feeding, the PCA model was applied to analyze the process data and the time evolution of powder feeding can be visualized. In another study at the University of Eastern Finland an end-to-end continuous direct compression (CDC) line was built.297 In these studies, multivariate modeling techniques, including PLS and OPLS models, were used to correlate spectral data with API contents. Hattori et al.298 developed a feed-forward control method to control a tablet continuous compression process based on a PLS model. Pauli et al.299 applied LVM to monitor the continuous manufacturing tablet process. In their study, a PLS model was used to predict the API content to evaluate blending uniformity and content uniformity in the tablets. The PCA model qualitatively describes the process monitoring. Stauffer et al.300 investigated the relationships between the variability of raw materials, CPPs and CQAs based on empirical models, such as PLS and MLR models. A design space of the CM line was developed, which can guarantee the consistency of the quality of the final product and process stability. Reports on the CM of tablets are presented in Table 11.

Although continuous manufacturing has many advantages and CM lines, such as that of GEA Pharma Systems, have been built,301 the progress towards CM is still slow.292 Furthermore, CM is facing many challenges, such as developmental challenges, and supervisory and quality control challenges.291 In the CM process, there is limited hold-up in each pharmaceutical operation unit. Thus, the variability of raw materials in upstream unit operations can rapidly and directly affect downstream pharmaceutical processes, which can significantly impact final pharmaceutical product quality.287 Furthermore, control engineering theories for pharmaceutical continuous manufacturing process have not yet achieved a common understanding or wider application in the pharmaceutical process. The material library, which was proposed to gain sufficient knowledge about raw material attributes,74,311,312 has not been studied for the CM process in solid oral dosages. In the future, based on the material library, multivariate modeling techniques or hybrid models should be established to study the relationship between critical equipment settings (CEEs), CMAs, CPPs and CQAs. Continuous improvement is another important
task in the CM process. However, relevant studies are rarely reported by the pharmaceutical industry. Based on the QbC concept, continuous improvement in pharmaceutical CM can easily be achieved at multiple levels, such as process performance monitoring, and predictive maintenance. To ensure the consistent quality of the product at every pharmaceutical operation unit, comprehensive control strategies and risk analysis should be proposed and implemented in the pharmaceutical process.

6. Conclusion

The tablet manufacturing process is inherently multivariate because it involves several basic pharmaceutical unit operations and produces multivariate data. However, the multivariate modeling technique is an appropriate tool to analyze these data. Consequently, this article provides a review of the applications of multivariate modeling techniques in tablet manufacturing. Multivariate modeling techniques can be used to understand the complex pharmaceutical phenomenon by investigating the relationships between measured variables. It can, to some extent, reveal the causes of process behavior, and plays an irreplaceable role in process optimization, monitoring, control, and continuous manufacturing processes. Through this review, we have aimed to provide more information to both researchers and manufacturers to improve tablet quality.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

This review was supported by the National Major Scientific and Technological Special Project for “Significant New Drugs Development” (2019ZX09301160, 2019ZX09721001-005-002).

