USING MULTIVARIATE ANALYSIS FOR PHARMACEUTICAL DRUG PRODUCT DEVELOPMENT

by

Yifan Wang

A dissertation submitted to the

Graduate School – New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Chemical and Biochemical Engineering

written under the direction of

Fernando J. Muzzio and Benjamin J. Glasser

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October 2016

ABSTRACT OF THE DISSERTATION
Using Multivariate Analysis for Pharmaceutical Drug Product Development

By Yifan Wang

Dissertation Directors:

Fernando J. Muzzio and Benjamin J. Glasser

Manufacturing of pharmaceutical products has a prominent role in the healthcare industry. Generally, the ultimate aim of pharmaceutical development is to release to the market products with acceptable quality. As advanced pharmaceutical manufacturing technologies such as continuous tablet manufacturing, are developed and embraced, it is essential to adopt a scientific, risk-based, and proactive approach for pharmaceutical development. The work presented in this dissertation focuses on using multivariate analysis tools to establish a predictive capability for pharmaceutical process and product development, especially when the amount of materials available is limited. Importantly, the methodologies developed in this dissertation can be applied easily to powder handling and processing in a wider range of industries, such as cosmetic, catalyst, chemical, petrochemical, and food.

In this work, methods for analyzing flow properties of raw materials and predict process performance were developed. A method to analyze shear cell data of powders measured under different initial consolidation stresses was introduced. The method was shown to reduce significantly the complexity of shear cell data, and to enabled comparison of materials measured under different initial consolidation stresses. In addition, a predictive correlation between material flow properties and feeder performance was developed. By

using multivariate models, the feeding performance of a material with given flow properties can be predicted and quantified.

Using a quality-by-design approach, the cohesion of a powder mixture can be predicted based on the concentration of each ingredient. The prediction model was further supplemented by a study investigating two mixing systems. Using statistical analysis, the effect of lubrication on blend flow properties was discussed. By quantifying the correlations between different flow property measurements, mixing systems that have different mixing mechanism were compared.

Disadvantages of widely used dissolution comparison methods were addressed. Statistically reliable methodologies to analyze, compare, and predict drug *in vitro* release profiles were proposed. The proposed methods were shown to be able to consider the self-correlated intrinsic nature of dissolution profiles, and to use within-group variability to estimate the reliability of observations. Additionally, the work presented a case study to improve real-time release testing for advanced tablet manufacturing processes by achieving predictive capability for nondestructive dissolution testing. Using hierarchical multivariate analysis, the validated prediction models were able to predict dissolution profile of an individual tablet based on its NIR spectrum.

Acknowledgements

Undertaking a PhD has been a truly life-changing experience for me, and I have been fortunate to have the support and guidance from many individuals. I would like to thank my thesis advisors, Prof. Fernando Muzzio, and Prof. Benjamin Glasser for their encouragement, mentoring, and trust. I would especially like to thank Fernando for always giving me opportunities to make mistakes, to learn, and to challenge myself, and Ben for his insight and dedication to improving my writing and presenting skills. I would like to thank Dr. Ronald Snee and Prof. German Drazer, for their suggestions and inputs.

I feel grateful for being involved in research projects of the National Science Foundation Engineering Research Center for Structured Organic Particulate Systems (C-SOPS), and the Catalyst Manufacturing Consortium (CMC). The interactions with industrial and academic collaborators have been invaluable and insightful. It also helped me to better understand and address industry needs for my dissertation.

It has been my privilege to be able to work with so many talented colleagues, who have made the course of my PhD studies filled with enjoyment, friendship and love: Min Guo, Sara Koynov, Bill Engisch, Juan Osorio, Sarang Oka, Pallavi Pawar, Krizia Karry, Sebastian Escotet, James Scicolone, Golshid Keyvan, Wei Meng, Zhanjie Liu, and Tianyi Li.

This dissertation is dedicated to my family for selflessly raising me and supporting me all the way: my parents, Guojie Wang and Shan Li, my grandpa, Fusan Li, and my grandmas, Yihui Yang and Guizhen Wang, and my aunt Lan Li. They have given me their endless love, for always, and no matter what.

Table of Contents

List of Tables

List of Figures

1. Introduction

1.1. Background

Growth of the pharmaceutical industry is highly correlated with general economic strength, health care spending levels, and regulatory environments.[1] During 2014-2018, global healthcare spending is expected to accelerate by an annual average of 5.2 percent, rising to 9.3 trillion USD, as the global economy continues to recover from the Great Recession.[2] In the US, the growth of the pharmaceutical industry has been partly driven by expanded consumer access to health care through the 2010 Patient Protection and Affordable Care Act (PPACA).[3] Meanwhile, as brand drugs came off-patent and generic products reduced profit margins, global competition in pharmaceuticals has been increased.[4]

Manufacturing of pharmaceutical products has an increasingly prominent role in the pharmaceutical industry.[5] In 2016, advanced pharmaceutical manufacturing has been recognized as one of the nation's priority technology areas by the US federal government.[6] In response to increasing competitions and potential opportunities, innovation and cutting edge knowledge has been explored and applied to pharmaceutical manufacturing.[7]

The regulatory perspective on pharmaceutical manufacturing has evolved significantly in the last two decades. Traditionally, pharmaceutical manufacturing operations are inefficient and costly.[8] While opportunities for improving efficiency and quality assurance through an engineering approach are often not well acknowledged,[9] the "desired state" of pharmaceutical manufacturing, defined by US Food and Drug Administration (FDA), emphasizes the improvement of knowledge on design and the

understanding of product and process and includes mention of many traditional engineering tools such as modeling and distributed process control. Regulatory agencies, including US FDA and European Medicines Agency (EMA), have adopted Quality by Design (QbD) paradigm. The QbD initiative requires a scientific, risk-based, and proactive approach to pharmaceutical development. Quality of a product is not assured by testing, but rather built into the design effort from product conception through commercialization.[10] In QbD paradigm, critical quality attributes (CQAs) are defined as physical, chemical, biological, or microbiological properties or characteristics that must be controlled directly or indirectly to ensure the quality of the product.[11] These CQAs are usually controlled by understanding and manipulating the effects of critical material attributes (CMAs) or critical process parameters (CPPs).

In the prevailing economic and regulatory environment, solid dosage forms (e.g. tablets and capsules) comprise roughly 80% of all pharmaceutical products for US consumption.[7, 12] In every tablet, active pharmaceutical ingredients (APIs) are formulated with excipients to generate the desired manufacturability and drug release attributes of the final drug products. The manufacturing process usually involves a series of unit operations, such as mixing, granulation, drying, milling, and tableting/capsule filling.[13] In a QbD perspective, the quality of the final products is heavily dependent on raw and intermediate material properties, formulation development, and process understanding.[11]

In terms of tablet manufacturing, the exact sequence of unit operations may vary depending on formulation of the drug. However, in general, tablet production is carried out by three main processes: direct compaction, wet granulation, and roller compaction.

Usually, a process starts with feeding API, excipients and lubricant into the manufacturing line. Optionally, the material may be processed in a Comil for delumping prior to mixing. In the direct compaction route, materials proceed directly to tableting after mixing. Alternatively, the blend can be granulated by wet or dry granulation. In wet granulation, a granule drying step is needed before further processing. When granulation is involved, an additional milling step may be required to reduce the granule size before tableting, followed by a second blending step as additional extragranular ingredients are added. A schematic illustration of tablet production using advanced continuous manufacturing is shown in **Figure 1-1**.

1.2. Motivation

Often, during process and product development, the amount of material required for testing is often used sparingly and limited to a few hundreds of grams due to cost of material, availability, and safety concerns. During early tablet formulation development, the amount of API available is often small.[14] For example, less than 10g of API is available for conducting drug-excipient compatibility studies. During early drug product formulation and manufacturing process development, the amount of API is often limited to 0.2-1.0 kg. Therefore, an approach addressing material sparing is highly encouraged to save resources and time, and to improve efficiency of dosage form development. Specifically, this approach should employ i) use of experimental design concepts to perform experiments and obtain data that lead to valid conclusions. ii) selection of material characterization techniques that minimizes required amount of materials for testing. iii) establishment of predictive models using reliable data analysis methods.

The ultimate prerequisite to ensure consistent product quality is the ability to understand and manage the effects of material properties and to monitor and control the manufacturing process continuously.[15] Multivariate tools for product and process design, data acquisition and analysis has been recognized by FDA as part of Process Analytical Technology (PAT) tools to enable process understanding for scientific, risk-managed pharmaceutical development, manufacturer, and quality assurance.[16] Methodological experiments based on statistical principles provide effective approaches to identify and understand the effects of formulation and process variables, as well as their interactions for final product quality. Multivariate mathematical approaches, in combination with implemented PAT tools, can be used to predict product quality and performance. The established prediction power helps to quickly identify process design space and to avoid potential failure modes. Importantly, knowledge obtained using this approach is self-updating during development, validation, and commercial manufacturing phase, which is able to support product life cycle management, and future development products.[17]

The work presented in this dissertation is motivated by the need to establish predictive capability for pharmaceutical process and product development, especially when the amount of materials available is limited. Specifically, raw material flow properties and their predictive correlations with gravimetric feeding performance are discussed in Specific Aim I. Shear cell test has been extensively used to characterize powder flow properties. While a number of studies have focused on simulating shear cell testing dynamics, examining the effects of various factors on the results, and correlating shear cell data with other characterization techniques,[18-20] there are very few published

studies that have developed analytical methodologies to reduce the complexity of shear cell data. Specifically, since the measurement is dependent on the initial consolidation stress, a material is usually characterized under different levels of initial consolidation stress in order to generate a family of curves from which parameters can be extracted. It is also quite common to develop a material library of the flow properties of various powder materials, which leads to a highly dimensional database which can be difficult to use and interpret. Meanwhile, as powder flow is a complex behavior and there is no unifying index to describe powder flowability, the data complexity often increases to higher degrees when multiple characterization techniques are used for testing. When the flow properties of a material are characterized by varying measurements, a multivariate approach to analyze data and extract information is needed for comparison between different materials. Such methods are required, for example, to identify materials with matching properties so that placebo or surrogate materials can be used during formulation development, saving money and reducing risk to personnel. More importantly, a method that develops models for predicting process performance based on material properties is needed for process design, development, optimization, and control.

For Specific Aim II, the powder mixing process and its effect on the flow properties of intermediate blends were studied. It has been known for decades that the intensity of the mixing process for blends with MgSt may impact blend flow properties significantly. Although lubrication effects have been widely studied, there is no published work that systematically quantifies and compares the lubrication effect in different mixing devices. Questions also remain unanswered regarding how to correlate material properties. Here, two mixing systems that have different mixing principles are compared. In addition, the

challenge of predicting blend flow properties based on the concentration of each ingredient needs to be addressed. Although attention has been paid to characterizing and predicting flow properties of mixtures, very few published cases have been presented to address the scenario during early formulation development when the amount of drug is limited and more than two ingredients are used in the formulation.

Specific Aim III is devoted to developing methods to analyze, compare, and predict tablet dissolution performance, which is one of the critical quality attributes of the finished solid dose products. The ability to compare in-vitro dissolution profiles and demonstrate similarity of dissolution profiles for tablets and other oral products is extremely important to the pharmaceutical industry.[21] The f_1 (difference factor) and f_2 (similarity factor) method has gained considerable popularity and is widely used to guide similarity decisions, mostly because of its simplicity and its adoption by various regulatory agencies. However, in spite their popularity, studies have shown that these indices have a dubious meaning, unknown reliability, and also poor sensitivity in a number of important situations.[22] Therefore, alternative methods are needed for comparing dissolution profiles that not only are significantly more rigorous and meaningful than f_2, but also enable the practitioner to assess the reliability of results using standard statistical metrics. Additionally, the ability to nondestructively predict *in vitro* drug release profiles is critical for implementing real-time release testing.[23] Previous studies have extensively studied the effects of formulation and process variable on dissolution performance.[24, 25] However, very few studies reported nondestructive methods for predicting entire dissolution curve, considering both formulation and process variables. Therefore,

systematic methods of using appropriate multivariate analysis tools to analyze and predict dissolution profiles are needed and are presented here.

1.3. Significance to other industries

The methodologies demonstrated in this dissertation are not limited to the pharmaceutical industry. Powder handling and processing are ubiquitous in a wide range of industries, such as cosmetic, catalyst, chemical, petrochemical, and food.[26] However, powder-based processed development, scale-up, and transfer are still based on empirical knowledge. Properties of powder and granular materials can be changed significantly during the manufacturing process, affecting final product characteristics. For example, flow properties of powder and granular materials have great impacts on process stability and product uniformity. Gravimetric powder feeding and powder mixing discussed in this dissertation are common unit operations in many industries such as oil refining, natural gas processing, chemical synthesizing, etc.[27] Understanding the interaction between material properties and process performance, as well as establishing predictive models, is powerful for robust and reliable process design and development. Therefore, the predictive modeling developed in the dissertation, the concept of Quality-by-Design, and use of multivariate analysis can be applied and transferred to other industries, offering a systematical approach for process and product development.

1.4. Scope of this dissertation

Given the preceding discussion, the three specific aims of this dissertation are summarized as follows:

- Specific Aim I: Flow properties of raw materials and their predictive correlation with feeder performance

- Specific Aim II: Powder mixing and predictive flow properties of intermediate blends

- Specific Aim III: Methods to analyze, compare, and predict drug *in vitro* release profiles

Each of the six main chapters in this dissertation develops a specific method and demonstrates its role in the development of a robust manufacturing process. Chapter 2 gives an introduction and overview of the statistical and multivariate analysis tools that are extensively practiced throughout the dissertation. Chapters 3 and 4 address flow properties of raw materials. In Chapter 3, a mathematical method to analyze shear cell data of powders measured under different initial consolidation stress is described. In Chapter 4, a predictive correlation between material flow properties and feeder performance is discussed using a multivariate approach. Chapter 5 and 6 look into flow properties of powder mixtures. In Chapter 5, a statistical mixture model is built. The model predicts cohesion of a four-component blend based on flow properties and concentration of each raw ingredient. Chapter 6 demonstrates use of experimental design and statistical methods to study mixing effect on blend flow properties. The lubrication effect during mixing on blend flow properties will be quantified and compared in two different mixing systems. Chapter 7 and 8 focus on drug in-vitro release performance. Chapter 7 addresses the need to use statistical methods to compare dissolution profiles. In Chapter 8, a case study of continuous direct compaction line is presented. Methods are

shown to predict tablet dissolution profiles using a combination of chemometrics and dissolution profile analysis.

1.5. Figures for Chapter 1
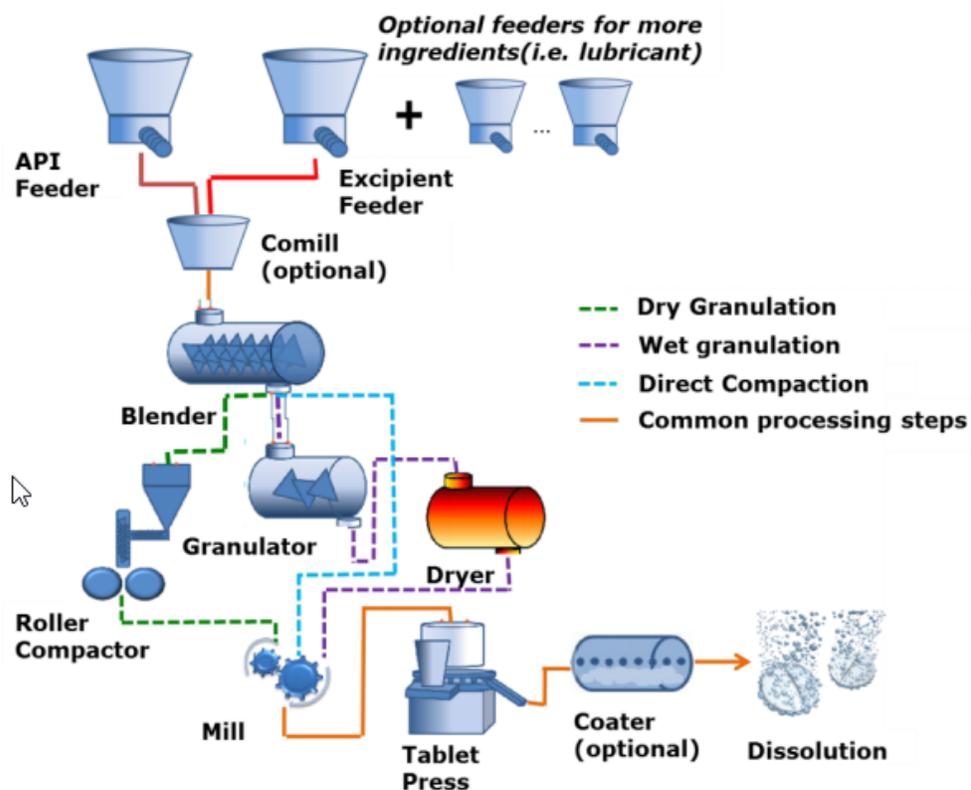
Figure 1-1 A schematic illustration of continuous tablet manufacturing using direct compression, wet granulation, and roller compaction. (Picture source: Sebastian Escotet, Fernando Muzzio.)

2. Overview of data analysis methods

2.1. Introduction

Data-drive models are effective tools to describe and predict data patterns. Formulation of statistical and data analytical problems, followed by proper data inference, is necessary for a comprehensive approach to process and product development.[28] Most common practices during pharmaceutical development seek underlying theoretical guidance in statistics.[29] For example, the use of experimental design concepts is common for screening, characterizing, and optimizing effects of critical variables on process performance and product quality.[30-32] For quality control and validation purposes, the sample size determination is strictly based on statistical hypothesis testing, population distribution models, and power analysis.[33, 34] Chemometrics, a critical enabling component of process analytical technology, is widely used for inline/online process monitoring and real time release testing.[35-37] Often, by using experimental design and statistical analysis, both effort and experimental error can be significantly reduced. In addition, the use of multivariate analysis has been widely used to explore datasets that involve multiple inputs or outputs, or both, especially when there is potential correlation among multiple variables.[38]

In this chapter, the data analysis methods used and demonstrated throughout the remaining chapters are introduced. The methods are classified into three categories: ANOVA-based methods, principal component analysis (PCA) -based methods, and regression methods. ANOVA-based methods test the statistical significance of observed effects by quantifying and comparing sources of variability, often estimated based on

experimental design.[39] PCA-based methods and regression methods explore, discover, and recognize the potential correlations in the dataset using machine learning algorithms, and use data to train models to learn the correlations and predict future performance. The difference between the two approaches is the procedure to identify input-output function in regression methods.[40]

2.2. ANOVA-based methods

2.2.1. ANOVA

Analysis of Variance (ANOVA) is used to test difference between two or more groups by comparing the variation between groups and variation within groups. When there is only one explanatory variable, for example, ANOVA measures the total amount of variability among data observations and splits the total variability into parts. The total variability can be expressed as:

$$\text{SST (sum of squares total)} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$$

The total variability has two sources: the variability between the group means, specifically, the variation around the overall mean, $\bar{x}$, expressed as:

$$\text{SSG} = \sum_{i=1}^{k} n_i (\bar{x}_i - \bar{x})^2$$

and the variability within the group means, or the variation of observations about the mean of individual groups, $\bar{x}_i$:

$$\text{SSE} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2 = \sum_{i=1}^{k}(n_i - 1)s_i^2$$

Where k is the number of groups, or levels of treatment, n_i is the sample size of group i, x_{ij} is the jth response sampled from the ith group, $\bar{x}_i$ is the sample mean of responses from the ith group $(\frac{1}{n_i}\sum_{j=1}^{n_i}x_{ij})$, s_i is the sample standard deviation from the ith group $(\frac{1}{n_i-1}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2)$, n is the total sample size $(\sum_{i=1}^{k}n_i)$, and $\bar{x}$ is the mean of all responses $(\frac{1}{n}\sum_{ij}x_{ij})$.

Based on the definition, an ANOVA table can be generated as shown in **Table 2-1**. Based on F statistic, in general, if the variability between groups is large relative to the variability within groups, it suggests that the means of the populations from where the data for the groups were drawn are significantly different.[41] If F is large, the variability between treatment levels (groups) is large relative to the variation within treatment levels, and therefore the null hypothesis of equal means should be rejected. If F is small, the variability between treatments is small relative to the variation within treatments, and the sample data is thus consistent with the null hypothesis that population means are equal between groups.

As an extension of one-way ANOVA, when two or more independent variables are involved, ANOVA assesses both the main effect of each independent variable, and any possible interactions between them. Use of ANOVA-based methods in different experimental designs has been extensively detailed by Montgomery.[42] In Chapter 6, two-way ANOVA is used to investigate the effect of mixing on blend flow properties.

Important to notice, ANOVA is based on several assumptions: i) The dependent variables (measured responses) should be continuous variables, as opposed to categorical variables. ii) The observations should be independent. In other words, there is no relationship between the observations in each group or between the groups. iii) The sample should be representative of the population, namely, there should be no significant outliers. iv) The variances between groups are equal. v) The errors are normally distributed.[43]

2.2.2. ω^2 effect size statistic

Calculation of effect size, which is often neglected in ANOVA-based analysis, measures the relative size of the effect or an interaction. Since the statistical significance of an effect, as described by its p-value, depends on both the effect size and the sample size, knowing the magnitude of an effect conveys information on practical significance, which cannot be obtained based on the p-value. A number of different statistics have been described in the literature to estimate effect size.[44] The Omega-squared index is an unbiased estimate of the proportion of variance in the population that is explained by a given treatment (or by an interaction).[45]

$$\omega^2 = \frac{SSG - df_{group}*MSE}{SST + MSE} \quad (3)$$

where SSG is the sum of squares for the treatment group, df_{group} is the degree of freedom of the treatment group, MSE is the mean squared error that can be referred to **Table 2-1**, and SST is the sum of squares total.

The values of omega squared are between 0 and 1 and often referred to the proportion of variance accounted by treatment. When F < 1, omega-squared becomes negative but can

be interpreted as $\omega^2=0$.[46] The use of ANOVA and effect size is further illustrated in Chapter 6 and Chapter 7.

2.2.3. MANOVA repeated measures

During pharmaceutical solid processing, many of critical material attributes and critical quality attributes are profile data, as opposed to single-point data. Specifically, many of the profile data are also associated with repeated measurements. Repeated measurements refer to measurements that are taken repeatedly at selected intervals on the same sample.[47] For example, compressibility profiles, obtained from a powder flow compressibility test, measure change of the same powder bed density by varying applied normal stress. Droplet penetration profiles, used to measure the wettability of porous powder beds, are intrinsically time-longitudinal data describing the change of liquid droplet volume over time during liquid penetrating into a powder bed. Dissolution profiles, which also belong to time series data, are measured to describe the trend in drug release over time. Since all the measurements described here are repeatedly taken on the same subject, there is an intrinsically associated autocorrelation in the data that violates the ANOVA assumption of incidence independence.[47] Earlier attempts to analyze repeated measures data used split-plot model that assumes a normal distribution of random effects and random errors, and that residual random error is constantly correlated across the within-group treatment (applied normal stress, or time), which is known as the sphericity assumption.[48] However, this univariate approach is not able to consider directly the nature of repeated measures data, and thus has limited applications.[49]

Alternatively, multivariate approach combines a multivariate general linear model and the testing of linear hypotheses using n-dimensional variate vector observations.[50] A general expression for MANOVA repeated measures is:

$$y_{hi} = \mu_j + \gamma_h + e_{hi}$$

$$h = 1, \dots s\ groups$$

$$i = 1, \dots N_h\ samples\ in\ group\ h$$

$$j = 1, \dots n\ repetaed\ measures\ (normal\ stresses, or\ time\ points)$$

$$N = \sum N_h\ total\ sample$$

y_{hi} is an n×1 vector of ith sample in the hth group, μ is an n×1 vector for mean response at j[th] within-subject treatment, γ_h is an n×1 vector for between-group effect of hth group, and e_{hi} is an n×1 vector of residuals.[51]

MANOVA is able to conduct multiple comparisons and directly tests the effect of significance. It is also flexible and straightforward to interpret from summary statistics. **Figure 2-1** uses dissolution profile as an example to show three scenarios: (a) only shows time effect, (b) adds treatment effect, and (c) further includes time*treatment interaction effect, meaning that the treatment effect is dependent on the time. MANOVA repeated measures analysis algorithms are available in most major statistical software packages, and O'Brien and Kaiser illustrated an extensive procedure with more technical details.[52] In Chapter 7, use of MAVOVA and its advantages for dissolution profile comparisons will be further discussed.

2.3. PCA-based methods

Often in multivariate analysis, the goal is to find hidden structure and correlation between variables. Different data-mining techniques haven been developed to address a wide range of industries. For example, during marketing basket analysis, association models can be used to identify common co-occurrences among a list of possible event [53]. Feature extraction algorithms are commonly employed in pattern recognition and image processing.[54] Some of the most open-ended data mining techniques are clustering algorithms such as principal component analysis (PCA). The aim of PCA is to find and group data points with similarities. This has been used extensively when there are no obvious natural groupings, in which case the data may be difficult to explore.[55]

2.3.1.   Principal Component Analysis

Principal Component Analysis (PCA) is the general name for a clustering technique which transforms a number of possibly correlated variables into a smaller number of variables called principal components. It uses a vector space transform to reduce the dimensionality of large data sets.[56] Specifically, a model is used to represent a data set (X) in a reduced dimension (latent variable space) such that the major axes of variability are identified [57] by calculating eigenvectors of the co-variance matrix. The raw data set X can be decomposed using the eigenvectors as a new orthogonal basis (a linear transformation that is interpreted simply as a rotation of coordinates in parameter space), based on the equation below, onto a set of scores (T) and loadings (P), while the remaining variability is modeled as random error (ε).

$$X = TP^T + \varepsilon$$

The columns T represent scores in the projected space (the new "coordinates"); loadings P represent the weights/ significance of each variable in each dimension. As mentioned earlier, use of PCA is often motivated by the need of identifying hidden structures in large data sets, which will be further demonstrated in Chapter 4 where materials with similar flow properties can be identified using PCA.

2.3.2.　　M-PCA for profile analysis

Modified PCA (M-PCA) is a method to analyze and compare a group of curves in terms of their level and shape. **Figure 2-2** uses dissolution profiles as examples to draw empirical understanding for the two properties: (a) and (c) shows that level and "shape" of a line (for example, when comparing $t^{1/2}$-transformed data) can be varied by slope and intercept, (b) and (d) are generalized schematics illustrating the comparison of non-linear curves. Mathematically, the level of the ith profile can be defined as the average of release percentage across n time points:

$$\bar{y_i} = \frac{\sum y_{ij}}{n}$$

Snee proposed a method to calculate the shape of a response curve in which the level is first analyzed, and then a residual matrix (R) is constructed by subtracting the grand mean ($y_{..}$), row and column means ($y_{i.}$ and $y_{.j}$, respectively):

$$R_{ij} = y_{ij} - y_{..} - (y_{i.} - y_{..}) - (y_{.j} - y_{..})$$

which is used to examine profile shape by principal component analysis (PCA).[58] The advantage of the PCA approach is that it does not require sphericity assumptions. For profile data from repeated measurements, a direct PCA on raw longitudinal data will

necessarily confound shape with factor effects (also known as confounding phenomenon).[59] The modified-PCA is able to isolate shape effects into the residual matrix and thus has better accuracy and sensitivity. The derived values representing level and shape effects can then be tested by standard analysis of variance (ANOVA) methods. Using this approach, comparing time series profile is able to transform to comparing its representative descriptive values. Application of M-PCA to dissolution profiles comparison and prediction is discussed in more details in Chapter 7 and Chapter 8.

2.4. Regression methods

Regression methods are often used for developing predictive models.[60] The objective of predictive modeling is to determine the relationship between several x-variables (independent variables, or explanatory variables) and one or more y-variables (dependent variables, or response variables). This objective can be achieved by establishing a model to estimate the relationships between independent variables and one or more dependent variables. In general, a regression model relates response variable (s) y to a function of x and β:

$$y \approx f(X, \beta)$$

where y is the dependent variable(s), X is the independent variable(s), and β are the model coefficients. In general, establishing a regression model is used to find statistically significant model coefficients that correlate y and X such that the residual error is minimized. Depending on selection of X as model inputs, the following regression methods are further introduced.

2.4.1.   Multiple linear regression (MLR)

Assuming that there is a linear relationship between the dependent variables and the independent variable, and that the independent variables do not have multicollinearity,[61] the most common form of regression is multiple linear regression:

$$y = \beta_0 + \sum_k \beta_k x_k + \epsilon$$

where y is a dependent variable, x_k is the kth independent variable, β_0 and β_k are the coefficients, ε is the error term, or the residual. The coefficients can be estimated by a standard least squares fit minimizing the sum of squared residuals. Once the coefficients are determined from sufficient calibration samples, the dependent variable can be predicted based on the independent variables in the model. Although MLR is easy to implement and interpret in some scenarios, a major limitation of MLR is that when there are many highly collinear variables, the stability of the established multiple linear regression model will decrease, and the derived model coefficients may become difficult to interpret.[62]

2.4.2. Principal component regression (PCR)

The principal component regression combines the principal component analysis with multiple linear regression. In other words, instead of regressing the dependent variable on the explanatory variables directly, the principal components of the explanatory variables are used as regressors. Since principal components are orthogonal to each other, PCR is more robust, compared to MLR, in dealing with the multicollinearity problem when two or more of the explanatory variables are being collinear. [63] In addition, PCR reduces the dimensionality of the explanatory variables and lowers the effective number of coefficients to characterize the model. This is particularly useful when high-dimensional

data matrix is used as model input.[64] The use of PCR will be described in Chapter 8 for nondestructive dissolution testing.

2.4.3.   Partial least squares regression

A criticism against PCR is that the major principal components may model variation in the explanatory variables of little relevance to the response variables. In other words, the calculation of principal components does not take into account of response variables of interest. Partial least squares regression (PLSR), on the other hand, improves the regression strategy. PLSR is similar to PCR in many ways. Historically, PCR predates PLS, with the latter appearing in published literature around 1983.[65, 66] Both methods model an independent variable, or response variable, when there are a large number of explanatory variables, and those variables are potentially highly correlated. Both methods project the variables to the most dominating dimensions by constructing new predictor variables, or the principal components, as linear combinations of the original explanatory variables. The difference between the two methods is that PLSR takes the response variable into account when finding the principal components, while PCR does not consider the response variable.[67] In other words, for each component to be computed, PCR maximizes variance of the linear combinations of the explanatory variables, while PLSR additionally maximizes covariance between linear combinations of the explanatory variable and the response variables. The use of PLSR is demonstrated in case study presented in Chapter 4 to predict gravimetric feeding performance based on material flow properties.

A schematic comparison between MLR, PCR, and PLSR is shown in **Figure 2-3**. The comparison between PLSR and PCR has been conducted extensively in literature. Studies

showed that PLS used fewer latent variables than PCR because it uses a linear combination of the eigenvectors to construct its loading vectors.[68] PLS also offers a greater interpretability of the loading results.[69] However, this did not reflect better predictive ability. Remarkably, PLSR can also be used as a supervised classification method. The response variable can be considered as a binary vector containing only zeros and ones, which are used to describe classification for each sample. This method is an extension of PLSR, called as PLS-discriminant analysis.[70]

2.5. Multivariate analysis for pharmaceuticals: A literature review

The majority of multivariate analysis application in pharmaceutical industry focuses on developing chemometric models to predict relevant parameters and quality attributes. Ultimately, the efforts are to achieve real-time release testing capabilities. This was also driven by the initiative of the US FDA Process Analytical Technology (PAT) guidance, which proposed the transition from testing at the end of batch processing to timely measurement. Process analytical instruments usually record high-dimensional data, and require multivariate techniques to extract the required information from obtained data. For example, multivariate analysis approach has been extensively practiced in drug substance manufacturing to monitor nucleation and crystallization process using attenuated total reflectance-Fourier transform infrared (ATR-FTIR) and focused beam reflectance measurement (FBRM).[71-73] In drug product manufacturing, use of NIR or Raman spectra, coupled with multivariate analysis, has been commonly used to monitor blend uniformity during mixing,[74-76] to determine end-point for wet granulation,[77-79]   to monitor tablet content uniformity and predict dissolution and tensile strength,[80-83]   and to determine coating thickness during coating.[84, 85]

Besides chemometrics, there are other possible applications of multivariate analysis to support process and product development. For example, a PLSR model was built based on process data to identify an optimal operation space of a continuous flow hydrogenation process.[86] To model the relationship between API and intermediate properties of a hard-capsule formulation process, a PLSR was used to remove irrelevant variables without significantly impacting the predictive power of the model.[87] To predict material flowability, PLSR was used to develop a model that relates particle properties to bulk flow properties.[88] The use of PCA can be applied to statistical process control and monitoring. For example, new manufacturing batches can be projected onto the model that is based on previous validation batches, and the projected trajectory with its expected ranges can signify possible deviations.[89, 90] A previous study also showed that PCA applied to analyze early cell culture process was more powerful than univariate analysis to reveal relevant process features.[91, 92] Importantly, some multivariate analysis applications involve multi-block and multiway extensions of PCA/PLS methods. The multi-block PLS (MB-PLS) allows the analysis of multiple types of data simultaneously. Therefore, raw material attributes, intermediate material attributes, final product quality attributes, and process data from several unit operations can be correlated and modelled.[93] In a recent study, MB-PLS was used to study the relationships between different process steps and identify variables that have greater impact on final product quality.[94] The multiway extension, on the other hand, allows analysis with one additional dimension. This approach is common for longitudinal data such as dissolution profiles. Studies demonstrated that multiway PCA is an effective way for dissolution profile analysis to better reveal within- and between-batch differences.[25]

In summary, use of multivariate analysis is helpful for developing calibration models using spectroscopic tools, analyzing material properties, identifying critical material and process variables, and improving process understanding and monitoring. In chapter 3-8, experimental design, statistical, and multivariate analysis is extensively applied to correlate process variables and critical material attributes such as powder flow properties, and critical quality attributes such as table dissolution profiles. In chapter 3, efforts to establish a high-dimension material property database is described and a mathematical method for shear cell data analysis is introduced. Chapter 4 uses PCA and PLSR to correlate material flow properties with feeder performance. Chapter 5 uses mixture model, an extension of MLR, to predict flow properties of powder mixtures. Chapter 6 applies experimental design, ANOVA, and effect size test to quantify and compare lubrication effects in two mixing systems. Chapter 7 proposes statistical methods to analyze dissolution profiles based on MANOVA and M-PCA. Chapter 8 uses PCR and chemometric tools to predict dissolution profiles based on NIR spectroscopy.

2.6. Figures for Chapter 2
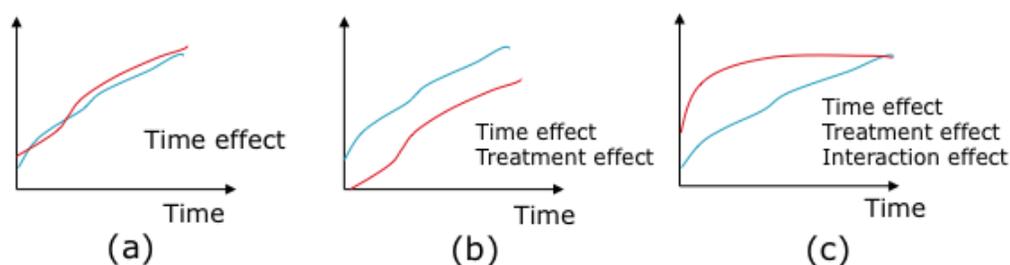
Figure 2-1 Schematic illustration for MANOVA repeated measures for dissolution profiles
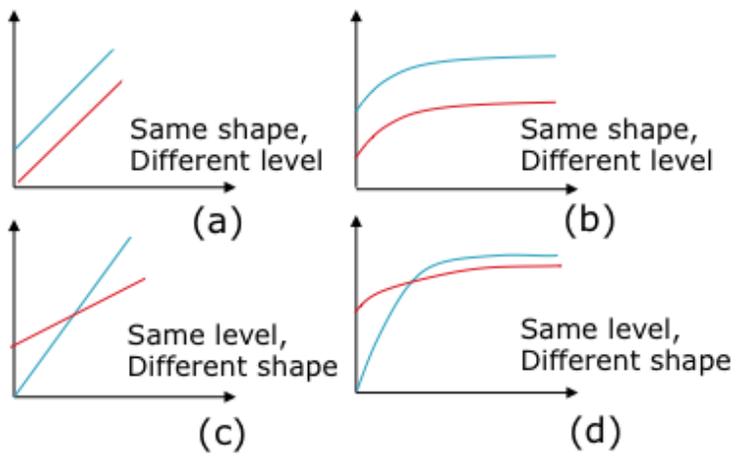
Figure 2-2 A schematic illustration on level and shape of a curve
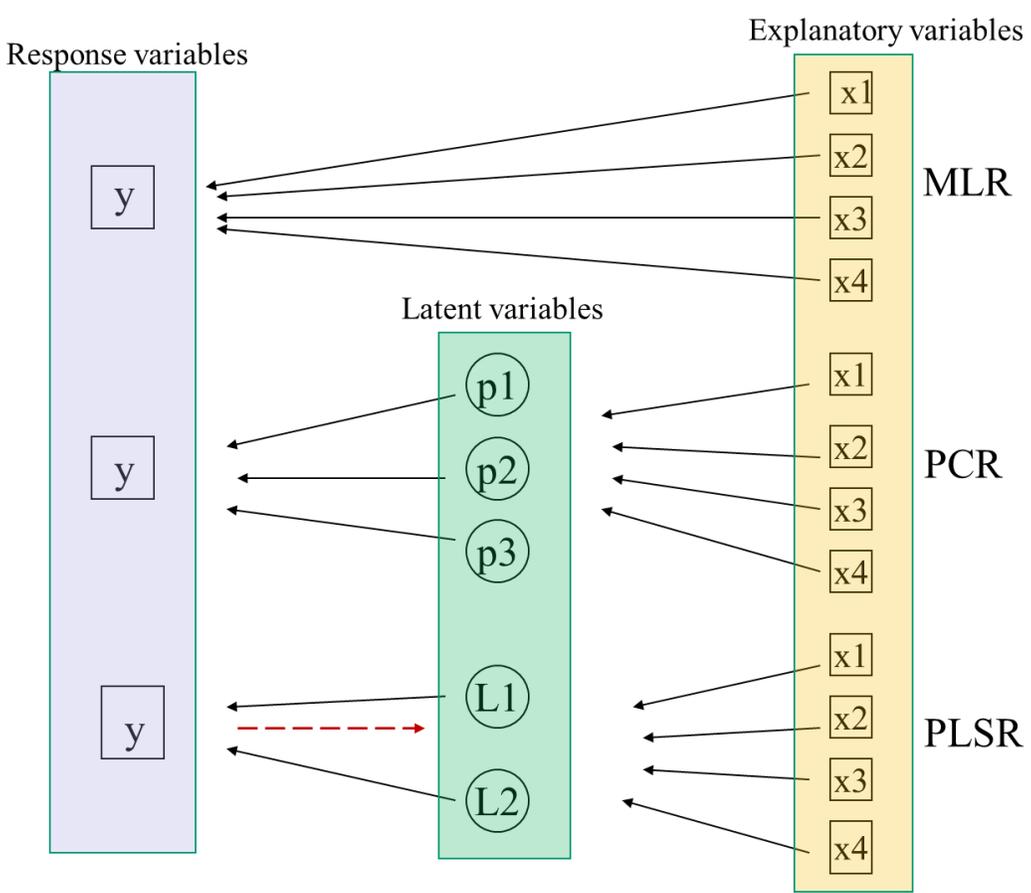
Figure 2-3 A schematic comparison of MLR, PCR, and PLSR. Both PCR and PLSR constructs principal components as linear combinations of the explanatory variable, and correlate the principal components to the response variables. The difference between PCR and PLSR is that PLSR considers the covariance between the response variable and the explanatory variables when computing the principal components.

2.7. Tables for Chapter 2

Table 2-1 Calculation of an ANOVA table using F-statistics.

Source	SS	df	MS	F
Model/Group	SSG	k-1	$MSG = \dfrac{SSG}{k-1}$	$\dfrac{MSG}{MSE}$
Residual/Error	SSE	n-k	$MSE = \dfrac{SSE}{n-k}$	
Total	SST	n-1		

3. A Method to Analyze Shear Cell Data of Powders Measured under Different Initial Consolidation Stresses

3.1. Introduction

Powders and granular materials are commonly handled and processed in a wide range of industries.[95] In the pharmaceutical industry, active pharmaceutical ingredients are mixed with excipient materials and processed through a series of powder-based unit operations (e.g, granulation, drying, milling, and compaction) in order to produce the final drug product.[96, 97] Similar powder processing is also widely practiced for cosmetic, catalyst, chemical, petrochemical and food products.[98-101]

Powder flow behavior, one of the most critical properties of powders and granular materials, can change significantly during various stages of processing,[102] and if the change is poorly understood, monitored, or controlled, process performance and product quality can be adversely affected.[7] Therefore, the ability to understand, measure, and predict powder flow at each process unit is vital to ensure controllable processes and reproducible products.[103] Ideally, this capability can be achieved by studying rigorously the flow of different materials in various types of process equipment. However, this can be complex and challenging because powder flowability, by definition, results from a combination of inherent material properties, environmental and process factors, and even choice of characterization method.[104-106]

In industrial practice, in order to predict the flow of a material in a given piece of equipment, or to compare the flow of a new material to an existing material, a practical way is to implement a standardized lab scale test to measure powder properties.[107] As

a result, many characterization techniques have been developed,[108] and efforts have been made to correlate and compare the results obtained in different measurement devices.[18, 109-112]

The shear cell methodology, which was originally developed by A. W. Jenike for design of hoppers and silos,[113] is one of the most common powder flow characterization techniques. It has been an engineering practice to classify and rank the flowability of different materials according to the extracted flow function coefficient (ffc).[114] The cohesion, the unconfined yield strength, the angle of internal friction, and other design parameters can be extracted from Mohr circle analysis.[115] The methodology has been widely applied for powder characterization, and the testing procedure has been detailed and documented as international standards.[116, 117] The shear cell test has thus been used extensively to characterize the flow properties of powders, and the effect of consolidation state on the powder flow properties has been examined.[118, 119] The effects of particle size, particle shape and the environmental factors, such as storage time, humidity and temperature, have been studied.[120-122] A recent study compared three commercially available shear cells in terms of powder flow properties.[123] In addition, multivariate analysis has been applied to analyze shear cell data along with other flow indices. For example, principal component analysis was used to develop a method for early pharmaceutical formulation development.[124] A partial least square regression model was established to predict powder flowability based on particle size and shape distribution.[88]

While a number of studies have focused on simulating the testing dynamics, examining the effects of various factors on the results, and correlating shear cell data with other

characterization techniques, there are very few published studies that have developed methodologies to reduce the complexity of the shear cell data. Specifically, since the measurement is dependent on the initial consolidation stress,[125] a material is usually characterized under different levels of initial consolidation stress in order to generate a family of curves from which parameters can be extracted. It is also quite common to develop a material library of the flow properties of various powder materials, which leads to a multi-dimensional database which can be difficult to interpret. In addition, comparing data measured from different shear cells is, in some cases, also needed. As the dimensionality of a shear cell database grows, challenges that inevitably need to be addressed are: 1) Correlating different flow parameters to reduce the complexity of the database. 2) Comparing different materials measured under different initial consolidation stresses. 3) Comparing and transferring the database between different types of shear cells.

The purpose of this chapter is to introduce such an analysis framework to address the above scenarios. The methodologies are demonstrated using an established shear cell database, which contains 9 catalyst support materials and 32 pharmaceutical materials measured under four different initial consolidation stresses. The remainder of this chapter is organized as follows: section 3.2 describes the materials in the database, the procedure for the shear cell test and the Mohr circle analysis used to extract the flow parameters; section 3.3 presents a correlation between different flow parameters, and methods to analyze the shear cell data for different materials, initial consolidation stresses, and testing devices; and finally, section 3.4 is devoted to conclusions.

3.2. Materials and Methods

The materials used in this study are listed in **Tables 3-1** and **3-2**. Particle size information and conditioned bulk density of the materials are listed in Table 3. The particle size of the materials was determined using a laser-diffraction (LS-13320) analyzer with a Tornado Dry Powder System (Beckmann-Coulter, Brea, California, US). Conditioned bulk density was measured after the standard conditioning cycle of the shear cell test, which is further detailed below.

3.2.1. Shear cell methodology

The flow properties of all blends were characterized using a rotational shear cell supplied as a component of the FT4 Powder Rheometer (Freeman Technology Inc., Worcestershire, UK), which is shown in **Figure 3-1**. The testing procedure consisted of four steps: conditioning, consolidation, preshearing, and shearing. The powder was first filled into a glass cylinder. During conditioning, a helical blade moved downwards, and then moved upwards to erase powder history and ensure a homogeneous reproducible state. In consolidation, a vented piston applied the initial consolidation stress (I) to the powder. The powder was then sheared to achieve a steady-state. The shear stress and normal stress were recorded and were indicated as a preshear point. After the preshear point was achieved, the normal stress was lowered and the sample was further sheared to obtain a yield point. The paired preshear-shear procedure was repeated 5 times in total at different normal stresses to obtain a yield locus.[116] Each test generated one preshear point and five yield points.

The shear stress was plotted in terms of the normal stress, which is also known as the τ-σ diagram. While models to take into account of shape of the yield locus have been developed,[126] a commonly used approach is the Mohr-Coulomb model to generate a

linearized yield locus. A schematic illustration of Mohr circle analysis for a yield locus is shown in **Figure 3-2**. Linear regression was used to fit the yield points to a mathematical expression for the linearized yield locus. The intercept of the yield locus on the shear stress τ -axis is defined as the cohesion (τ_1), which can be interpreted as the shear stress required to deform the powder when no normal stress is applied. The slope of the linearized yield locus is tan (φ), where (φ) is the called the angle of internal friction (see **Fig. 3-2**), which is a measure of ease with which the powder particles will slide past one another.[127] To extract other flow indices, a Mohr circle was then plotted passing through origin and tangent to the yield locus. The intercept of this Mohr circle with the normal stress σ-axis is defined as the unconfined yield strength (σ_c). A second Mohr circle was drawn tangent to yield locus and passing through the preshear point (σ_p, τ_p), with its intercept on the σ-normal stress axis called the major principal stress (σ_1). The flow function coefficient (ff_c), which is often correlated to the arching phenomenon in hoppers, is defined as the ratio between the major principal stress (σ_1) and the unconfined yield strength (σ_c):[128]

$$\text{ffc} = {\sigma_1}/{\sigma_c} \quad (1)$$

While these parameters are typically calculated and reported as independent measurements, they can, in fact, be calculated from the mathematical expression in the Mohr circle analysis. The Mohr-Coulomb model in the τ-σ diagram can be described as:

$$\tau = \sigma * \tan\varphi + \tau_1 \quad (2)$$

For a straight line with slope equal to tan(φ) and intercept (cohesion) τ_1, the unconfined yield strength can be expressed as a function of the cohesion and the angle of internal friction:

$$\sigma_c = \tau_1 * 2 * \tan(45 + \varphi/2) \quad (3)$$

For a circle tangent to the yield locus and passing through the preshear point (σ_p, τ_p), with its intercept on the σ-axis equal to the major principal stress (σ_1), one can show that the major principal stress can be related to the cohesion, the angle of internal friction, and the coordinate of the preshear point (σ_p, τ_p):

$$\sigma_1 = (1 + \sin\varphi)\left(\left(A - \sqrt{A^2\sin^2\varphi - \tau_p^2\cos^2\varphi}\right)\Big/\cos^2\varphi\right) - \tau_1/\tan\varphi \quad (4)$$

where $A = \sigma_p + \tau_1/\tan\varphi$

In this study, the normal stresses during the shear cell test, and Mohr-Couloumb model to extract flow properties were selected as are typical for material characterization purposes.[129] The four levels of initial consolidation stresses were 3 kPa, 6 kPa, 9 kPa, and 15 kPa for all materials studied. At the initial consolidation stress of 3kPa, the shear stresses to achieve incipient flow were obtained at normal stresses of 2 kPa, 1.75 kPa, 1.5 kPa, 1.25 kPa, and 1 kPa. Similarly, at the initial consolidation of 6 kPa, shear stresses under normal stresses of 4 kPa, 3.5 kPa, 3 kPa, 2.5 kPa, and 2 kPa were recorded. At the initial consolidation stress of 9 kpa, shear stresses under normal stresses of 7 kPa, 6 kPa, 5 kpa, 4 kPa, and 3kPa were obtained. For the initial consolidation stress of 15kPa, shear stresses at normal stresses of 9 kPa, 8 kPa, 7 kPa, 6 kPa, and 5 kPa were recorded. Before

obtaining each yield point, the materials reached a steady-state flow at the pre-defined preshear consolidation stress. Sampling replicates of 3 shear stresses were conducted at each normal stress.