References

1. S. Patel, A. M. Kaushal and A. K. Bansal, Crit. Rev. Ther. Drug Carrier Syst., 2006, 23, 1–66.
2. Solid dose: under-hyped but not under-represented, https://www.pharmamanufacturing.com/articles/2019/solid-dose-under-hyped-but-not-under-represented/
3. L. X. Yu and M. Kopcha, Int. J. Pharm., 2017, 528, 354–359.
4. V. K. Ojha, S. Schiano, C. Wu, V. Snasel and A. Abraham, Neural. Comput. Appl., 2018, 29, 467–481.
5. T. Kourti and B. Davis, Pharmaceut. Eng., 2012, 32, 52–62.
6. M. A. Behjani, Y. G. Motlagh, A. E. Bayly and A. Hassanpour, Powder Technol., 2020, 366, 73–81.
7. C. Hildebrandt, S. R. Gopireddy, R. Scherlies and N. A. Urbanetz, Adv. Powder Technol., 2020, 31, 755–769.
8. K. Takayama, T. Sato, K. Sato, H. Todo, Y. Obata and K. Sugibayashi, J. Drug Delivery Sci. Technol., 2019, 52, 1021–1031.
9. A. Kumar, J. Dhome, J. Bertels, D. Klingeleers, K. V. Gernaey, I. Nopens and T. De Beer, 2017, https://biblio.ugent.be/publication/8542476.
10. M. G. Ierapetritou and R. Ramachandran, Process simulation and data modeling in solid oral drug development and manufacture, Springer, 2016.
11. G. M. Troup and C. Georgakis, Comput. Chem. Eng., 2013, 51, 157–171.
12. J. Gabrielson, N. O. Lindberg and T. Lundstedt, J. Chemos., 2002, 16, 141–160.
13. T. Rajalahti and O. M. Kvalheim, Int. J. Pharm., 2011, 417, 280–290.
14. B. R. Kowalski, J. Chem. Inf. Comput. Sci., 1975, 15, 201–203.
15. E. Tomba, P. Facco, F. Bezzo and M. Barolo, Int. J. Pharm., 2013, 457, 283–297.
16. P. Geladi and B. R. Kowalski, Anal. Chim. Acta, 1986, 185, 1–17.
17. A. P. Ferreira and M. Tobyn, Pharm. Dev. Technol., 2015, 20, 513–527.
18. J. A. Westerhuis, T. Kourti and J. F. Macgregor, J. Chemom., 1998, 12, 301–321.
19. N. Kettanehwold, Chemom. Intell. Lab. Syst., 1992, 14, 57–69.
20. S. G. Munoz, J. F. Macgregor and T. Kourti, Chemom. Intell. Lab. Syst., 2005, 79, 101–114.
21. S. Ibric, J. Djurisic, J. Parojcic and Z. Djuric, Pharmaceutics, 2012, 4, 531–550.
22. V. Vutariya, A. Groshev, P. Sadana, D. Bhatia and Y. Pathak, Open Bioinf. J., 2013, 7, 49–62.
23. J. Bourquin, H. Schmidli, P. Van Hoogeveest and H. Leuenberger, Pharm. Dev. Technol., 1997, 2, 95–109.
24. J. Bourquin, H. Schmidli, P. Van Hoogeveest and H. Leuenberger, Pharm. Dev. Technol., 1997, 2, 111–121.
25. A. K. Jain, J. Mao and K. Mohiuddin, Computer, 1996, 29(3), 31–44.
26. M. F. Simoes, G. Silva, A. C. Pinto, M. Fonseca, N. E. Silva, R. M. Pinto and S. Simoes, Eur. J. Pharm. Biopharm., 2020, 152, 282–295.
27. G. Sin, K. V. Gernaey and A. E. Lantz, Biotechnol. Prog., 2009, 25, 1043–1053.
28. Good modelling practice, https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1072&context=usepapapers.
29. N. Crout, T. Kokkonen, A. J. Jakeman, J. P. Norton and P. Whitfield, Good Modelling Practice, 2008, vol. 3, pp. 15–31.
30. ICH-Endorsed Guide for ICH Q8/Q9/Q10 Implementation, https://database.ich.org/sites/default/files/Q8_Q9_Q10_Q% 26As_R4_Points_to_Consider_2.pdf.
31. T. Saffaj, B. Ihssane, F. Jihal, H. Bouchafa, S. Laslami and S. A. Soesse, Analyst, 2013, 138, 4677–4691.
32. Z. Wu, B. Xu, M. Du, C. Sui, X. Shi and Y. Qiao, J. Pharm. Biomed. Anal., 2012, 62, 1–6.
33. O. Scheibelhofer, N. Balak, P. Wahl, D. M. Koller, B. J. Glasser and J. Khinast, AAPS PharmSciTech, 2013, 14, 234–244.
34. R. De Maeschalck, F. C. Sanchez, D. L. Massart, P. Doherty and P. A. Hailey, Appl. Spectrosc., 1998, 52, 725–731.
279 A. F. Silva, M. C. Sarraguca, M. Fonteyne, J. Vercruysse, F. De Leersnyder, V. Vanhoorne, N. Bostijn, M. Verstraeten, C. Vervaet and J. P. Remon, *Int. J. Pharm.*, 2017, *528*, 242–252.

280 R. Kona, H. Qu, R. Mattes, B. Jancsik, R. M. Fahmy and S. W. Hoag, *Int. J. Pharm.*, 2013, *452*, 63–72.

281 S. Matero, S.-P. Reinikainen, M. Lahtela-Kakkonen, O. Korhonen, J. Ketolainen and A. Poso, *J. Chemom.*, 2008, *22*, 653–660.

282 Y. Hattori and M. Otsuka, *Int. J. Pharm.*, 2017, *524*, 407–413.

283 R. Singh, A. D. Romanospino, R. J. Romanach, M. G. Ierapetritou and R. Ramachandran, *Int. J. Pharm.*, 2015, *495*, 612–625.

284 T. Kourti, *Am. Pharm. Rev.*, 2009, *12*, 118.

285 Q. Su, M. Moreno, A. Giridhar, G. V. Reklaitis and Z. K. Nagy, *J. Pharm. Innov.*, 2017, *12*, 327–346.

286 R. Singh, F. J. Muzzio, M. Ierapetritou and R. Ramachandran, *Processes*, 2015, *3*, 339–356.

287 Q. Su, S. Ganesh, M. Moreno, Y. Bommireddy, M. Gonzalez, G. V. Reklaitis and Z. K. Nagy, *Comput. Chem. Eng.*, 2019, *125*, 216–231.

288 S. M. L. Coalition, *Implementing 21st Century Smart Manufacturing*, [www.academia.edu/1984950/Implementing_21st_Century_Smart_Manufacturing](http://www.academia.edu/1984950/Implementing_21st_Century_Smart_Manufacturing).