3.3. Results and discussion

3.3.1.   Correlation between the flow function coefficient and the cohesion

The materials in the database covered a wide range of flow behavior. **Figure 3-3** examines the correlation between the flow function coefficient and the cohesion (kPa) for all the 41 materials at the same initial consolidation stress, and shows that the two parameters have an intrinsic inverse correlation (power law with exponent of -1). The inverse correlation at all the initial consolidation stresses is statistically significant ($p < 0.05$). The inverse correlation between the flow function coefficient and the cohesion is consistent with a previous study on solid mixtures.[130] Importantly, the pre-factor coefficient of the power correlation is a material-independent characteristic, and only depends on the initial consolidation stress. Moreover, as shown in **Figure 3-4**, the pre-factor coefficient is linearly proportional to the initial consolidation stress (Adjusted R-squared $= 0.996$, $p < 0.05$).

The flow function coefficient, which, by definition, is obtained from the intercepts of the two Mohr circles, can also be calculated through its correlation with the cohesion. To further investigate this correlation, a dimensionless cohesion, C*, was defined as:

$$C^* = {\tau_1 \, (\text{kPa})} \Big/ {I \, (\text{kPa})} \quad (5)$$

where I is the initial consolidation stress. As shown in **Figure 3-5**, the flow function coefficient for all materials, for all values of the initial consolidation stress, collapse to a

single curve when plotted as a function of the dimensionless cohesion C*, once again exhibiting inverse proportionality. The correlation between them is statistically significant (adjusted R-squared = 0.979, p < 0.05). The correlation can be described as:

$$ffc = {}^{E}\!/_{C^*} \quad (6)$$

where E is a constant. Based on the usual interpretation of the flow function, materials can now be classified based on the dimensionless cohesion, C*:

C* ≤ 0.048, free-flowing

0.048 < C* ≤ 0.121, easy-flowing

0.121 < C* ≤ 0.242, cohesive

C* > 0.242, very cohesive or not flowing

The distinction between very cohesive materials and non-flowing materials has not been made in this study because materials with C* higher than 0.484 were not tested. The advantage of this classification is that while the criteria based on the flow function coefficient does not take into account the initial consolidation stress, calculation of C* by definition includes the effect of the initial consolidation stress.

The proportionality constant E = 0.485 in the **Figure 3-5** is independent of both material characteristics and the initial consolidation stress, but it depends on geometric factors of the testing device during consolidation, and can be used as an equipment characteristic (or to compare measurements obtained in different devices). By means of a simple

logarithmic transformation, the correlation between the flow function coefficient and the dimensionless cohesion C* can be linearized, as shown in **Figure 3-6**.

Combining the experimental results with the mathematical relation introduced in section II, the constant E can be used to describe another material property:

$$\sigma_1/_I = 2 * E * \tan(45 + \varphi/_2) \quad (7)$$

where σ_1 is the major principal stress, which is the relative consolidation state when the material is in steady-state flow under the preshear normal stress; φ is the angle of internal friction, and the plane $(45° + \varphi/_2)$ to the major principal stress is the direction where the shear band with lowest deformation resistance is most likely to occur.[131] Therefore, the change of the major principal stress is related to the equipment characteristic, E, and the material property, φ. This is consistent with the previous studies. It has been observed that when powder reaches a steady-state, the consolidation state is a material property, and only depends on the applied stress.[132] In addition, previous researchers showed that the impeller torque, which is related to the shear stress, increased linearly with the fill level or applied normal stress.[133, 134]

3.3.2. Correlation between the unconfined yield strength and the cohesion

As described in section II, the ratio between σ_c, the unconfined yield strength (UYS) and the cohesion is only dependent on the angle of internal friction. Experimental results in this study show that the unconfined yield strength (σ_c) is linearly correlated with the cohesion. **Figure 3-7** shows the yield loci of maize starch measured under four initial consolidation stresses. By plotting the UYS as a function of the cohesion, all the yield points at different initial consolidation stresses can be described by a single line passing

through the origin. In other words, as the initial consolidation stress changes, the ratio between the change of the UYS and the change of the cohesion is constant. The linear correlation is statistically significant (Adjusted R-squared = 0.997, $p < 0.05$). The coefficient 3.303 is independent of the initial consolidation stress, and thus this slope can be considered to be a characteristic of the material and to be independent of the testing conditions for a particular shear cell.

Hereon, we describe the relationship between UYS and the cohesion as the "material characteristic line". **Figure 3-7** also shows that the characteristic line can be transformed back to the shear stress-normal stress (τ-σ) diagram to describe the linearized yield loci. In other words, while the UYS-cohesion plot (right) reduces the complexity of the shear cell data by collapsing all yield loci into a single line, it contains all the information needed, such as the intercept (τ_1) and the slope (φ), to transform back to the linearized yield loci family (left). In addition, the cohesion is highly correlated to the flow function coefficient, as discussed earlier in this section, and therefore the information regarding the preshear point, for example the major principal stress, is also stored in the UYS-cohesion plot. The material characteristic line for the individual materials examined is reported in the **Figure 3-8 and Figure 3-9**. For all the materials investigated, a linear relation between UYS and cohesion is obtained (which is statistically significant, $p < 0.05$).

The material characteristic line proposed here is consistent with previous studies. Teoman *et al.* investigated seven bulk solids and found that the angle of internal friction did not change significantly with initial consolidation stresses.[135] In addition, simulations have

demonstrated that the ratio between the shear stress and the normal stress is constant when the powder flow is in the elastic-quasistatic regime.[136]

The methodology used to generate the material characteristic line can be used to compare different materials. For demonstration, the yield points at all consolidation states and the material characteristic lines of the catalyst support materials in the database are shown in **Figure 3-10**. A yield point with lower value of the cohesion indicates lower inter-particulate cohesive forces during incipient flow, and a material with lower value of the slope generally experiences less frictional forces to flow.[127] While the material characteristic line describes a material at different consolidation states, the differences between the lines are independent of the initial consolidation, which enables the comparison of different materials tested under different initial consolidation stresses.

3.3.3.   Comparison of different shear cells

To summarize the previous discussion, the UYS – cohesion plot gives a material characteristic line that is independent of the initial consolidation stress, and the ffc – C* plot enables one to classify the material. This relationship can also be used to compare measurements obtained in different types of equipment. This section compares different types of rotational shear cells to further investigate the proposed methodology. Three commercially available rotational shear cells were compared in a previous study[123]: The RST-XS (Dietmar Schulze, Wolffenbuttel, Germany), the FT4 (Freeman Technology, Tewkesbury, Gloucestershire, UK), and the PFT (Brookfield Engineering Laboratories, Inc., Middleboro, MA, USA). The same coarse alumina and fine alumina materials described above and used in the present study were used in the previous study. 3 kPa, 6 kPa, and 9 kPa were selected as initial consolidation stresses for all the shear

cells, and an additional 15 kPa was used for both FT4 and Schulze shear cell. A previous study showed that the reproducibility of measurements of the free-flowing material is lower than that of cohesive materials.[123] The yield locus measured for the coarse alumina using the Brookfield shear cell had negative cohesion values at initial consolidation stresses of 6 kPa and 9 kPa, which was an artifact of the extrapolation of the linearized yield locus, and thus were not considered here.

Figure 3-11 shows that the equipment characteristic E of the Schulze (E=0.447) and the FT4 (E=0.485) shear cell are reasonably similar. The Brookfield shear cell, on the other hand, had different testing performance to achieve the steady-state flow. As shown in **Figure 3-12**, the angle of internal friction measured from the Schulze shear cell was larger than the one measured using the FT4 shear cell for both materials. The Brookfield shear cell gave similar results for the angle of internal friction to the FT4 shear cell for fine alumina. For both materials, the cohesion measured by the FT4 shear cell had a wider span as the initial consolidation stress increases. The UYS – cohesion plot suggests that the devices had different testing performance in achieving the incipient flow of materials as well. The difference in achieving both the steady-state flow and the incipient flow is consistent with the statistical analysis in the previously published paper.[123]

3.4. Conclusion

In this chapter, shear cell methodologies used to characterize powder flow properties were examined. A shear cell database containing 41 powder and granular materials at 4 different initial consolidation stresses was presented. The flow indices were extrapolated based on the Morh-Coulomb model to generate the linearized yield locus. In order to

reduce the complexity of the dataset, the correlation between different parameters was examined. For the cases studied here, an inverse correlation between the flow function coefficient and the cohesion was identified. The correlation pre-factor is material independent and increases linearly with the initial consolidation stress. A dimensionless cohesion, C*, was defined using the cohesion value and initial consolidation stress. The ratio between the flow function coefficient and C* was independent of the initial consolidation stress, and can be considered to be a testing equipment characteristic value, E. The characteristic E can also be used to describe the relationship between the applied normal stress and the characteristic state when materials are critically consolidated.

The mathematical relation between the unconfined yield strength and the cohesion is a function of the angle of internal friction. For the cases studied here, by finding that the angle of internal friction did not change significantly with the initial consolidation stress for all the 41 materials in the database, a material characteristic line was generated to collapse all the yield loci for a given material into a single line independent of the initial consolidation stress. The slope of the material characteristic line is a material property required to achieve incipient flow. Material characteristic lines can be used to compare different materials, even when measurements were taken under different initial consolidation stresses.

The methodology presented here augments the shear cell data analysis through understanding the relationships of the different measurements. The identified correlation can be used to compare different types of shear cells as shown in the case study presented here. For the cases studied here, the comparison suggests that shear cells have different

testing dynamics to achieve the incipient and the steady-state flow for different materials. The results presented here are for a limited set of materials and further work is needed to see if additional materials will follow the same behavior. If the observed behavior carries over to other materials, then the proposed methods could facilitate understanding of shear cell measurements and provide valuable ways to analyze and compare them. Based on the results presented in this chapter, the angle of internal friction will be used as an important material flowability input in the following chapters. Chapter 4 combines data from the shear cell test and other flow characterization techniques to predict feeding performance of raw materials. Chapter 5 supplements the proposed shear cell data analysis method by investigating powder mixture systems.

3.5. Figures for Chapter 3
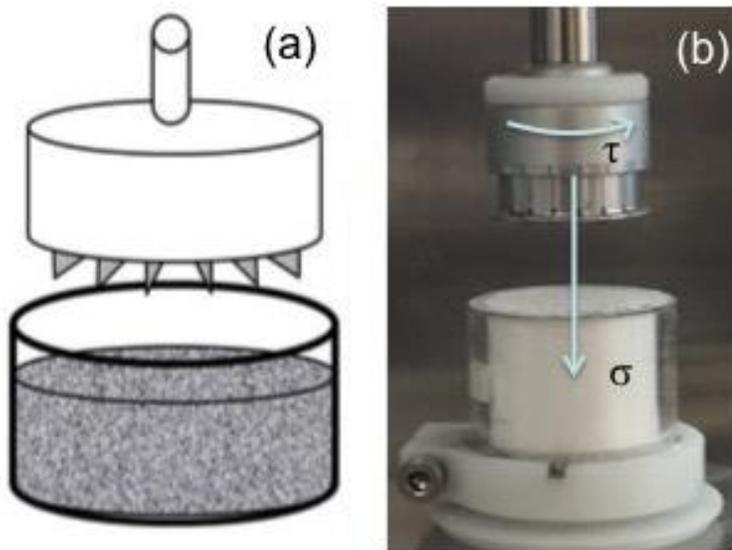
Figure 3-1 The schematic (a) and image (b) of the shear cell setup supplied with the FT4 Powder Rheometer
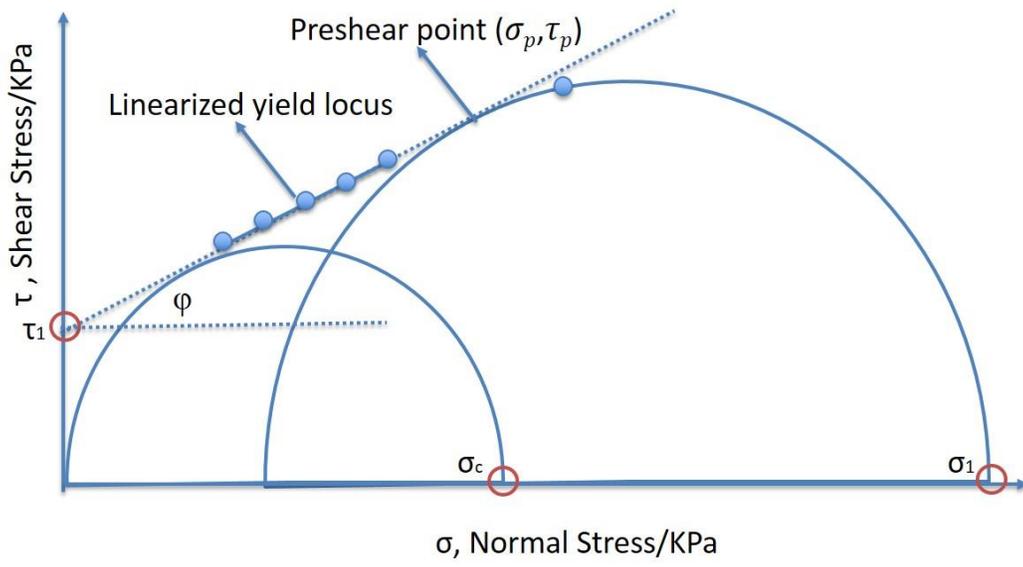
Figure 3-2 Mohr circle analysis for shear cell data.
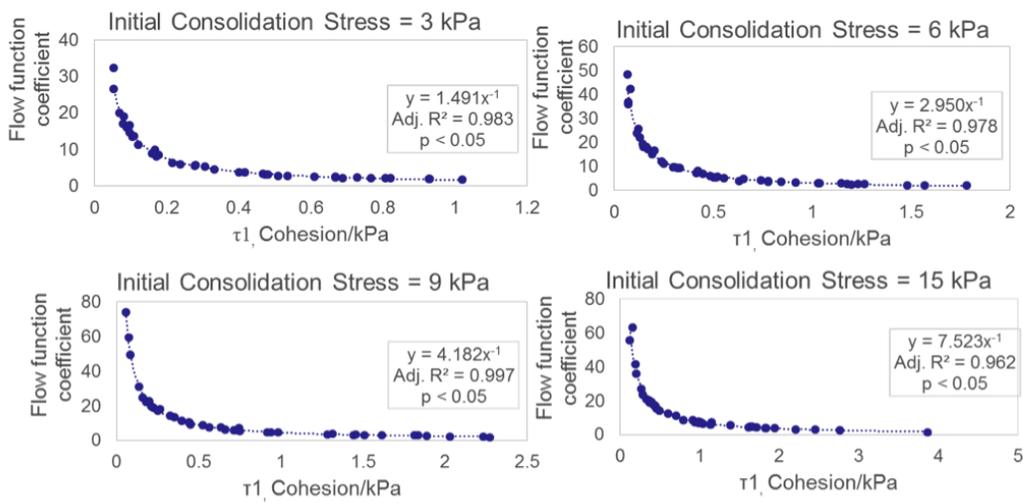
Figure 3-3 Correlation between flow function coefficient and cohesion (kPa) for 41 materials under four initial consolidation stresses.
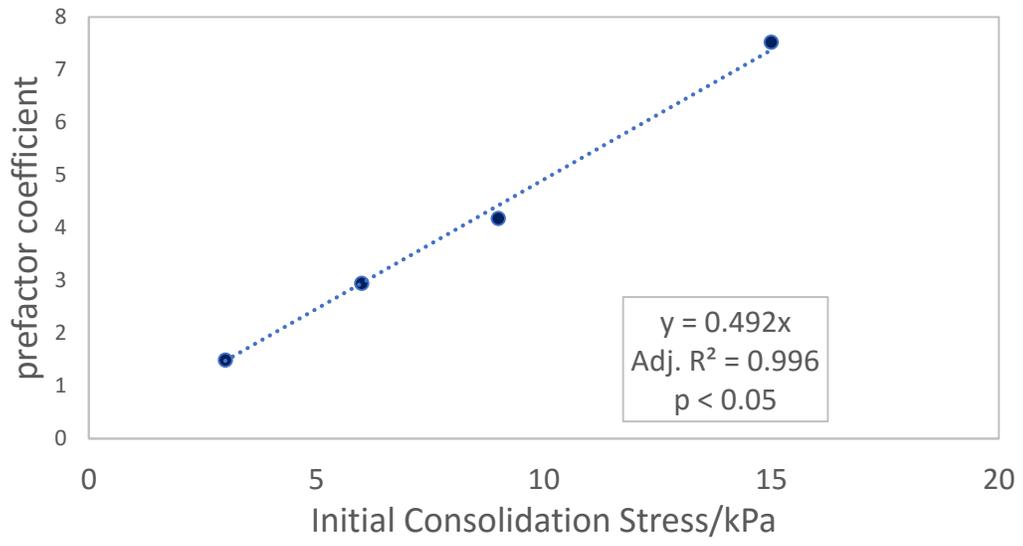
Figure 3-4 The power correlation coefficient of flow function coefficient and cohesion vs.the initial consolidation stresses.
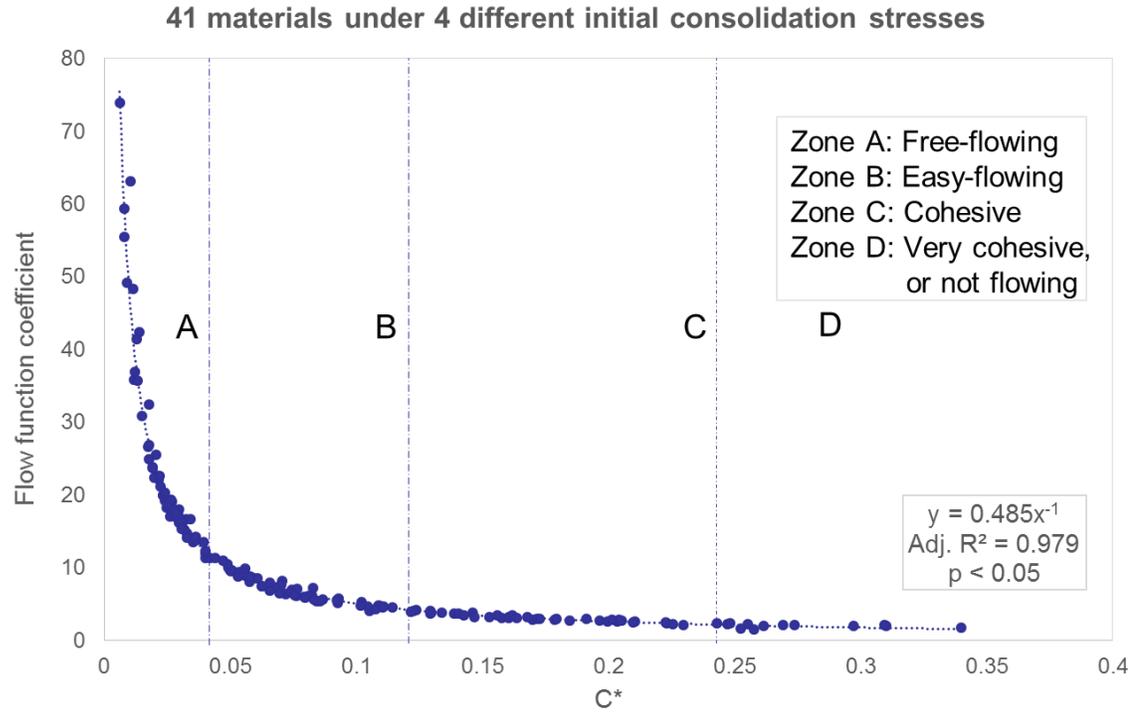
Figure 3-5 The correlation between the flow function coefficient and log of C* for 41 materials under 4 initial consolidation stresses.
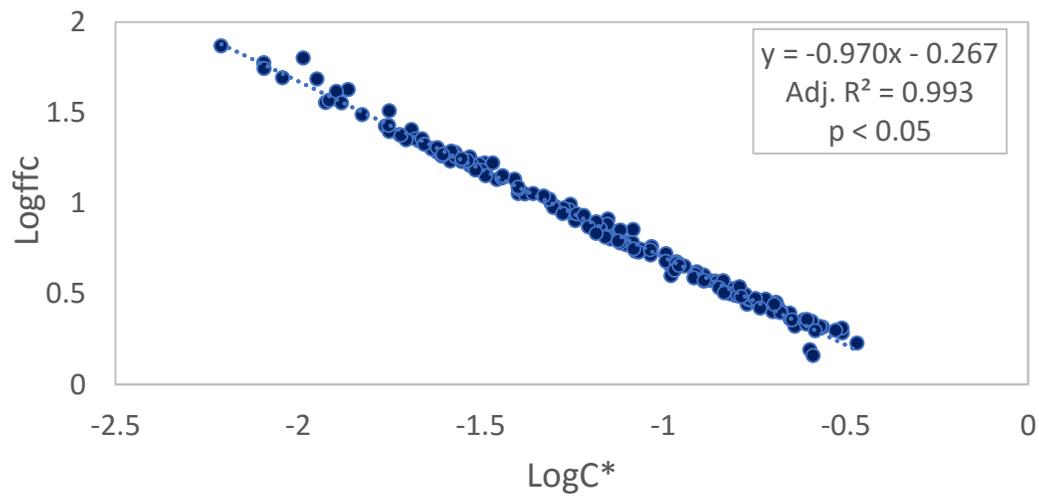
Figure 3-6 The correlation between the log of the flow function coefficient and log of C* for 41 materials under 4 initial consolidation stresses.
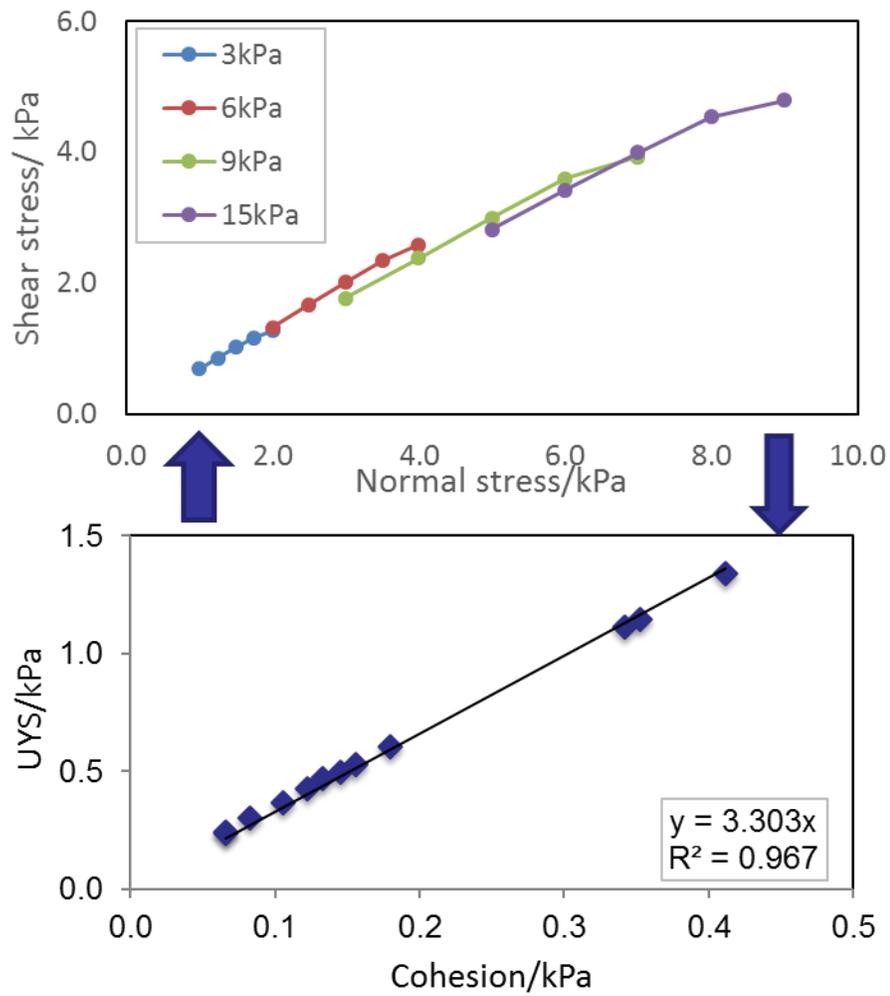
Figure 3-7 The yield points of maize starch are shown. Top: Yield loci of coarse alumina measured under four different initial consolidation stresses. Bottom: The unconfined yield strength (UYS) as a function of the cohesion.
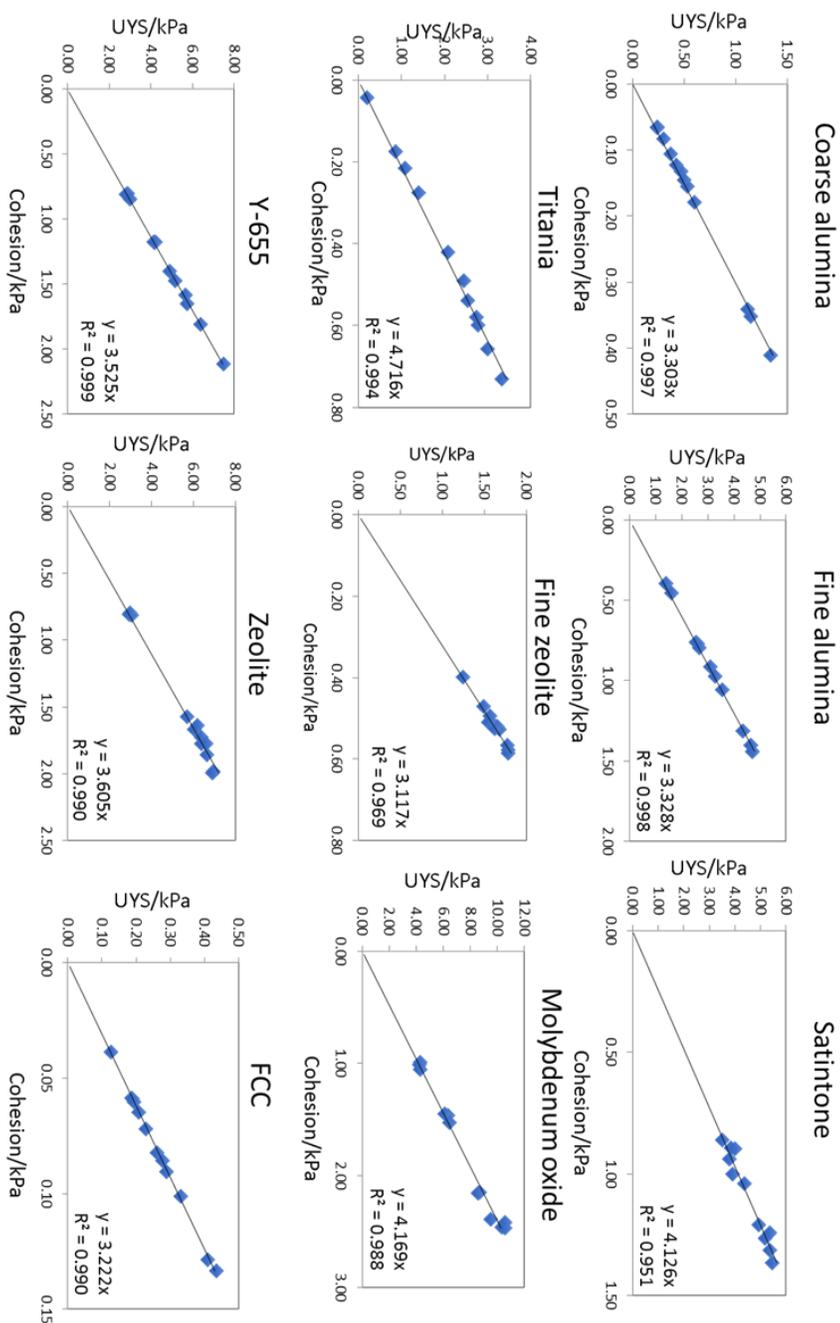
Figure 3-8 Plots of unconfined yield strength and cohesion showing the material characteristic lines for catalyst support materials.
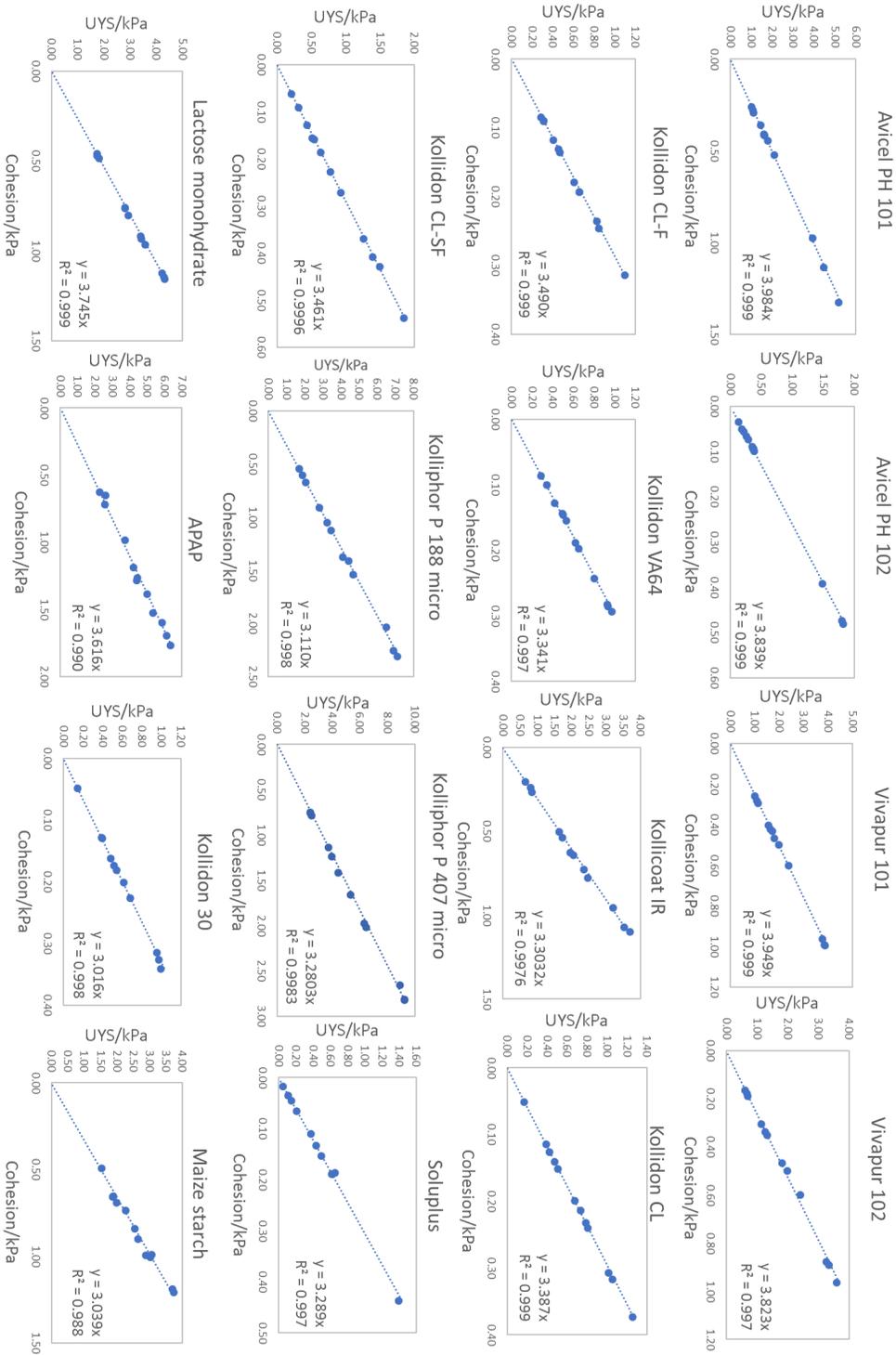
Figure 3-8 Plots of unconfined yield strength and cohesion showing the material characteristic lines for pharmaceutical materials.
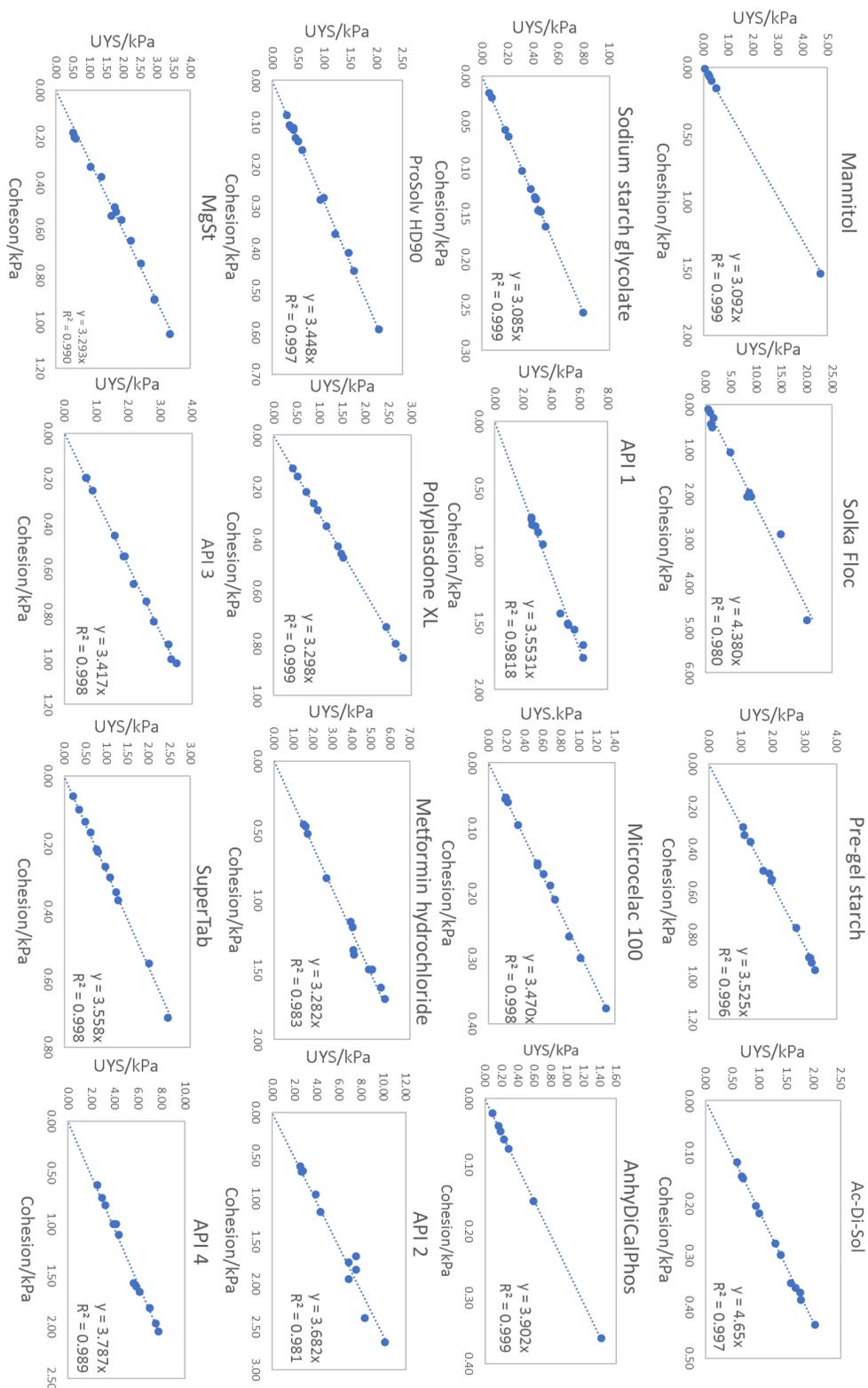
Figure 3-8-continued Plots of unconfined yield strength and cohesion showing the material characteristic lines for pharmaceutical materials.
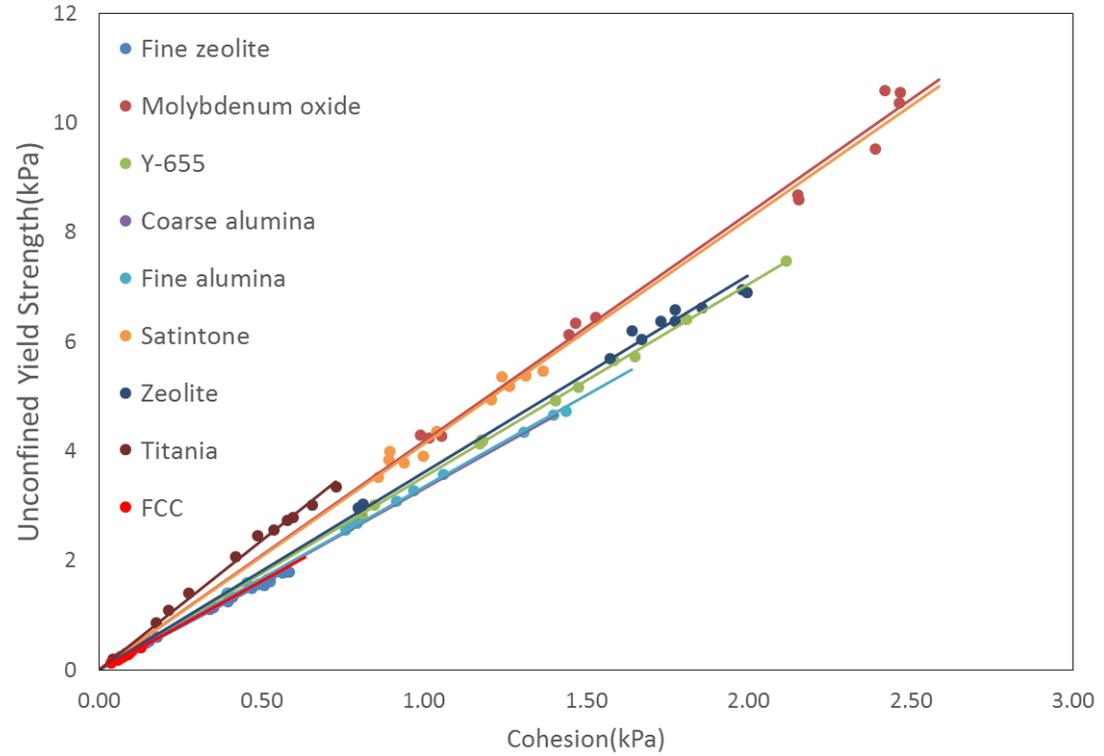
Figure 3-9 Using material characteristic lines to compare different materials. The unconfined yield strength (UYS) as a function of the cohesion.
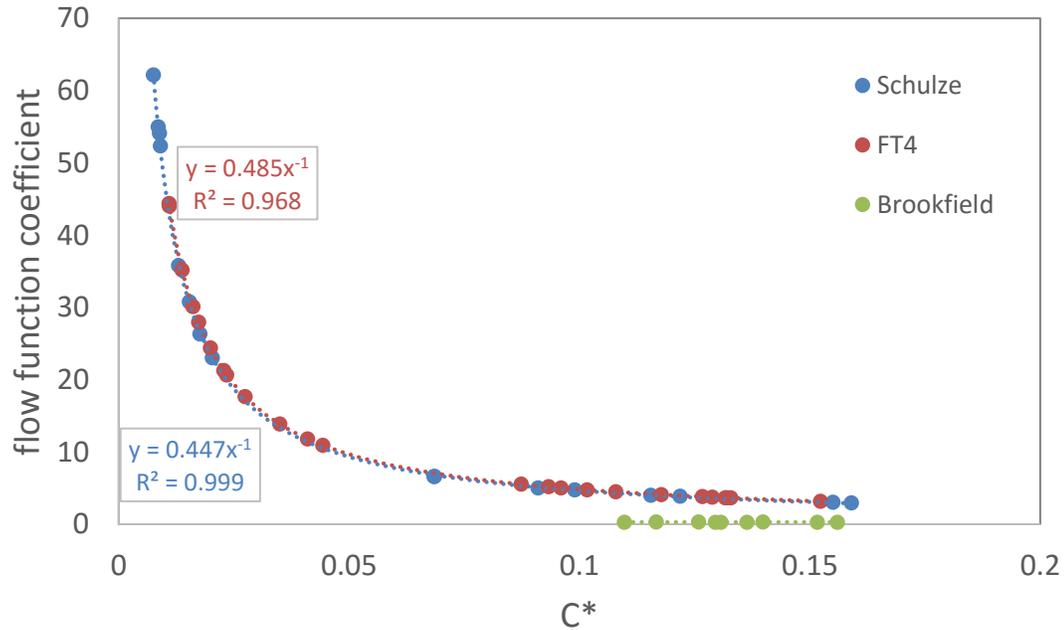
Figure 3-10 Comparing three shear cells by using ffc-C* correlation
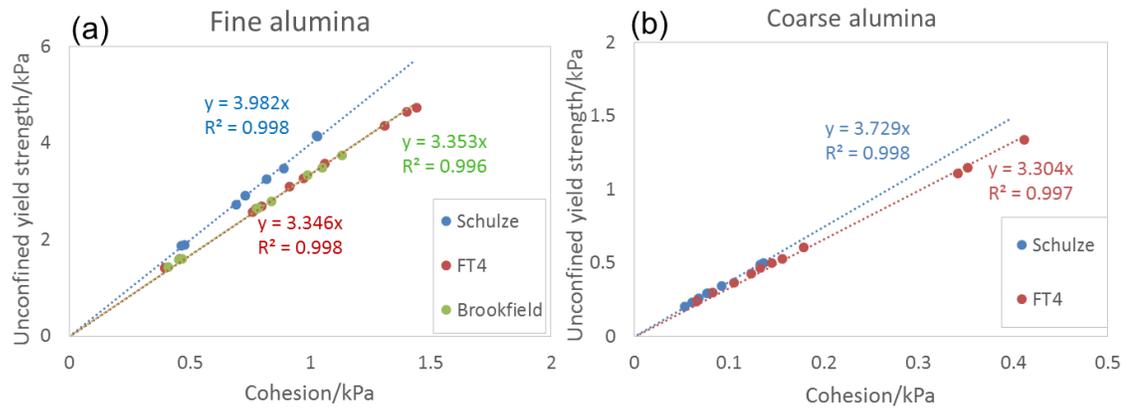
Figure 3-11 Comparing three shear cells by using the material characteristic line for fine alumina (a) and the coarse alumina (b).

3.6. Tables for Chapter 3

Table 3-1 Information for the 32 pharmaceutical materials in the study.

Material	Supplier
Kollidon® CL-F	BASF, Ludwigshafen, Germany
Kollidon® VA 64	BASF, Ludwigshafen, Germany
Kollidon® CL	BASF, Ludwigshafen, Germany
Kollidon® CL-SF	BASF, Ludwigshafen, Germany
Kollidon® 30	BASF, Ludwigshafen, Germany
Kollicoat® IR	BASF, Ludwigshafen, Germany
Kolliphor® P 188 micro	BASF, Ludwigshafen, Germany
Kolliphor® P 407 micro	BASF, Ludwigshafen, Germany
Soluplus®	BASF, Ludwigshafen, Germany
Avicel PH 101®	FMC Biopolymer, Newark Delaware, US
Avicel® PH 102	FMC Biopolymer, Newark Delaware, US
Vivapur 101	JRS Pharma LP, IA, US
Vivapur 102	JRS Pharma LP, IA, US
Lactose monohydrate	Foremost Farms, WI, US
Acetaminophen, anhydrous basis	Mallinckrodt Inc. IL, US
Maize starch	Mitsubishi Corporation, Tokoy, Japan
Mannitol, type SD 100	Roquette Pharma, Lestrem, France
Solka Floc	International Fiber Corportation, OH, US
Pre-gel starch	Ingredion Incorporated, IL, US
Ac-Di-Sol	FMC Biopolymer, Newark Delaware, US
Sodium starch glycolate, type A	Roquette Pharma, Lestrem, France
Microcelac® 100	Meggal, Wasserburg, Germany
AnhyDiCalPhos	JRS Pharma LP, IA, US
SuperTab® 11SD	DFE Pharma, Norten-Hardenberg, Germany
ProSolv® HD90	JRS Pharma LP, IA, US
Polyplasdone® XL	ISP Chemicals LLC. KY, US
Metformin hydrochloride	Ganules India Ltd, India
Magnesium stearate	Mallinckrodt, MO, US
API 1	Janssen Pharmaceutica N.V, Beerse, Belgium
API 2	Jassen Pharmaceutica, NV, Beerse, Belgium
API 3	Janssen Ortho LLC. Puerto Rico
API 4	Jassen Pharmaceutica N.V., Beerse, Belgium

Table 3-2 Information for the 9 catalyst support materials in the study.

Material	Supplier
Coarse γ-alumina	Albermarle, Amsterdam, the Netherlands
Fine γ-alumina	Albermarle, Amsterdam, the Netherlands
Fine zeolite	Albermarle, Amsterdam, the Netherlands
Molybdenum oxide (fine, sublimed type)	Albermarle, Amsterdam, the Netherlands
Zeolite Y-655 (calcined)	Albermarle, Amsterdam, the Netherlands
Zeolite Y-type CBV100	BASF Corporation, NJ, US
Titania ISK grade MC-150	BASF Corporation, NJ, US
Satintone (calcined kaolin SP-33)	BASF Corporation, NJ, US
FCC	BASF Corporation, NJ, US

Table 3-3 Particle size information and conditioned bulk density results for the materials used in the study.