289 D. Zuehlke, *Annu. Rev. Contr.*, 2010, *34*, 129–138.

290 J. Woo, S.-J. Shin, W. Seo and P. Melanitasari, *Int. J. Adv. Manuf. Technol.*, 2018, *99*, 2193–2217.

291 N. A. Helal, O. Elnoweam, H. A. Eassa, A. M. Amer, M. A. Eltikhy, M. A. Helal, H. A. Fayyaz and M. I. Nounou, *Pharm. Pat. Anal.*, 2019, *8*, 139–161.

292 C. Badman, C. L. Cooney, A. J. Florence, K. Konstantinov, M. Krumme, S. Mascia, M. Nasr and B. L. Trout, *J. Pharm. Sci.*, 2019, *108*, 3521–3523.

293 S. Laske, A. Paudel, O. Scheibelhofer, S. Sacher, T. Hoermann, J. Khinast, A. L. Kelly, J. Rantannen, O. Korhonen and F. Stauffer, *J. Pharm. Sci.*, 2017, *106*, 667–712.

294 V. Vanhoorne and C. Vervaet, *Int. J. Pharm.*, 2020, *579*, 119194.

295 *Global Pharmaceutical Continuous Manufacturing Technology Market: Snapshot*, [https://www.transparencymarketresearch.com/pharmaceutical-continuous-manufacturing-technology-market.html](https://www.transparencymarketresearch.com/pharmaceutical-continuous-manufacturing-technology-market.html).

296 S. P. Simonaho, J. Ketolainen, T. Ervasti, M. Toiviainen and O. Korhonen, *Eur. J. Pharm. Sci.*, 2016, *90*, 38–46.

297 S. Lakio, T. Ervasti, P. Tajarobi, H. Wikstrom, M. Fransson, A. Karttunen, J. Ketolainen, S. Folestad, S. Abrahamsenalami and O. Korhonen, *Eur. J. Pharm. Sci.*, 2017, *109*, 514–524.

298 V. Pauli, Y. Roggo, L. Pellegratti, N. Q. N. Trung, F. Elbaz, S. Ensslin, P. Kleinebudde and M. Krumme, *J. Pharm. Biomed. Anal.*, 2019, *162*, 101–111.

299 F. Stauffer, V. Vanhoorne, G. Pilcer, P. Chavez, C. Vervaet and T. De Beer, *Int. J. Pharm.*, 2019, *11*, 118525.

300 B. Nagy, A. Farkas, K. Magyar, B. Demuth, Z. K. Nagy and G. Marosi, *Eur. J. Pharm. Sci.*, 2018, *123*, 10–19.

301 N. O. Sierravega, A. D. Romanospino, J. V. Scicolone, F. J. Muzzio, R. J. Romanach and R. Mendez, *Int. J. Pharm.*, 2019, *560*, 322–333.

302 Y. Roggo, V. Pauli, M. Jelsch, L. Pellegratti, F. Elbaz, S. Ensslin, P. Kleinebudde and M. Krumme, *J Pharm. Biomed. Anal.*, 2020, *179*, 112971.

303 R. Singh, A. Sahay, F. Muzzio, M. Ierapetritou and R. Ramachandran, *Comput. Chem. Eng.*, 2014, *66*, 186–200.

304 S. Zomer, J. Zhang, S. Talwar, S. Chattoraj and C. Hewitt, *Int. J. Pharm.*, 2018, *547*, 506–519.

305 A. F. Silva, J. Vercruysse, C. Vervaet, J. P. Remon, J. A. Lopes, T. De Beer and M. C. Sarraguça, *Eur. J. Pharm. Biopharm.*, 2018, *128*, 36–47.

306 A. D. Romanospino, R. Singh, M. G. Ierapetritou, R. Ramachandran, R. Mendez, C. Ortegazuniga, F. J. Muzzio and R. J. Romanach, *Int. J. Pharm.*, 2016, *512*, 61–74.

307 E. M. Hetrick, Z. Shi, L. E. Barnes, A. W. Garrett, R. G. Rupard, T. T. Kramer, T. M. Cooper, D. P. Myers and B. C. Castle, *Anal. Chem.*, 2017, *89*, 9175–9183.

308 V. Pauli, P. Kleinebudde and M. Krumme, *Eur. J. Pharm. Biopharm.*, 2020, *153*, 200–210.

309 K. Taipalekovalainen, A. Karttunen, J. Ketolainen and O. Korhonen, *Eur. J. Pharm. Sci.*, 2018, *115*, 1–10.

310 L. Manley and Z. Shi, *Int. J. Pharm.*, 2018, *551*, 60–66.

311 B. Van Snick, J. Dhondt, K. Pandelaere, J. Bertels, R. Mertens, D. Klingeleers, G. Di Pretoro, J. P. Remon, C. Vervaet and T. De Beer, *Int. J. Pharm.*, 2018, *549*, 415–435.

312 Y. Wang, T. O’Connor, T. Li, M. Ashraf and C. N. Cruz, *Int. J. Pharm.*, 2019, *569*, 118551.