Material	d10 (μm)	d50 (μm)	d90 (μm)	Conditioned Bulk Density (g/mL)
Kollidon® CL-F	26.77	37.31	86.89	0.29
Kollidon® VA 64	12.17	47.07	87.97	0.44
Kollidon® CL	17.90	76.06	154.90	0.42
Kollidon® CL-SF	11.26	30.14	68.94	0.21
Kollidon® 30	34.14	86.92	141.10	0.49
Kollicoat® IR	19.16	77.95	155.67	0.42
Kolliphor® P 188 micro	13.28	40.90	74.38	0.54
Kolliphor® P 407 micro	14.20	46.64	84.76	0.40
Soluplus®	197.10	325.40	476.10	0.81
Avicel PH 101®	22.93	69.98	154.89	0.34
Avicel® PH 102	34.63	113.39	220.95	0.36
Vivapur 101	20.14	64.09	131.15	0.35
Vivapur 102	31.46	123.75	224.95	0.37
Lactose monohydrate	10.32	63.54	157.78	0.66
Acetaminophen, anhydrous basis	9.93	57.40	166.00	0.41
Maize starch	9.95	14.40	24.20	0.51
Mannitol, type SD 100	59.60	97.90	147.00	0.46
Solka Floc	29.80	147.00	625.00	0.16
Pre-gel starch	8.89	36.60	104.00	0.51
Ac-Di-Sol	19.52	45.72	117.87	0.54
Sodium starch glycolate, type A	25.40	49.20	76.20	0.80
Microcelac® 100	54.10	158.00	264.00	0.51
AnhyDiCalPhos	68.60	185.00	304.00	0.75
SuperTab® 11SD	47.00	122.00	213.00	0.62
ProSolv® HD90	26.70	127.00	241.00	0.48
Polyplasdone® XL	9.65	23.60	54.16	0.42
Metformin hydrochloride	50.46	136.75	248.144	0.52
Magnesium stearate	2.1	7.8	8.8	0.27
API 1	9.66	63.60	185.00	0.36
API 2	3.15	11.90	24.40	0.37
API 3	14.60	47.10	95.90	0.36

API 4	9.75	41.50	153.00	0.36
Coarse γ-alumina	11.16	59.37	122.19	0.97
Fine γ-alumina	0.99	4.07	11.46	0.31
Fine zeolite	0.94	3.85	5.81	0.86
Molybdenum oxide	0.76	3.71	10.31	1.4
Zeolite Y-655 (calcined)	0.80	2.33	7.59	0.31
Zeolite Y-type CBV100	0.77	2.99	10.44	0.25
Titania ISK grade MC-150	0.69	2.99	10.13	0.35
Satintone (calcined kaolin SP-33)	0.92	6.09	21.14	0.52
FCC	38.48	79.65	134.30	0.83

4. Predictive Feeder Performance Based on Material Flow Properties

4.1. Introduction

Continuous manufacturing has been widely adopted across many industries, such as oil refining, natural gas processing and bulk chemical production [137, 138]. There are many advantages of continuous processing, including smaller scale equipment, higher flexibility, enhanced controllability, and reduced labor requirements. Studies have shown that continuous manufacturing can significantly decrease production costs while improving product quality [8]. In continuous processing, input raw materials and energy are fed into the system at a constant rate, and at the same time, a constant extraction of output products is achieved. The process performance is heavily dependent on stability of the material flowrate. For powder-based continuous processes, it is critical to feed powders consistently and accurately into subsequent unit operations of the process line, as feeding is typically the first unit operation. Feeder consistency also has an impact on mixing efficiency [139]. In particular, when multiple ingredients are fed to the process, inability to maintain constant material concentrations in the process stream can potentially lead to product failure [140]. Therefore, accurate gravimetric feeder performance has been recognized as crucial to success of the overall process.

Feeders have been designed to achieve performance reliability, feed rate accuracy, and minimal disturbances [141-143]. Loss-in-weight feeders control material dispensing by weight at a precise rate, and are often selected to minimize the flowrate variability that is caused by change of fill level and material bulk density [144]. Engisch *et al.* proposed a method to characterize and evaluate the steady state performance of a loss-in-weight

feeder for different powders [145]. Methods to measure feeding performance during hopper refill and to optimize scheduling of refilling have also been presented [146]. Kehlenbeck *et al.* showed improved feeder dosing consistency by using a single proportioning device at the discharge [147]. Tardos *et al.* used vibratory hopper agitation to improve overall flow and feeding precision [148].

In general, a particular feeder cannot handle all materials [149]. For a certain feeder, the feeding performance depends not only on equipment design and control systems and target flow rate, but also on material properties, specifically, material flow properties. For example, cohesive powders can adhere to screws, or can bridge across feeder hopper, causing phenomena called rat holes". Free flowing materials, on the other hand, often have higher densities and thus requires more energy for the feeder to displace the materials. They can also "flush" through the feeder, causing pulsating flow rates. A poorly paired powder-feeder combination typically leads to powder sticking to equipment walls, and high feed rate variability [150]. Therefore, the selection of an appropriate feeder and feeder tooling for a powder material usually starts with characterizing the powder flow properties. In fact, some flow property measurements were initially developed to bring a mathematical approach for powder handling equipment design, for example using shear cell tests for hopper specification [113]. However, currently, feeder tooling selection for a material to achieve good feeding performance heavily relies on empirical knowledge.

Raw material properties often have substantial impact on the performance of the manufacturing process [104, 109]. Identifying the relationship between raw material properties and process performance is critical for determining the effect of material

property variations on downstream processes [151]. However, there is still a lack of published work on using a systematic approach to predict feeder performance based on material flow properties. Owen *et al.* demonstrated use of the Discrete Element Method (DEM) to predict screw conveyor performance as a function of operating conditions [152]. Unfortunately, material properties are not considered in the model. Freeman *et al.* used a dynamic powder characterization test to correlate two dynamic properties to the volumetric flow rate of different screw feeders [153]. Importantly, the "flowability"(i.e., flow-related behavior) of a powder is a multi-dimensional characteristic [154]. Relying on a single characterizing technique to correlate to process performance is not sufficient [155]. Questions remain regarding which measurement or measurements provide the best predictive capability for feeder performance.

The purpose of this paper is the development of a multivariate approach to correlate feeder performance with material flow properties. Particularly, the aim is to answer the following three questions: i) How can we compare a given new material to previously characterized materials? ii) For a material with given properties, can we predict its feeder performance? iii) Furthermore, for a material with given properties, can we predict the optimal tooling that achieves a specified feeding performance?

The method to be presented in this paper includes characterization of material flow properties, characterization of feeder performance, and predictive multivariate analysis. As there is no unifying framework to describe powder flow behavior, materials were characterized by multiple flow property measurements so that each reflects a different aspect of flow behavior. Based on the methods to analyze shear cell data, introduced in Chapter 3, the angle of internal friction was included as one of material flow properties.

As discussed in Chapter 3, dimensionless cohesion, C*, was used to classify material flowability. Materials were thus selected based on measurement results from Chapter 3 to represent varying flow regimes, as shown in Figure 3-5. To characterize feeder performance, volumetric studies were initially performed to determine feeder capability, followed by gravimetric studies to evaluate overall performance. For multivariate analysis, two approaches were carried out. Principal component analysis (PCA) was firstly used to project material properties onto reduced dimensions; the predictive capability can then be achieved by calculating similarity scores to find materials with similar flow behavior. Secondly, partial least squares regression (PLSR) was used to directly predict feeder performance based on material property inputs. The method was demonstrated by using a commercially available feeder, a K-Tron KT20 loss-in-weight feeder, for seven powder materials that were characterized by four material flow characterization techniques, represented by 30 flow indices.

4.2. Materials and methods

4.2.1. Materials

The materials that were characterized in this study were: Coarse γ-alumina (Albermarle Inc., Amsterdam, the Netherlands), Fine γ-alumina (Albermarle Inc. , Amsterdam, the Netherlands), Fine zeolite (Albermarle Inc., Amsterdam, the Netherlands), Zeolite (Y-type CBV 100, BASF Corporation, NJ, USA), Satintone (calcined kaolin SP-33, BASF Corporation, NJ, USA), and Lactose monohydrate (Foremost Farms, WI, USA). The material used to demonstrate the predictive model was a calcined zeolite which we term "Material A" (Albermarle Inc., Amsterdam, the Netherlands). The particle size information of the materials was discussed in Chapter 3.

4.2.2.   Material flow property measurements

The shear cell test has been extensively discussed in Chapter 3, therefore is not introduced here. In addition to shear cell test, the following test was also performed including compressibility test, permeability test, shear cell test, and dynamic flow test.

4.2.2.1.   Compressibility test

The compressibility test is part of the Freeman Technology FT4 Powder Rheometer suite (Freeman Technology Inc., Worcestershire, UK). The test measured change of powder bulk density as normal consolidation stress changes. The powder was first conditioned by a helical blade in order to create a uniform and reproducible packing state. A normal force was then slowly applied by a vented piston. The normal forces ranges from 0.5 to 15 kPa, holding each load for 60 seconds. The change of volume due to compression was measured, and the compressibility (CPS %) was calculated as the percent change in volume after compression:

$$\text{CPS \%} = 100 * \frac{V_c - V_p}{V_c}$$

where V_c is the bulk volume after the conditioning step and V_p is the powder volume after compression.

4.2.2.2.   Permeability test

Permeability test, also achieved by the FT4 powder Rheometer, measures how well air passes through a powder bed. After conditioning step mentioned previously, an upward air velocity of 2mm/s was introduced from the bottom of the powder bed. Meanwhile, a

normal stress was applied by the vented piston over a range of 0.5 to 15kPa. At each normal force, the pressure drop (PD) across the bed was recorded.

4.2.2.3. *Dynamic flow test*

Flow energy E is defined as the energy required to move a helical blade through a powder bed. The Freeman Technology FT4 dynamic test measures this energy as a function of time and shear rate [18]. The powder sample was loaded into the vessel and then conditioned using a helical blade. The blade, moving on the downward and upward traverse, was used for the testing cycle and the energy consumed to induce the powder to flow was measured. The conditioning step followed by a testing cycle was repeated seven times with measured flow energy E_1 – E_7. The flow energy required on the seventh downward blade pass is known as the basic flow energy (BFE). The energy consumed during the seventh upward traverse is called the specific energy (SE). The stability index, SI, reflects the change in flow energy over time:

$$SI = \frac{E_7}{E_1}$$

where E_1 is the flow energy required on the first downward blade pass, and E_7 is the flow energy required on the seventh downward blade pass. A SI value of 1 suggests that the material is stable and non-friable. If SI is larger than 1, it indicates that a material requires more energy to flow over time, possibly due to de-aeration, agglomeration, moisture uptake, or electrostatic charges. If SI is smaller than 1, it may be caused by material attrition or de-agglomeration [18].

Four additional repetitions were performed with different shear rates (i.e., different blade tip speeds). The four repetitions were performed with a tip speed ranging from 10 to 100mm/s. The flow rate index (FRI) is the ratio between the required flow energy at 100mm/s and 10mm/s:

$$\text{FRI} = \frac{E_{11}}{E_8}$$

Where E_{11} is the basic flow energy at a blade tip speed of 100mm/s, and E_8 is the basic flow energy at 10mm/s. If materials are more sensitive to changes in flow rate, for example cohesive materials due to higher air content, the FRI values are typically higher.

In total, 30 different (but related) flow measurements were used as material property inputs for feeder performance prediction, including conditioned bulk density (CBD), compressibility% from the compressibility test, pressure drop at 15kPa, consolidation stress from the permeability test, cohesion, unconfined yield strength, major principal stress, angle of internal friction, flow function coefficient at 3kPa, 6kPa, 9kPa, and 15kPa from the shear cell tests, basic flow energy, stability index, flow rate index, specific energy from the dynamic flow test, and d10, d50, and d90 from the particle size distribution. Three replications were performed for each test that was performed.

4.2.3.   Feeder characterization

A commercially available twin screw loss-in-weight feeder K-Tron KT20 (Coperion K-Tron Pitman, Inc. Sewell, NJ, US) was used in the study. A loss-in-feeder consists of three parts: volumetric feeder, weighing platform, and a gravimetric controller, as shown in **Figure 4-1.** The weighing platform measures the mass of the feeder and contained

powder materials in the hopper. The bottom of the feeder hopper contains a horizontal agitator that helps to fill the flights of feed screws. The gearbox controlling the screws is a type C gearbox with maximum speed of 154 RPM.

As feeder delivers powder materials, the gravimetric controller acquires signals from the weighting platform over time, and determines real time feed rate. The feed rate is thus controlled by adjusting rotation speed of the screw that dispense the material accordingly. The feeder was placed on a plane lab bench. Feeder characterization experiments were performed by using a gain-in-weight catch scale (OHAUS adventurer, OHAUS Corporation, Parsippany, NJ, US) to record the weight of powder dispensed by the feeder every 1s for all tests. Buckets were used to collect the samples. The feeder was connected to an external laptop. Besides data from the catch scale, feeder data from the gravimetric control box was also recorded to ensure consistent control capability, including time, set point, mass flow, screw speed, drive command, and perturbation value. Three feeder screws were available in this study: fine concave screw (FCS), coarse concave screw (CCS), and fine auger screw (FAS)**.** The experimental set-up and screws are shown in **Figure 4-2.** The feeder was firstly calibrated with 100% fill level at volumetric mode, without engaging the feeder gravimetric control system, to determine the maximal volumetric capacity. The calibration procedure was performed three times and the average was recorded as the initial average feed factor. The averaged initial feeder factor is the control value that refers to the capacity of the feeder at 100% of the control magnitude.[145]

After calibration at volumetric mode, the feeder was then performed in gravimetric mode starting at 80% fill level. Fine concave screw was firstly used because it has the smallest

capacity among the three screws. For each material, the set point for the three screws tested was kept constant and set at 50% and 80% of the initial average feed factor of fine concave screw. For each screw, the experiment was run for 20 min. The data obtained from the catch scale and the control box were used for screw comparison. Defining the target mass feed rate as m, the actual mass feed rate, $\dot{m}_\iota$, at the ith time point can be calculated by:

$$\dot{m}_\iota = \frac{\Delta m_i}{\Delta t}$$

Since the set point for each material may not be the same, relative standard deviation (RSD) and relative deviation between the set point and the mean (RDM) were used as criteria for feeder performance:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(\dot{m}_i - \bar{\dot{m}})^2}{n-1}}$$

$$RSD = \frac{\sigma}{\bar{\dot{m}}}$$

$$RDM = \frac{|\dot{m}_i - \bar{\dot{m}}|}{\bar{\dot{m}}}$$

where σ is standard deviation, $\bar{\dot{m}}$ is the arithmetic mean of the actual mass feed rate, and n is the number of time points. A moving average of 5 seconds of the obtained data from catch scale was used to calculate the RSD and RDM (Δt=5). RSD reflects the consistency of the mass feed rate, and RDM suggests the difference between the actual mass feed rate and the target mass feed rate. In this study, the aim is to find the screw that provides the best feeder performance, namely performance with lowest RSD and RDM. For data

analysis, experiment data of the first 5 seconds and the last 10 seconds were removed since there may be significant perturbation. For the same reason, data of a period of 20 seconds were also removed if changing bucket was needed during the experiment [145].

4.2.4. Multivariate analysis

Multivariate data analysis was performed using The unscramble X 10.2 (Camo, Oslo, Norway). Since material properties have different units and theirs ranges have different scales, the data was firstly preprocessed using Z-score normalization. The z scores were calculated as follows:

$$z = \frac{x - \mu}{\sigma}$$

where μ is the average of each property and σ is the standard deviation from the average. Data after z-score normalization are centered around 0 with a standard deviation of 1, which is a common requirement for most machine learning estimators, especially for clustering analysis, such as PCA, when comparing similarities between samples is needed based on certain distance measures.[156, 157]

Two approaches to predict feeder performance are discussed in this study: Principal component analysis followed by similarity scoring (PCA-SS), and Partial least squares regression (PLSR). Both approaches identify hidden structures in the material property data set and use a vector space transform to reduce the dimensionality of large data sets [158]. The difference between the two approaches is that PLSR directly regresses the data in reduced dimensions to feeder performance, while PCA-SS firstly identifies similarity between materials and uses feeder performance of the most similar material to predict results.

In the PCA-SS approach, a PCA model was used to represent the z-score normalized material property data set (X) in a reduced dimension (principal component space) such that the major axes of variability are identified [57]. The data set X can be decomposed, based on the equation below, into a set of scores (T) and loadings (P), while the remaining variability is modeled as random error (ε).

$$X = TP^T + \varepsilon$$

The columns T represent principal component (PC) scores of each material in the projected space; loadings P represent the significance of each material property in each principal component. Both T and P are obtained from eigenvectors and eigenvalues of the covariance matrix of X. A good indication of similarity between two materials is a measure of the distance between them [159]. Principal components (PCs) [61] are orthogonal to each other, and each is associated with the value that explains the proportion of variability in the data set. Usually, only a few PCs are retained based on their statistical significance. The similarity scores between material a and material b can thus be calculated based on weighted Euclidean distance d_w:

$$d_{w\,(a,b)} = \sqrt{\sum_{i=1}^{n} w_i\,(a_i - b_i)^2}$$

where n is the total number of principal components selected in the model, a_i is the score of material a in the ith principal component, b_i is the score of material b in the ith principal component, w_i is the weight of the ith principal component, namely the relative variability explained by the ith principal component:

$$0 < w_i < 1$$

A relatively small value of d_w suggests similarity between materials.

Partial least squares regression, also known as projection to latent structure regression, was used also to correlate a number of predictor variables X with response variables Y by finding the latent variables in the data set. The calculation of latent variables in the PLSR approach takes into account the response variables so that the linear combinations have maximal covariance [61]. In this study, the aim is to understand how varying flow indices impact feeder performance for different materials. Therefore, flow properties of six powder materials, covering a wide range of flow properties, were used to develop the calibration model for RSD and RDM obtained from feeder characterization experiments. Cross-validation (CV) was used for initial method evaluation. CV was performed by developing six parallel regression models from reduced data with one of the materials deleted, which is also called leave-one-out cross validation. In addition, the model was used to predict feeder RSD and RDM of one additional material, material A. A confirmatory feeder experiment for material A was then performed using the same procedure described in section 2.3. The predicted results were compared with the experimental results.

4.3. Results and discussion

To demonstrate the application of proposed approaches, Material A was used as the prediction material. This material belongs to a group of calcined zeolite materials for catalyst manufacturing processes. Flow properties of Material A was firstly characterized, and each flow parameter was compared with all other materials that were already characterized. Because both approaches are, in general, linear regressions, it is important to ensure that each flow parameter of Material A is within the range of existing

materials.[160] **Figure 4-3** compares the results from shear cell test at initial consolidation stress of 9kPa. The major principal stress relates to a material property during steady-steady flow under a certain normal stress. Unconfined yield strength, cohesion, and angle of internal friction are known as characteristic values when material achieves incipient flow.[161] The flow function coefficient (ffc) is often used to rank the flowability of a material, where a larger ffc value indicates a more free-flowing material. The materials can be ranked in terms of ffc from highest to lowest as follows: coarse alumina, zeolite, lactose, fine alumina, satintone, Mateiral A, and fine zeolite. Shear cell test results at initial consolidation stress of 3kPa, 6kPa, and 15kPa were not shown in **Figure 4-3,** but it was confirmed that, for each flow index, data for Material A was between the minimum value and maximum value of six calibration materials.

Figure 4-4 shows results from compressibility test and permeability test for all materials. Typically, a lower compressibility percent corresponds to a material with lower cohesion and better flowability. Based on compressibility result, the material flowability can be ranked from highest to lowest as follows: Coarse alumina, zeolite, lactose, Material A, fine alumina, fine zeolite, and satintone. Interestingly, material flowability ranking in terms of compressibility percent was different from the one based on ffc. In other words, material with higher ffc does not necessarily have higher value of compressibility percent. Similarly, pressure drop, obtained from permeability test, did not have a linear correlation with compressibility percent. For example, while both fine alumina and fine zeolite had relatively higher value of compressibility percent, fine alumina had much larger pressure drop than fine zeolite had, possibly due to difference of pore volume inside the particles.

Figure 4-5 includes results from dynamic test. As mentioned previously, the basic flow energy (BFE) measures the energy requires to stir a blade through a sample. Generally, if a material is free flowing, it requires low energy to move the powder, resulting in a low BFE. For porous materials, BFE will also be low, as the blade is in contact with low density particle bed. Interestingly, among all materials characterized, while coarse alumina was the least compressible material, and satintone as the most compressible material, both had relatively higher value of BFE compared to other materials. Both materials had relatively high value of conditioned bulk density, suggesting that bulk density has potentially an importantly role in determining the basic flow energy. All materials in this study had stability index (SI) close to 1.0, indicating materials are relatively stable and non-friable. Flow rate index (FRI) measures material sensitivity to flow rate. Most materials in this study had FRI value between 1.5 to 3.0, which corresponds to a modest increase in flow energy at higher flow rate. Lactose, with a FRI value of 1.0, was considered as a flow rate insensitive material.

In summary, two conclusions can be made based on material characterization results: i) For all flow indices considered in this study, Material A, was within the range between the maximal value and minimal value of other six materials. Material A can thus be used for prediction based on material properties. ii) Comparing materials based on a single flow index, also called as a univariate approach, is not sufficient. As described earlier, each flow parameter reflects a different aspect in terms of flow property. Furthermore, results from conducting multiple univariate comparisons may often be perplexing to interpret. In fact, univariate comparison approach does not take into account the correlation between different parameters. Results have shown that flow parameters do not

all vary independently, and how they correlate with each other is also a critical material characteristic. Therefore, multivariate analysis was performed to compare materials with different flow properties.

In multivariate analysis, not only the isolated individual flow index was included, but also the relationships among different indices. All material property data can be considered and analyzed simultaneously, providing greater statistical power compared to univariate comparison.[162] A PCA model was developed including Material A and six calibration materials. As the aim in this study is to identify materials with similar properties, the number of principal components was retained to ensure the model captured at least 95% of the overall variability in the data set. In addition, cross-validation was used to avoid model overfitting when most of the variability was included in the model. The first four principal components were found to explain 96.4% of total variability in the dataset, as shown in **Figure 4-6.** This was also the component with the highest explained variability in validation, and for this reason, four components were retained in the model to be used in subsequent calculations. A score plot was used to visualize how each material is projected into the reduced dimension space. As shown in **Figure 4-7**, a cubic score plot, including the first three principal components that explained 91.4% variability of the data set, suggests that the six calibration materials represented a wide range of materials with varying flow properties. If materials have the same properties, they will be projected to the same location in the score plot. Materials that are projected to be close to each other indicate similar flow behavior. To quantify distance between two materials, a weighted Euclidean distance was calculated based on the first four principal components, as described in section 2.4. The similarity score

outputs are presented in **Figure 4-8**. The material most similar to Material A was fine zeolite. The identified similarity between Material A and fine zeolite using PCA-SS was further confirmed in the following feeder characterization experiment.

The relative standard deviation (RSD) and relative deviation from the mean (RDM) using different screws were used for comparison of feeder performance for each calibration material, as shown in **Figure 4-9**, and **Figure 4-10**, respectively. Consistent with previous knowledge, material with different flow properties lead to different feeding performance, and therefore different screw is selected in order to achieve the consistent and accurate delivery of materials.[163] For example, for coarse alumina, which is very free-flowing and can easily fill the flights of the screws, all screws were found to be able to achieve stable feeding at both predefined feed rate. The feeding performance of some less free-flowing materials, such as lactose monohydrate, zeolite, was found to be heavily dependent on selection of screws and feed rate. For lactose, the fine concave screw (FCS) achieved best performance at feed rate of 7.0 kg/hr, and coarse concave screw (CCS) became the optimal when feed rate increased to 11.2kg/hr . For zeolite, FCS had the lowest RSD and RDM at both feed rate of 3.0 kg/hr and 4.9 kg/hr, however, use of fine auger screw (FAS) at feed rate of 4.9 kg/hr significantly increased RSD and RDM. For cohesive materials studied here, the selection of optimal screw was consistent at two feed rate investigated in this study. Specifically, FAS was considered as the optimal selection for satintone at feed rate of 2.1 kg/hr and 3.4 kg/hr, and fine alumina at 4.3 kg/hr and 6.8kg/hr. FCS was selected for fine zeolite both at 4.7kg/hr and at 7.5kg/hr.

As fine zeolite had the shortest weighted Euclidean distance, FCS was predicted to be the optimal screw for Material A. To validate the hypothesized prediction, feeder experiment

was performed using Material A, and the results were compared to fine zeolite. As shown in **Figure 4-11**, FCS was considered the optimal screw for both materials. However, it was noticed in **Figure** 4-**11(a)** that FAS had a higher RSD than CCS for fine zeolite, while for Material A, the order reversed. Nevertheless, the overall performance of Material A and fine zeolite was relatively agreed.

Surprisingly, a strong linear correlation was observed between the initial feed factor and score of the first principal component for all three screws, as shown in **Figure 4-12**. Since the first principal component is a linear combination of measured flow indices, the initial feed factor can be directly predicted based on observed linear correlation. The initial feed factor is an important process parameter that refers to maximal capacity for each screw. A typical operation set point during feeding process for each screw is usually 20-80% of its maximal capacity. Therefore, the initial feed factor is required to determine operational space. To obtain the initial feed factor, hopper needs to achieve 100% fill level, which requires a comparable large amount of materials, relative to the amount required for material property measurements. In the case presented here, 2-6 kg of materials were used, depending on material bulk density, to obtain the initial feed factor, while the amount of materials required for characterization was 0.2-0.5kg. It is common that in early process and product development, the amount of materials is often limited. In this case, the predictive correlation can be used to determine the initial feed factor based on material properties, represented by scores of the first principal component.

The workflow of PCA-SS approach is described in **Figure 4-13.** In general, when a new material is received, it is first characterized by flow property measurements described in section 2.2. A PCA is performed including the new material and all existing materials

that have been previously characterized. By knowing the PC score of each material and weight of each PC, the weighted Euclidean distance can be calculated. Distances between the new material and all other materials can be ranked, and the material with the smallest distance can be identified. Based on the assumption that, under the same feeder parameter settings, materials with similar flow properties have similar feeder performance, the optimal feeder screw option for the new material can be predicted based on existing knowledge of the material with the smallest weighted Euclidean distance. More importantly, new material property is automatically added to material library. Knowledge of materials with various flow properties are accumulated. As the number of characterized materials is expanding, the accuracy of identifying similar materials will also be improved.

The advantage of PCA-SS approach is the ability to quickly identify material similarity considering all available flow property measurements. The prediction material is also included when developing PCA model, which enables the model to fully explore the overall data structure and pattern. The prediction error can be reduced significantly by using similarity scores, instead of projecting prediction material data to existing PCA models. When material is sparing, this approach can be powerful to find placebo materials, and use placebo materials to identify operation space. However, there also limitations for PCA-SS approach. The approach is based on an assumption that materials with similar flow properties have similar screw selection in order to achieve good feeding performance. If calibration materials do not cover a wide range of flow properties, the difficulty to find materials of similarity may potentially increase prediction error. In addition, the optimal screw was predicted only based on the material that has the highest

similarity. To consider the potential quantitative correlation between material properties and feeder performance, PLSR models can be developed.

In this study, four factor non-hierarchal PLS models were fitted to material property data consisting six calibration materials and 30 flow parameters. The response variables were RSD and RDM for each screw at feed rates that correspond to 80% of initial feed factor of FCS. The cumulative percent variance explained by the four factor in all models were above 95%. **Figure 4-14** suggest that good parity was observed between the predicted and actual y-values (RSD or RDM). The root mean square error of calibration (RMSEC) and root mean square error of cross validation (RMSECV) was also shown. The regression coefficients for RSD models are depicted in **Figure 4-15**. The regression coefficients indicate the contribution of each material flow property to the prediction models, with higher absolute value of coefficients indicating larger contributions. Positive value of coefficients indicates a positive correlation with the RSD, and negative coefficient values indicate inverse correlation. Material flow property with relatively small coefficients suggests its small contribution in the prediction. **Figure 4-15** suggests that material property contributes to feeder performance differently when different screws are selected. For example, pressure drop (PD) from permeability test had significant effect on RSD when using FCS and FAS, while its contribution when using CCS was relatively small. Also interestingly, PD increased RSD when using FCS, but decreased RSD when using FAS. Importantly, flow property variables that have small contributions can be removed without losing predictive capability. That work will be the focus of a subsequent publication.

The RSD and RDM of a material during feeding process can be predicted based on obtained regression coefficients and material flow properties. **Figure 4-16** shows the prediction results of Material A in comparison to measured results. It was validated that the developed PLSR models were able to predict RSD and RDM that are reasonably similar to experimental results. By comparing predicted RSD and RDM values using different screws, FCS was predicted to be the optimal screw for Material A, which was consistent with predicted results using PCA-SS approach, as well as experimental results. The workflow using PLSR approach can be summarized in **Figure 4-17**. When a new material is received, its properties were firstly characterized. The material property data is then projected to the PLSR model. The model outputs are RSD and RDM using different screws. The optimal screw, with the lowest predicted RSD and RDM value, is then selected. Confirmatory experiments using the screw from prediction results are performed. Importantly, models are consistently maintained and updated by adding confirmatory experimental results to the model.

While multivariate analysis has been used to develop the predictive process performance based on material properties, it is important to note that the size of the data set available is important to both the approaches discussed in this study. For the PCA-SS approach, as more materials are added, model knowledge to predict feeder performance is consistently expanded and accumulated, increasing the likelihood of finding similar materials for a given new material. For the PLSR approach, models are maintained and updated when feeder performance of more materials are obtained during process development, continuously improving the prediction accuracy for future materials. Further work is required to test if the proposed approaches can be validated across a wider range of

material flow properties that were not considered in this study, for example electrostatics [164, 165], surface energy, etc. Additionally, as more flow property measurements are considered, it would be useful to use the multivariate analysis and similarity scores as tools to meaningfully select the measurements that have higher weights in terms of classification of the materials.

4.4. Conclusions

This paper examines two methodologies, PCA and PLS, to develop predictive correlations between material flow properties and material feeder performance. The methods include characterization of material flow properties, characterization of feeder performance, and predictive multivariate analysis. Six calibration materials with varying flow properties were firstly characterized by five techniques. The flow properties of each material were represented by 30 flow indices. The calibration materials were characterized by feeder performance. Multivariate analysis was used to find the correlation between material properties and feeding performance. Specifically, two approaches were discussed and compared. The PCA-SS approach is based on weighted Euclidean distance, calculated after performing the principal component analysis on the entire data set. The material that has the shortest distance to a given new material can be considered to have similar flow behavior. The PLSR approach further quantifies the correlation between material properties, projected into a reduced dimensional space, and feeder performance, represented by the RSD and RDM.

This study shows that material flow properties affect feeder performance, and therefore the screw that achieves the best feeding performance is heavily dependent on the material flow properties. Experimental results in combination with multivariate analysis suggest

that material flow properties and feeding performance are highly correlated. Both approaches discussed in this study were confirmed by cross-validation and prediction of one additional material. The predicted feeding results were, in general, in good agreement with the experimental results. The work presented here has shown an efficient approach to correlate material properties with process performance using multivariate analysis. This approach is especially powerful when the amount of a given new material is limited or if the new material is expensive or dangerous.

The initial feed factor and scores of the first principal component for each material were found to be linearly correlated. By characterizing flow properties, the initial feed factor, or maximal capacity, of each screw can be predicted, which enables a quick identification of operational ranges. By using the PCA-SS method, a placebo material with similar flow behavior can be identified. When a material is not available in large amounts, placebo materials can be used for process development. Additionally, instead of selecting optimal tooling based on empirical knowledge and a trial-and-error method, feeding performance using different screws can be directly quantified, predicted and compared by developing predictive PLSR models.

Importantly, for any data-driven models, the size of the data set is important for prediction. Therefore, the methodology should be implemented as a self-updating knowledge reservoir. The more materials and properties that the data base includes, the better will be the predictive power of the models. It is also important to notice that the same type of models can be augmented by additional measurements performed during process scale-up, validation, and commercial manufacturing in order to capture process knowledge and increase predictability. At the same time, this study has been for a limited

number of materials and for one feeder. Future work should examine additional materials and other feeders in order to examine if the results observed in this work carry over to other types of feeders. Future work should also test the validity of the proposed methods across a wider range of material properties, characterized by additional techniques.

In Chapter 3, the intrinsic power correlation between the cohesion and the flow function coefficient from the shear cell test has been discussed. As this chapter investigates the correlation between material flow properties and feeding process, it is also important to understand the correlation between different flow indices. Therefore, results of powder mixtures are discussed in Chapter 5 to validate the correlation that was found in raw materials. Chapter 6 examines the correlation between different flow indices of powder mixtures, and uses the correlation to compare different mixing systems. As more characterization techniques become available, PCA-SS can also be used to reduce the number of measurements, which will be discussed in future work recommendation section.

4.5. Figures for Chapter 4
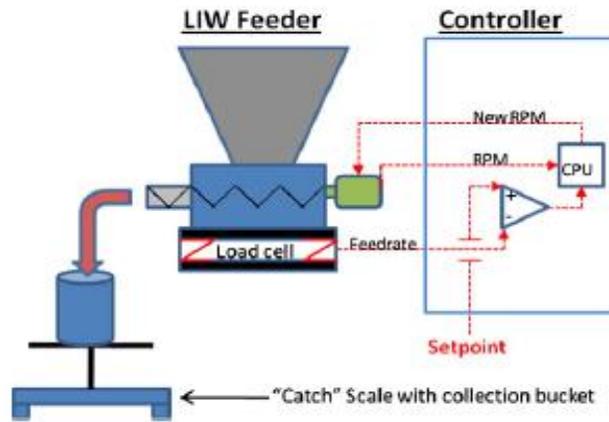
Figure 4-1 A schematic illustration of a loss-in-weight feeder. (picture source: W. Engisch, F. Muzzio)
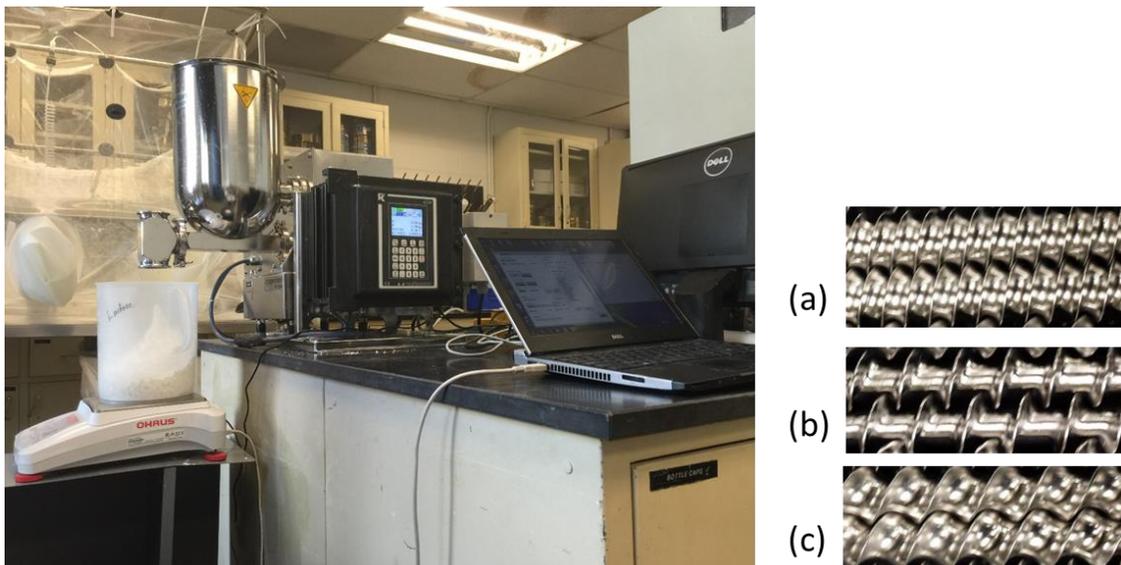
Figure 4-2 Experimental set-up of characterizing feeding performance for each material. Available screws in this study are (a) fine concave screw, (b) fine auger screw, and (c) coarse concave screw.
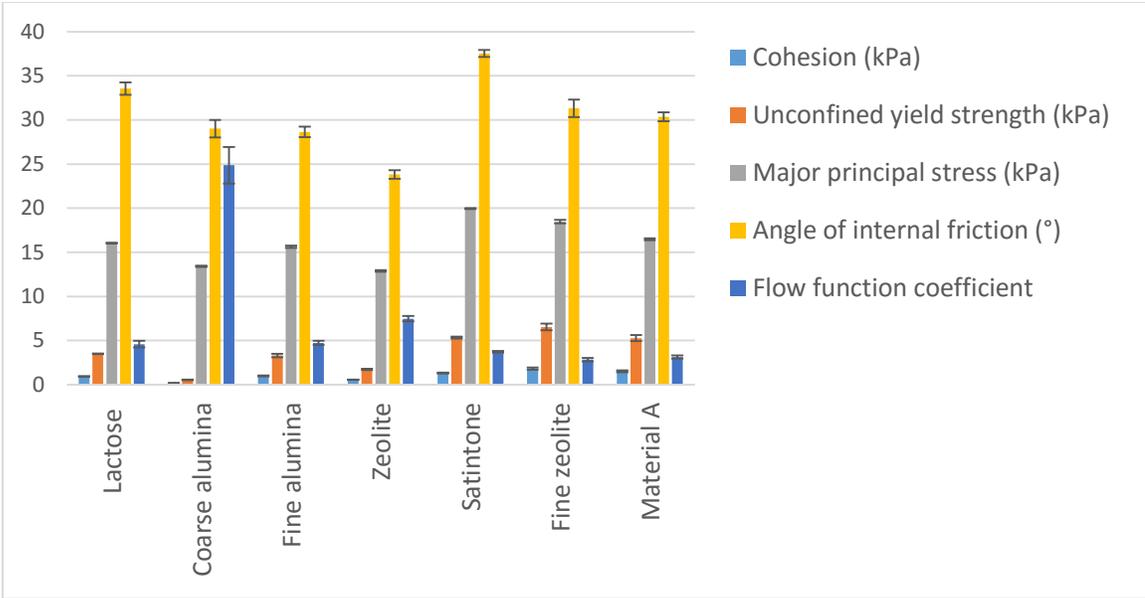
Figure 4-3 Material characterization results from shear cell test at initial consolidation stress of 9 kPa.
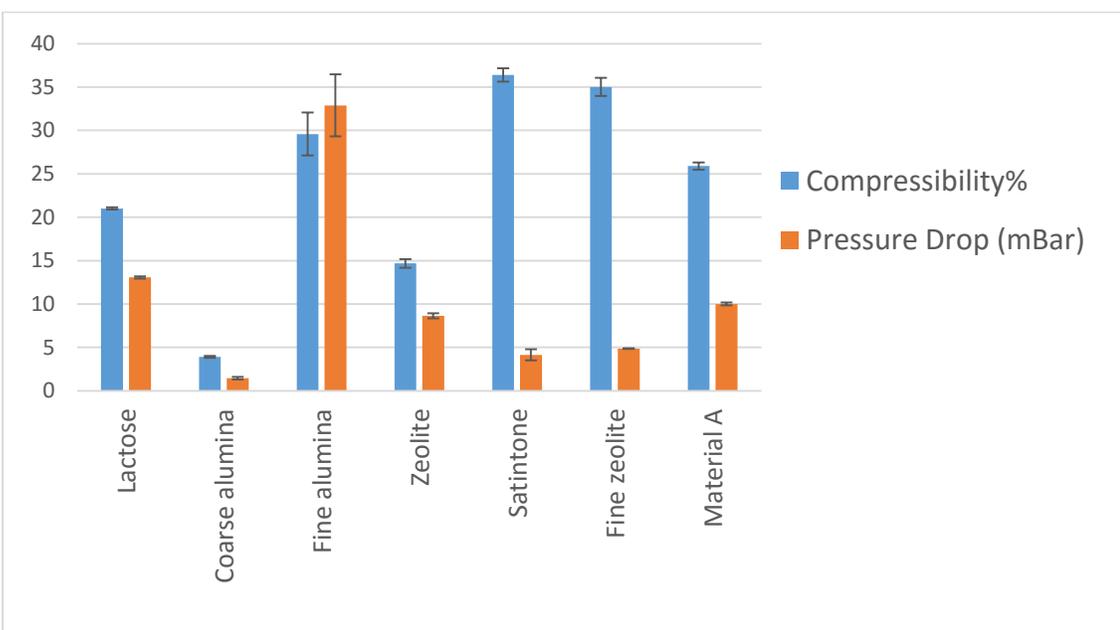
Figure 4-4 Material characterization results from compressibility test and permeability test.
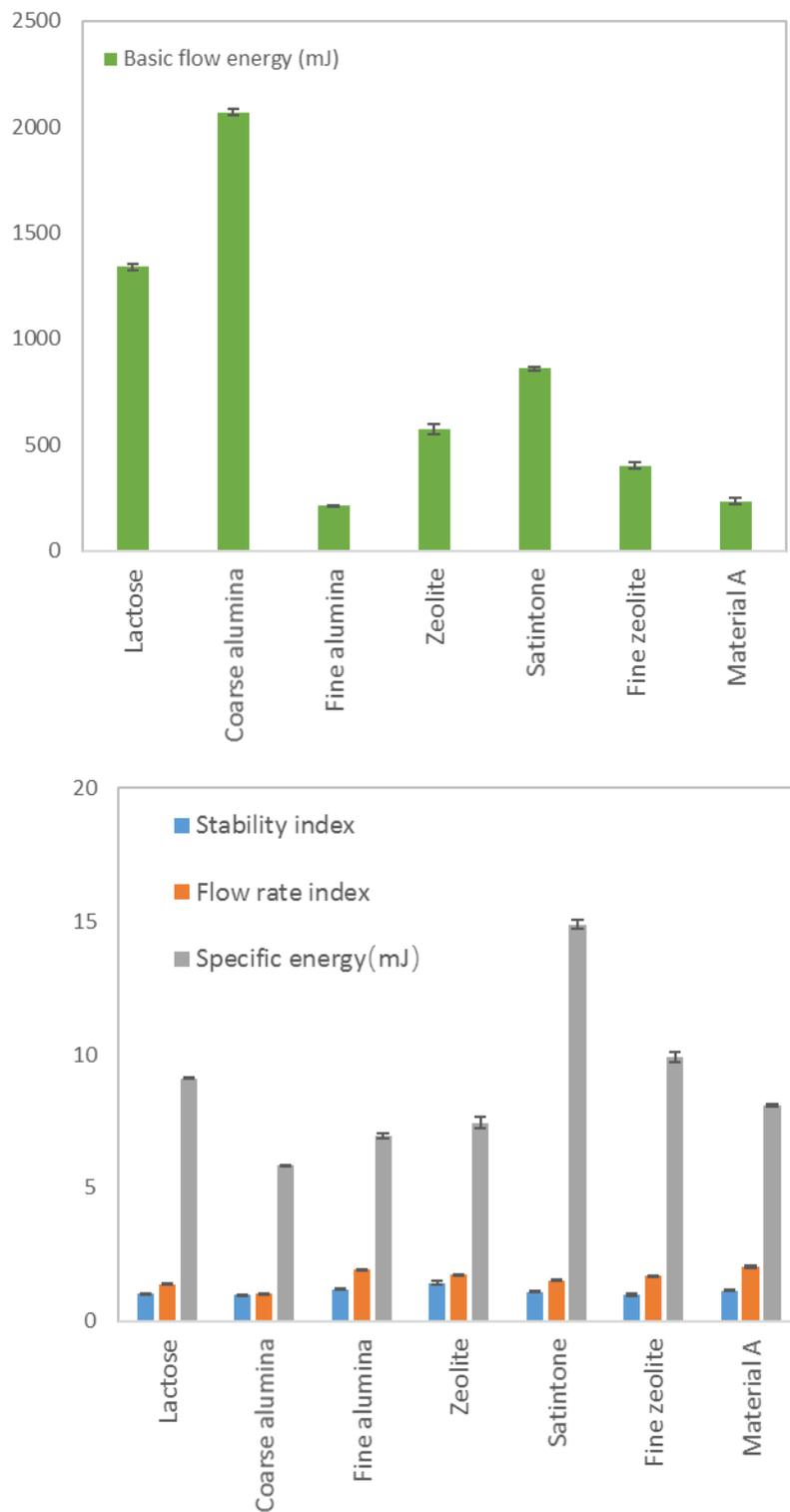
Figure 4-5 Material characterization results from dynamic flow test.
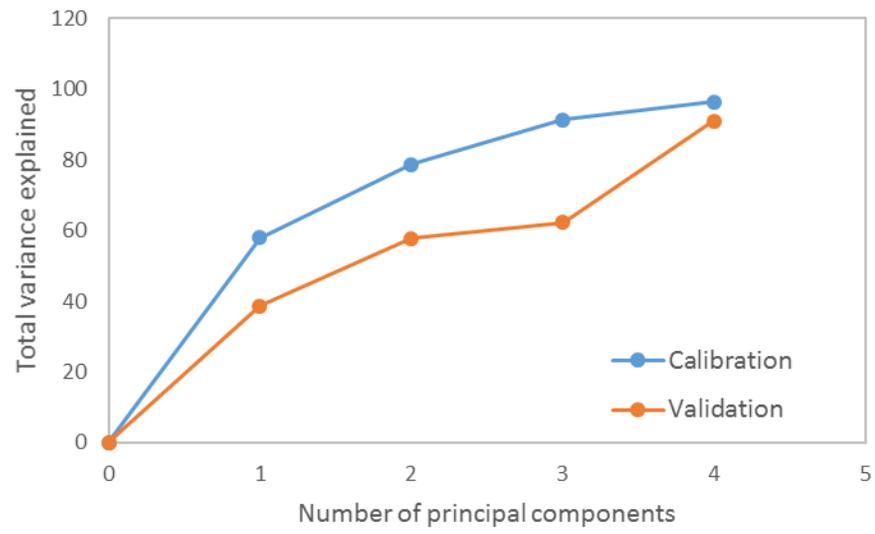
Figure 4-6 The first four principal components were selected for the PCA model.
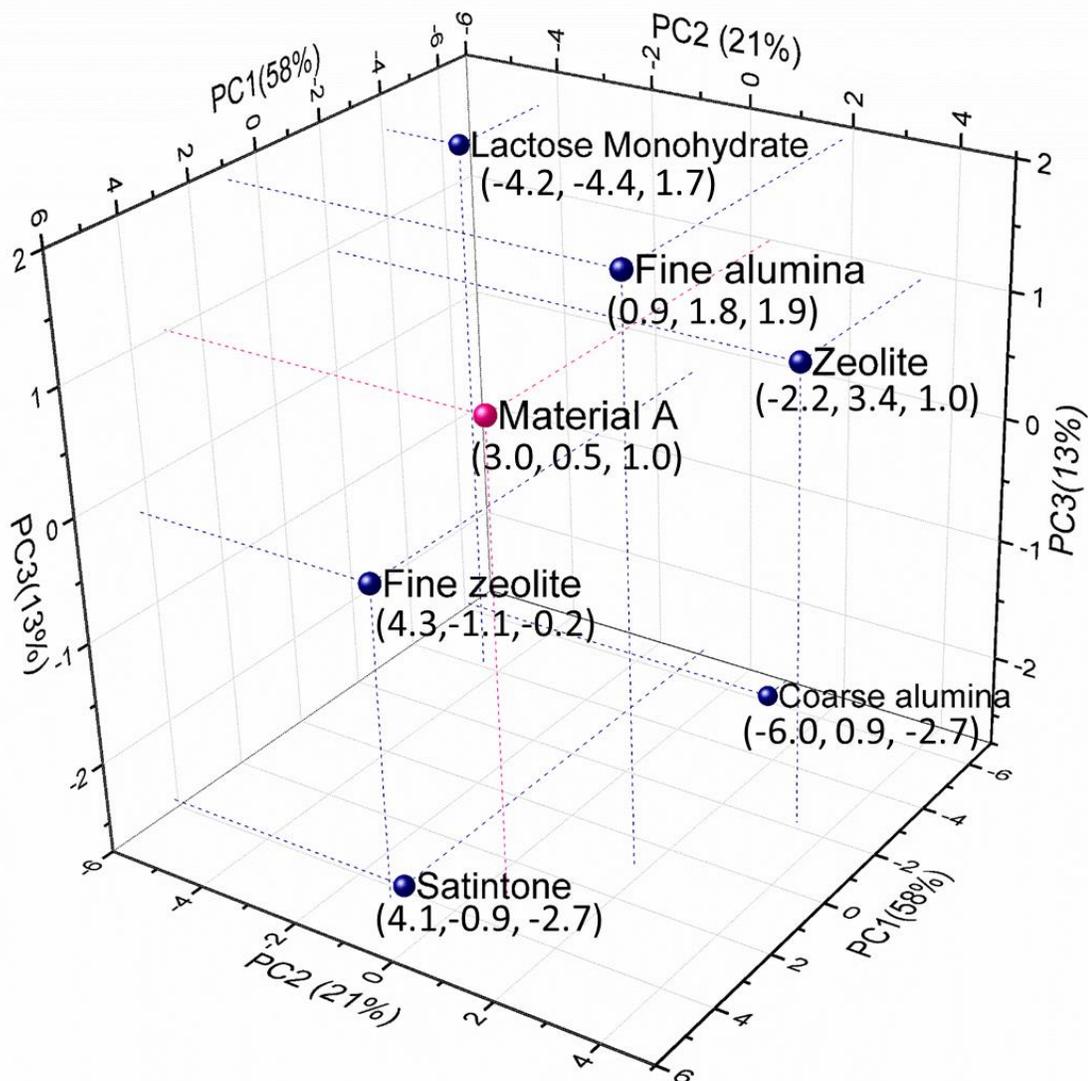
Figure 4-7 A cubic score plot was used to visualize materials in the projected spaces. Only the first three principal components were shown in cubic plot.
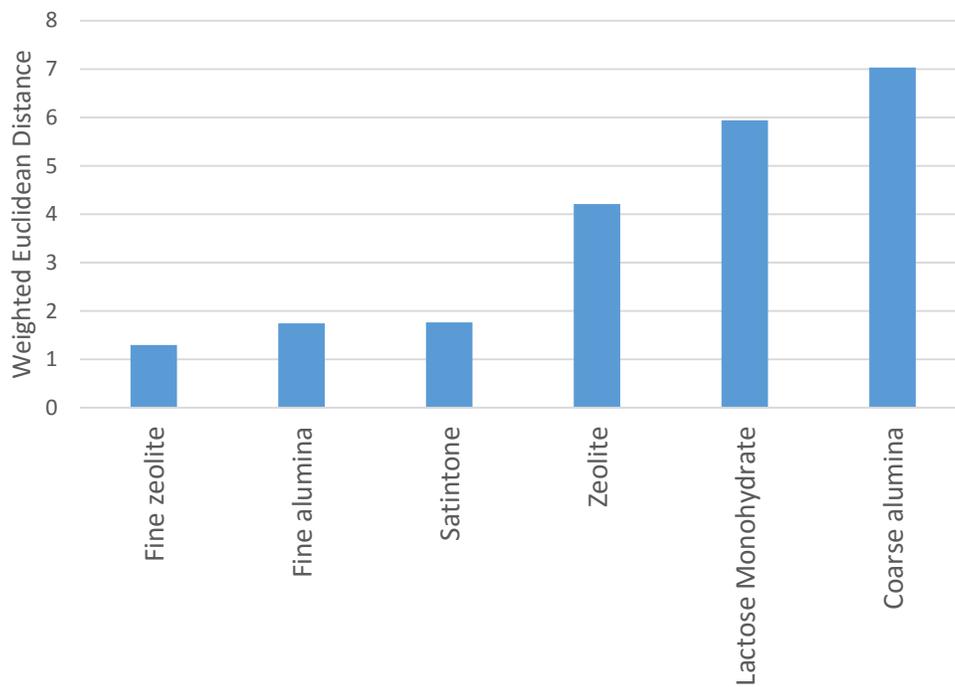
Figure 4-8 Based on similarity score using weighted Euclidean distance, fine zeolite was identified to be the most similar material to Material A.
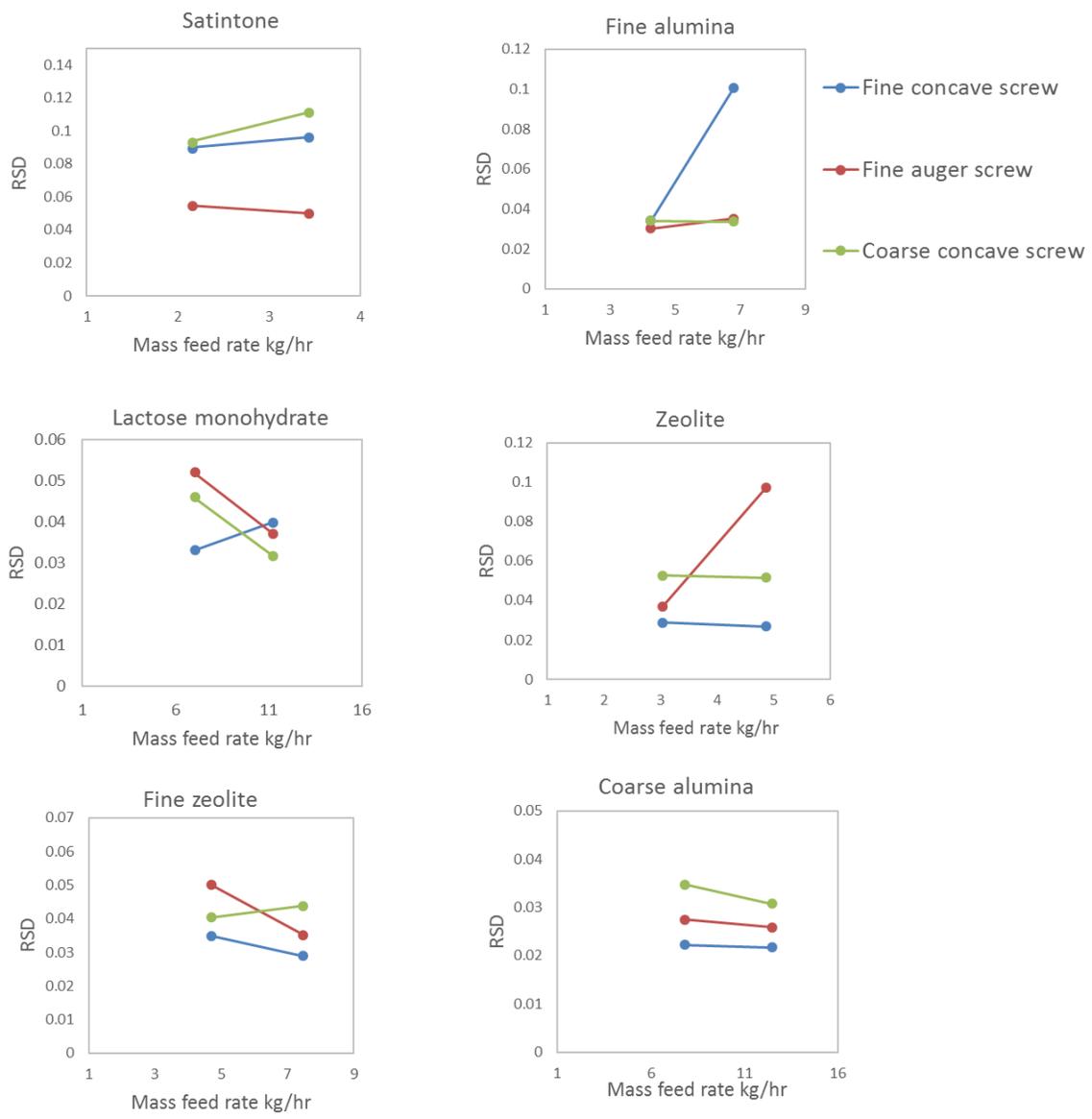
Figure 4-9 Gravimetric feeding performance, represented by relative standard deviation (RSD), for each material using three available screws.
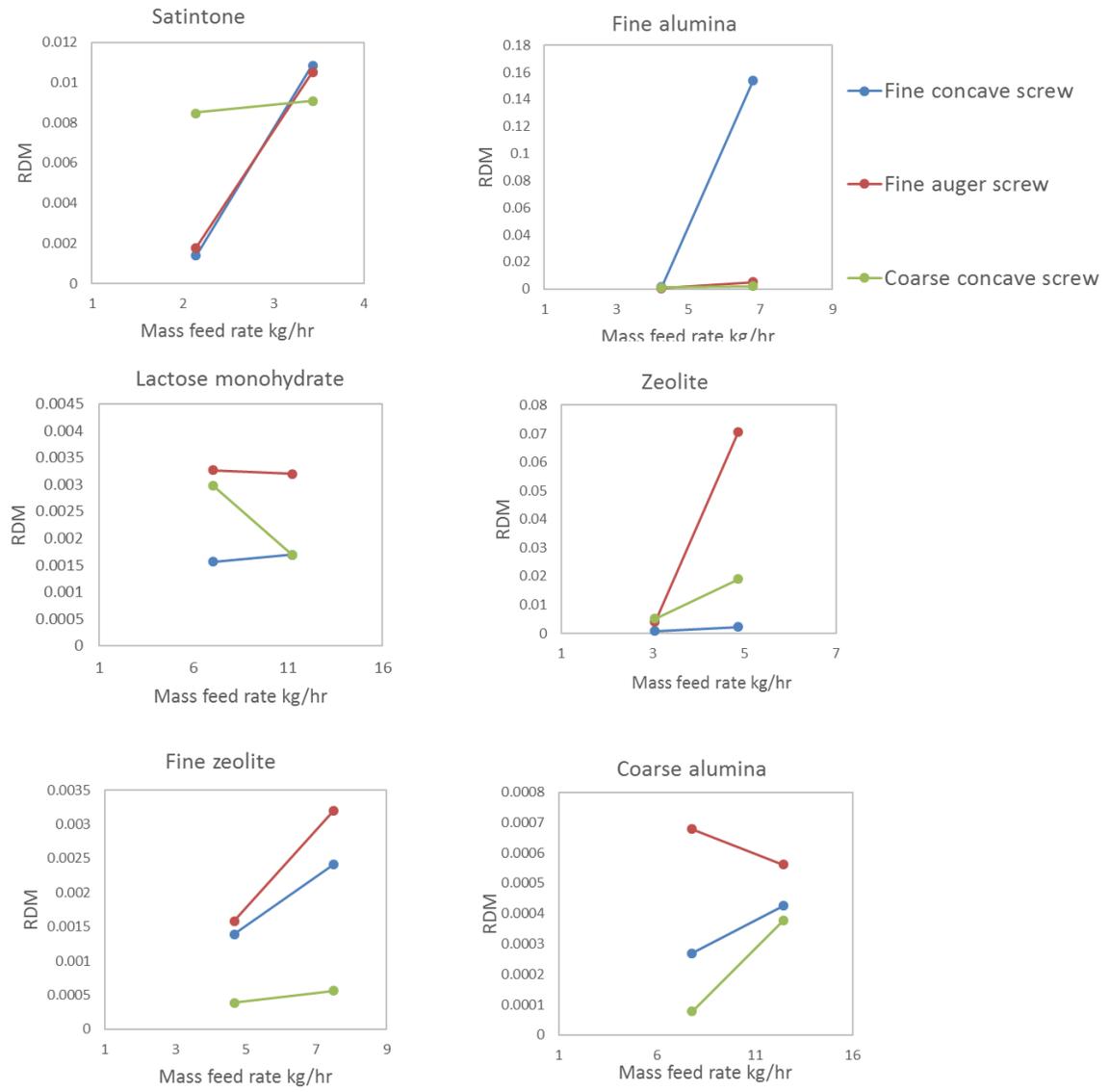
Figure 4-10 Gravimetric feeding performance, quantified by relative deviation from the mean (RDM), for each material using three available screws.
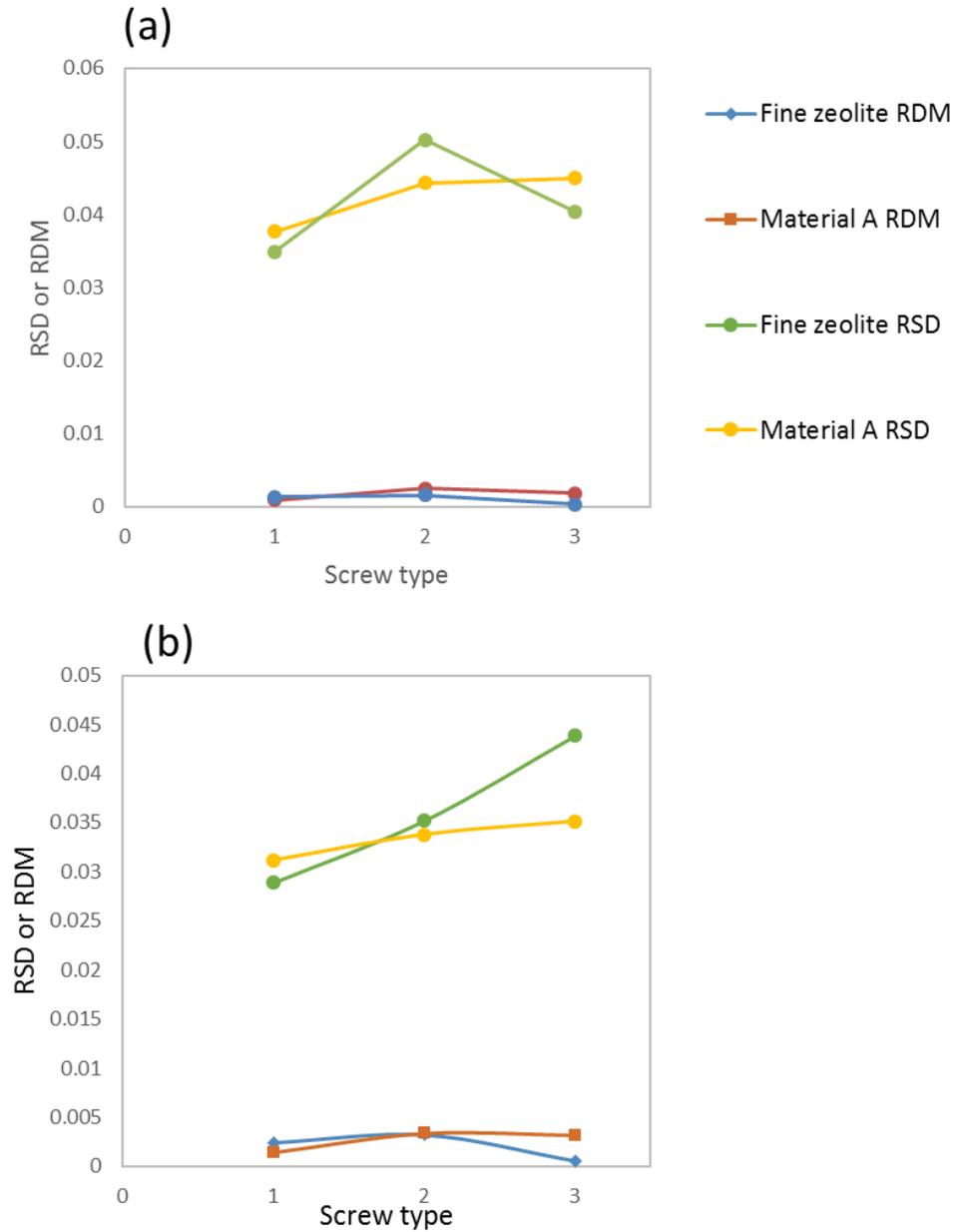
Figure 4-11 Comparison of feeding performance between fine zeolite and Material A at (a) 50% initial feed factor of fine concave screw and at (b) 80% initial feed factor of fine concave screw. Screw type 1 corresponds to fine concave screw, screw type 2 corresponds to fine auger screw, and screw type 3 corresponds to coarse concave screw.
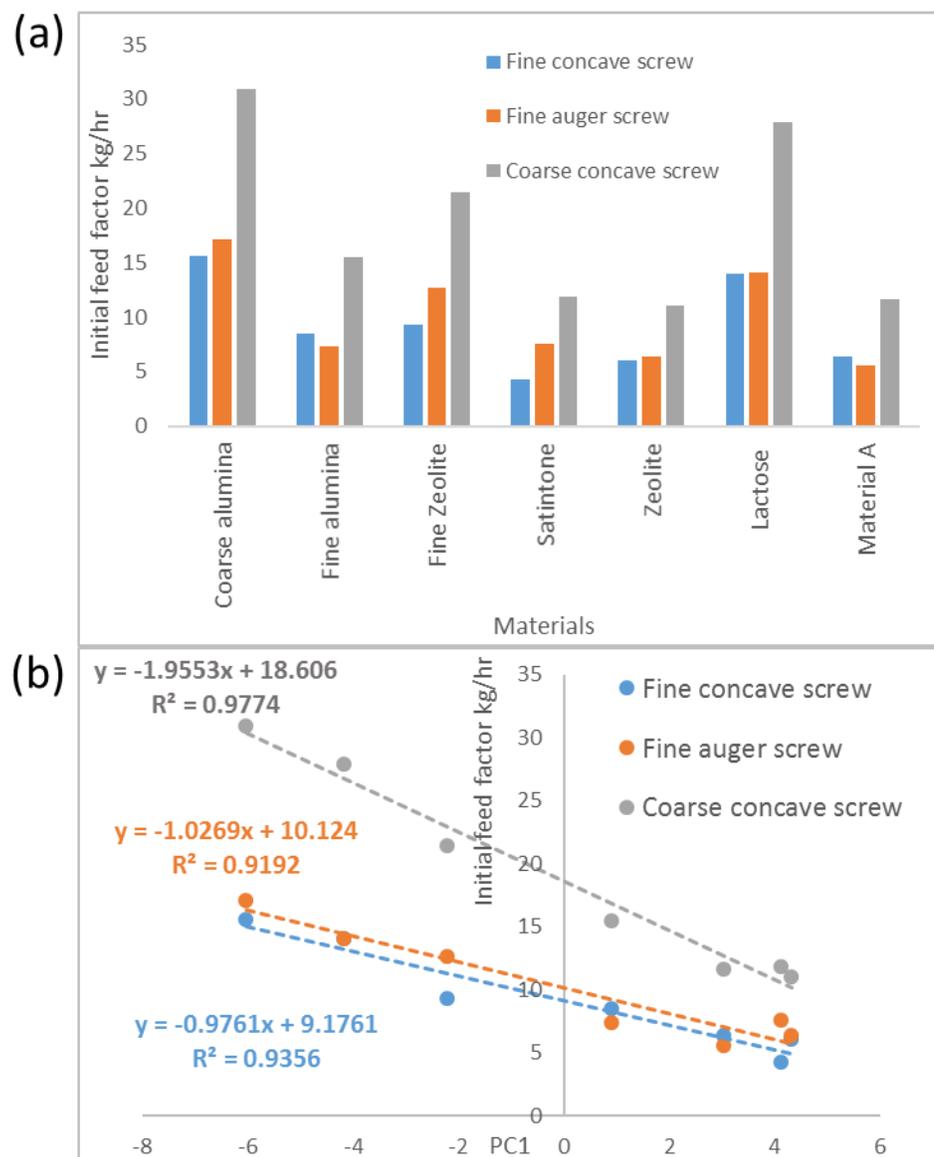
Figure 4-12 (a) Initial feed factor during volumetric calibration for each material using three available screws. (b) A strong linear relation between the obtained initial feed factor and the scores of the first principal component.
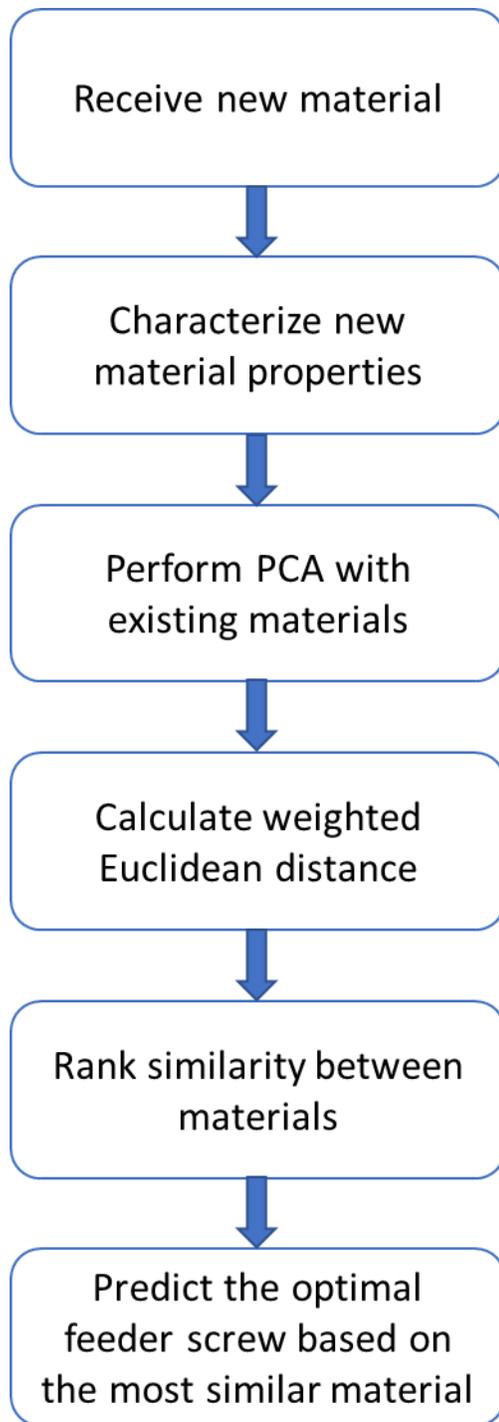
Figure 4-13 Workflow of principal component analysis - similarity score (PCA-SS) approach.
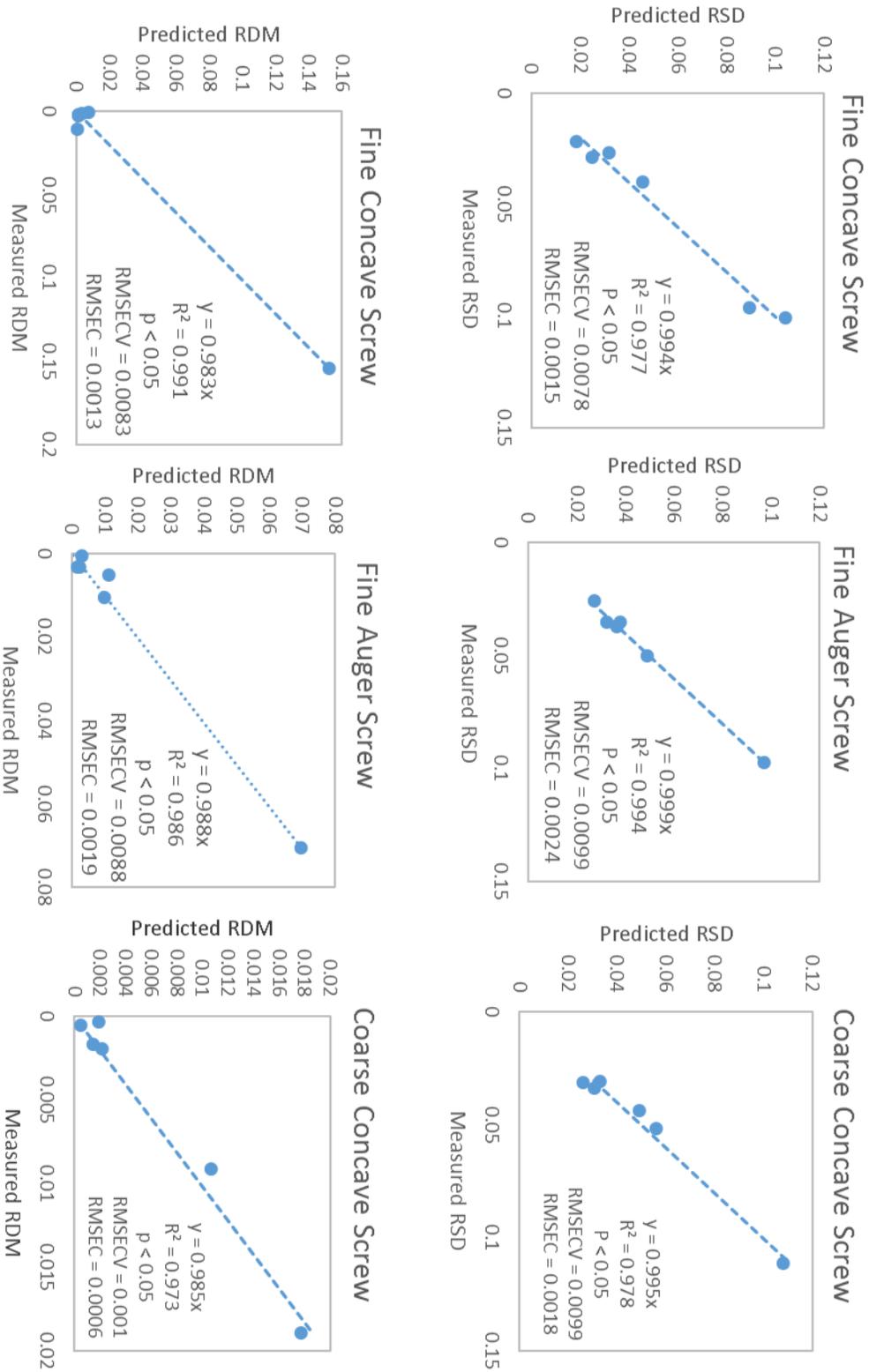
Figure 4-14 Predicted versus reference parity plot for feeding performance, represented by RSD and RDM, for three screws.
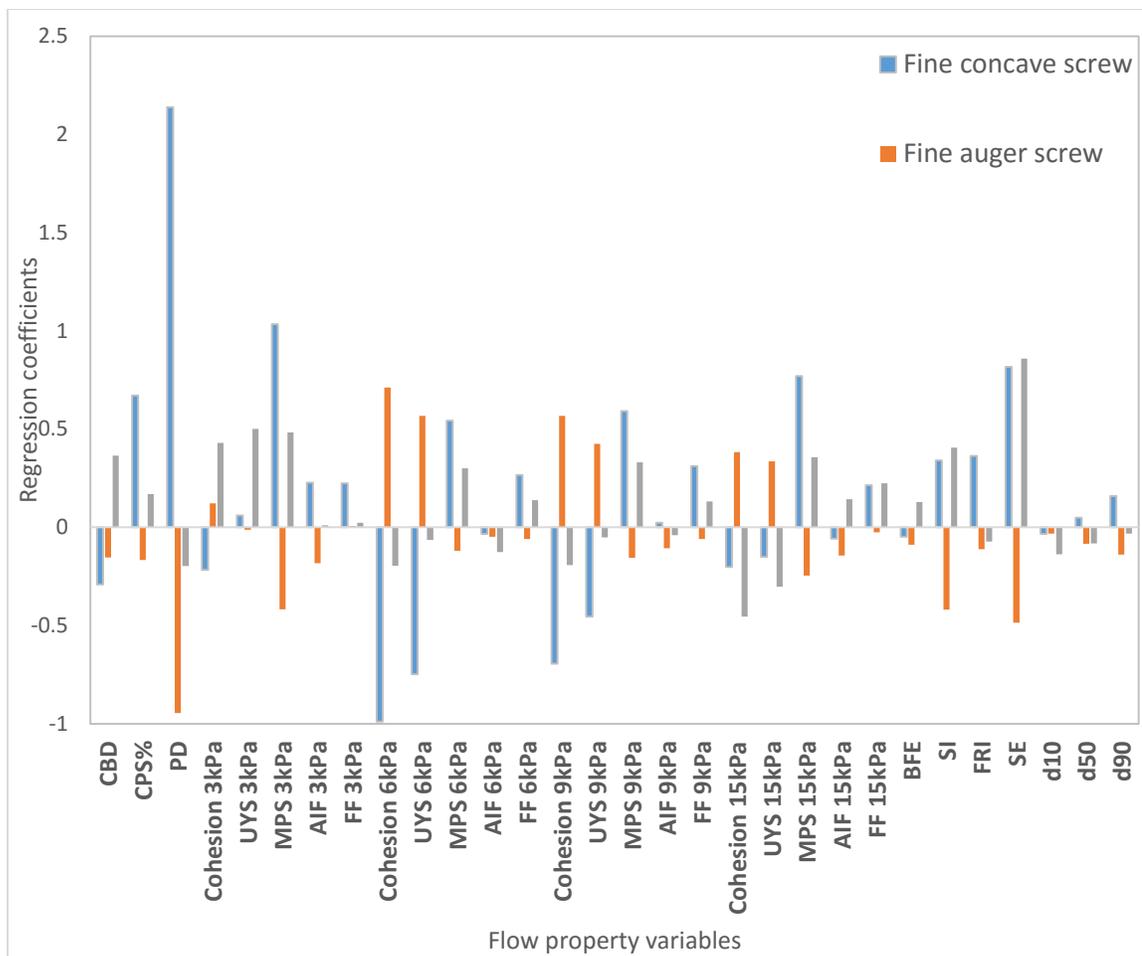
Figure 4-15 Regression coefficients for PLSR models of RSD prediction.
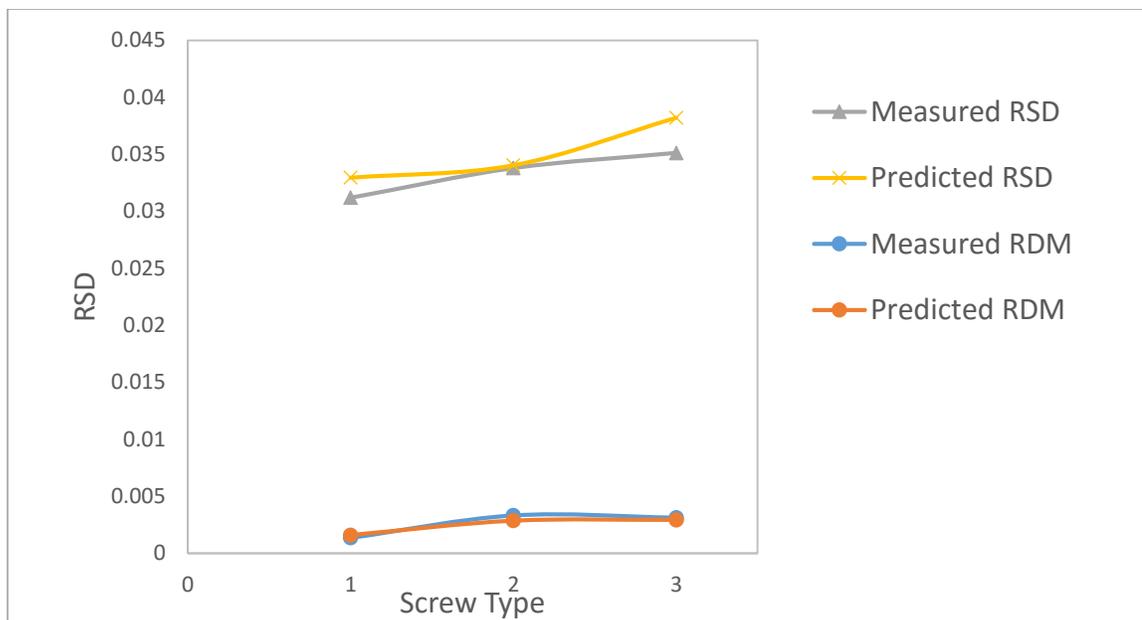
Figure 4-16 Predicted feeding results for Material A, in comparison with experimental results. Screw type 1 corresponds to fine concave screw, screw type 2 corresponds to fine auger screw, and screw type 3 corresponds to coarse concave screw.
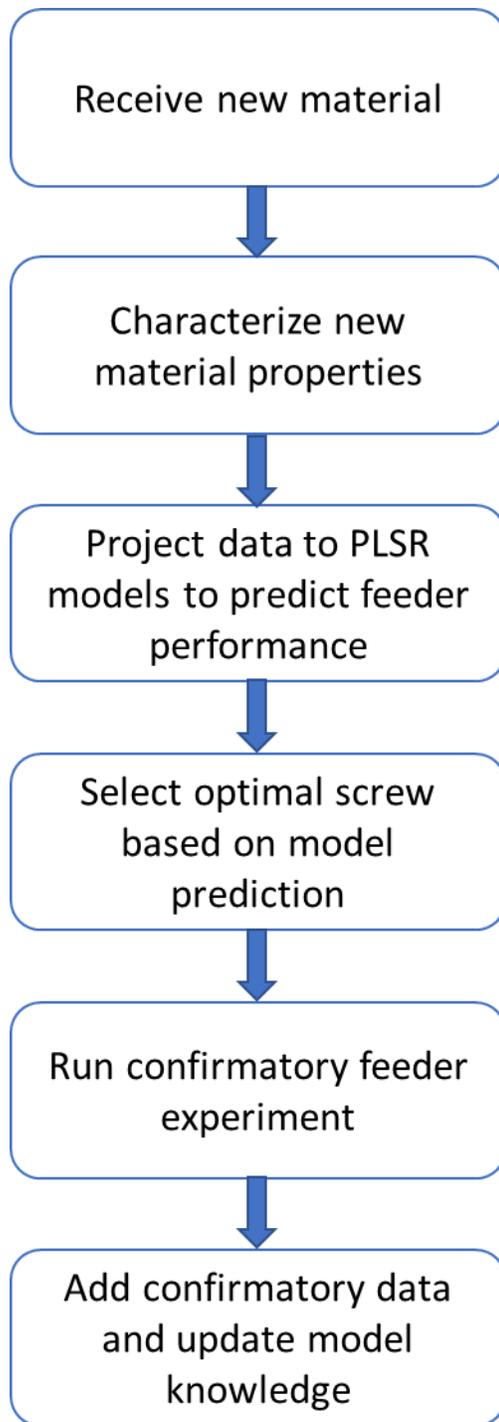
Figure 4-17 Workflow of partial least squares regression (PLSR) approach to predict feeder performance based on material flow properties.

5. Predicting Flow Behavior of Pharmaceutical Blends Using Shear Cell Methodology: A Quality by Design Approach

5.1. Introduction

The ability to understand, evaluate, and most importantly predict powder flow performance is critical for formulation development and process design of solid dosage forms in the pharmaceutical industry. During product development, flow properties are tested routinely to achieve desired manufacturability. This process often includes evaluation the flow performance of APIs, blends, and sometimes even excipients. Poor flow performance can result in multiple problems during bulk solids processing, such as arching in hoppers, segregation and/or agglomeration after mixing, and content variability in final dosage forms.[166] Powders are complex materials. Flow measurement is more of a functional performance test than a physical test,[167] and there have been many indices developed to indicate powder flow properties, such as angle of repose, compressibility index, Hausner ratio, flow through an orifice, and parameters from the shear cell tests. Another widely used dynamic test, the avalanching test, has been correlated to compressibility and shear cell results in a recent comparative study.[109]

Among the various characterization techniques, one of the commonly used is the shear cell methodology, which was originally developed by Jenike for design of hoppers and silos.[113, 168] A flow regime was proposed to describe the limit when powders jam and form an arch at the opening of a hopper. Useful information, such as cohesion, flow factor, unconfined yield strength and angle of internal friction, can be extracted from the test to guide process design. Jenike's mathematical analysis to determine hopper angle

and opening size has become an engineering standard practice.[169] Shear cell testing has been therefore used extensively for flow properties measurement. The effect of particle size, shape, and density on flow properties have been well studied using the shear cell methodology.[120], [121] Effects of the storage time and environmental factors, such as relative humidity and temperature, have also been reported.[122, 170] For example, Freeman showed repeatability in shear cell measurement,[18] and examined the effect of consolidation on shear properties and normal stresses.[118]

The quality by design initiative (QbD) of the U.S. Food and Drug Administration requires a process to be controllable and predictable.[11] Theories and methods to characterize powder flow have facilitated the implementation of QbD approaches to predict powder flow. Taylor *et al.* used principal component analysis based on five flow characterization methods to develop a method for material screening in early formulation development stage.[124] Both Niklas *et al.* and Yu *et al.* used principal component analysis and partial least square regression to predict powder flowability as a function of particle size and shape distribution.[88, 171] Although attention has been paid to characterizing and predicting flow properties of a mixture, very few cases were presented to address the scenario during early formulation development when the amount of drug is limited and more than two ingredients are in the formulation.

The intention of this chapter is to demonstrate a general QbD approach to quickly classify flow properties of a mixture during formulation development. In addition, this chapter also aims to validate the analysis methods, proposed in Chapter 3, by including measurement results of powder mixtures. Four ingredients, including one model drug and three excipients, were used to represent a realistic formulation design process. This

approach also addresses the challenge in using minimal materials during early formulation development.

5.2. Material and Methods

5.2.1.   Material

The materials used in this study were: Semi-fine acetaminophen (Mallinckrodt, Raleigh, North Carolina, USA), microcrystalline cellulose (Avicel PH102, FMC Biopolymer, Newark Delaware, USA), Fast-Flo® lactose (Monohydrate N.F. modified-spray dried, Foremost Farms USA, Rothschild, Wisconsin, USA), and Regular lactose (Monohydrate, Foremost Farms USA, Rothschild, Wisconsin, USA). Particle size information of the reported materials is listed in **Table 5-1**. The particle size distribution of each ingredient was measured using a laser diffraction analyzer (LS-13320) with a Tornado Dry Powder System (Beckmann-Coulter, Brea, California, USA).

5.2.2.   Design of experiments

A mixture extreme vertices design was used to characterize flow properties of the four-component system.[172] Constraints for concentration of Acetaminophen (APAP, x_1), MCC(x_2), Fast-Flo lactose (x_3) and Regular lactose (x_4) are expressed as follows:

$x_1 \leq 0.45$

$x_2 \leq 0.40$

$x_1 + x_2 + x_3 + x_4 = 1.0$

The MINITAB® Release 16 (Minitab Inc.) software was used to aid the design and 18 conditions were generated. The four raw materials were also included to the design

adding up to 22 conditions in the study. **Figure 5-1** shows the mixture design structure. Detailed formulation for each blend is listed in **Table 5-2**. Blends 13 and 14, 15 and 16, 17 and 18, 19 and 20 were four pairs of replications to evaluate experimental variations.

5.2.3. Blend preparation

To minimize amounts of materials used in the study (as is often the requirement in early formulation development), 100 g of each blend was prepared. Before blending, each material was passed through a no. 18 sieve (sieve size of 1.0mm) to enhance blend homogeneity by breaking up agglomerates. Blends were prepared in a laboratory scale ResonantAcoutic ® Mixer (RAM, Resodyn Acoustic Mixers, Butte, Montana, USA) which uses low frequency and high intensity acoustic energy to induce mixing and allows for sufficient mixing for small-scale blends. [173, 174] The effect of LabRAM mixing on blend flow properties will be further investigated in Chapter 6.

5.2.4. Shear cell methodology

The flow properties of the blend were characterized by shear cell test at an initial consolidation stress of 9 kpa. A detailed description of the shear cell test can be found in Chapter 3. The value of ff_c can be used to characterize flowability numerically. Larger ff_c indicates better flow performance. A criterion has been defined to classify flow behavior by Schulze: [175]

$ff_c < 1$, not flowing;

$1 < ff_c < 2$, very cohesive;

$2 < ff_c < 4$, cohesive;

$4 < ff_c < 10$, easy-flowing;

10 < ff$_c$, free-flowing.

Importantly, values of ff$_c$ depend on the actual device used to measure it, although ranking orders of materials tend to be preserved[123, 176]. Both flow factor and cohesion were calculated in the experiment. The consolidation stress and pre-shearing normal stress in the experiments was 9kPa. Normal stress for shearing at 7kPa, 6kPa, 5kPa, 4kPa, and 3kPa were used, the pre-shearing/shearing process is repeated to generate five yield points. To minimize samples required, a standard 25mm x 25ml split-vessel was used and sampling replicates of 3 were conducted for each condition.

5.3. Results and Discussion

5.3.1.   Correlation of Measured Flow Indices

In chapter 3, the power correlation between the flow function coefficient and cohesion was discussed. **Figure 5-2** validates the correlation observed in powder mixtures by showing that the flow factor and the cohesion (kPa) follow an inverse relation with a proportionality constant of 4.2. As the initial consolidation stress performed during the test was 9kPa, the coefficient from the ffc-C* plot in this study is 0.467, which is reasonably close to the coefficient (0.485) in Figure 3-5 for raw materials. In other words, the intrinsic power correlation between the ffc and C* can be expanded to powder mixtures. This observed relationship can be explained by the definition of the flow factor. As introduced previously, the flow factor is defined as the ratio between major principal stress (MPS) and unconfined yield strength (UYS). Major principal stress was kept constant due to the same initial consolidation stress. For blends with a linear yield locus, the unconfined yield strength has a strong linear relationship with the cohesion, as shown in **Figure 5-3**. The flow factor can be rewritten as the reciprocal of the cohesion with a

constant related to MPS and UYS. The classification of flow behavior can thus be extended to cohesion as follows:

Cohesion < 0.42 kPa, free-flowing;

0.42 kPa < Cohesion < 1.05 kPa, easy- flowing;

1.05 kPa < Cohesion < 2.10 kPa, cohesive;

2.10 kPa < Cohesion < 4.20 kPa, very cohesive;

Cohesion > 4.20 kPa, not flowing.

5.3.2.   Model Selection

Four regression models were compared, as shown in **Table 5-3**. Having four components in the formulation made linear and quadratic models less robust. The full cubic model generated the best fitting results by including 18 terms. Since the calibration model was built from 22 sample blends and over-fitting needs to be avoided, the full cubic model was not selected in our case. The special cubic model was selected because it was sufficient for the purpose of prediction flow classification.

A probability plot is shown in **Figure 5-4** to test the normality assumption of residuals for analysis of variance (ANOVA). The probability plot can provide a graphical view for the normal distribution. The linearity of the scattered points in the plot suggested a normal distribution for cohesion residuals. The figure also shows that the residuals of the flow factor model were not distributed normally. The model was less robust to predict flow factor than cohesion unless certain transformation technique was performed. Since the flow factor and the cohesion were highly correlated, the regression model for cohesion was used. Analysis of Variance (ANOVA) was used for data analysis, which

was discussed in Chapter 2. **Table 5-4** shows ANOVA results including linear, quadratic and special cubic terms. Term APAP*FastFlo (p=0.31) and MCC*FastFlo*Regular (p=0.56) were excluded for model reduction leaving 12 terms in total. Adjusted R-squared showed that 74.95% of total variability in the dataset could be explained by the model. The reduced model had higher adjusted R-squared value compared to the full special cubic model (R-squared = 64.87%). Based on the ANOVA analysis, cohesion can be predicted with the following equation:

$$Cohesion = 1.71\, x1 + 0.23\, x2 + 0.60\, x3 + 0.74\, x4 + 5.71\, x1\, x2 + 1.62\, x1\, x4$$
$$+ 4.11\, x2\, x3 + 3.86\, x2\, x4 - 14.00\, x3\, x4 - 25.80\, x1\, x2\, x3$$
$$- 47.97\, x1\, x2\, x4 + 40.05\, x1\, x3\, x4$$

Experimental measurements were compared to the estimated values as shown in **Figure 5-5**. The diagonal dashed line represents the points at which the measured values equals to the predicted values. The regression parameters show that the model prediction is statistically consistent with the reference values ($p < 0.05$)

5.3.3.   Model applicability

The model can be used to estimate flow performance during formulation screening. For example, Figure 8 shows contour plots of predicted cohesion with varying excipient combinations at constant APAP concentration. **Figure 5-6(a)** shows the flow regime when concentration of APAP was kept constant at 27% and 5-**6(b)** shows corresponding results at 45%. Two highlighted points were selected as external validation to predict flow classification. Blends were prepared and tested with shear cell methodology using the same procedure described previously. The predicted and measured cohesion values

are shown in **Table 5-5**, along with formulation details. Formulation 1 was confirmed by experiment to be easy flowing, while formulation 2 was found to be cohesive. This approach can be particularly useful when cohesive formulations are avoided during formulation screening. For the purpose of model maintenance, once a new data point is tested and validated, it can be added to the model to help extend the applicable range and improve model robustness. In addition, if each calibration sample is measured under different levels of initial consolidation stresses, the model can be expanded to predict the blend cohesion at a certain consolidation stress for process development.

5.4. Conclusion

A model was built to predict the flow regime of pharmaceutical blends based on a four-component mixture design. A special cubic model was established using 22 calibration blends including four-component powder mixtures and raw materials. The measured flow indices from the shear cell test were highly correlated. For example, Cohesion was in a power correlation to the flow factor and in linear relation with unconfined yield strength. The flow classification based on flow factor was thus extended to cohesion. Normality test of residuals was performed to examine model robustness. It was found that the regression model for flow factor did not have normal distribution of residuals.

The work presented here has shown an effective approach to characterize flow properties of powder mixture using statistical methods. The motivation for Quality by Design is to quickly determine the design space using minimum materials and time. Shear cell methodology requires a relatively small amount of material and a mixture design assures minimum amounts of blend prepared. This approach can be used to screen through a large number of formulations to avoid cohesive formulations with little materials or

manpower cost. To illustrate model applicability, two external validation points were prepared. The predicted and measured cohesion values were consistent.

In this chapter, the prediction model only includes formulation variables since mixing of the powder mixtures was set constant. Chapter 6 further supplements the prediction model by considering the effects of lubrication variables on blend flow properties. Also important to notice is that the same type of data can be augmented by additional measurements performed during scale up and validation of the blending process, and during commercial manufacturing of a product. In this context, it might also be possible to expand models to include process variables. Model parameters can then be refined and/or the model itself can be upgraded, consistent with the spirit of the recent continual process verification regulation, to capture process knowledge and increase predictability.

5.5. Figures for Chapter 5
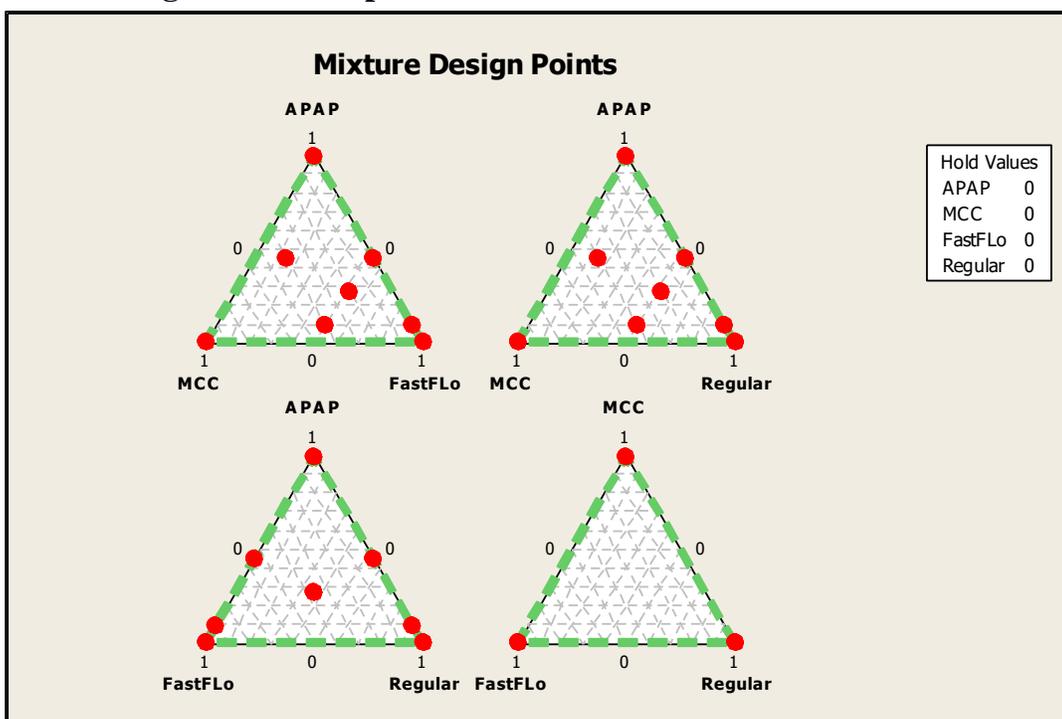
Figure 5-1 Design of experiments. 18 blends were generated from extreme vertices design. 4 raw materials were also included adding up to 22 design points.
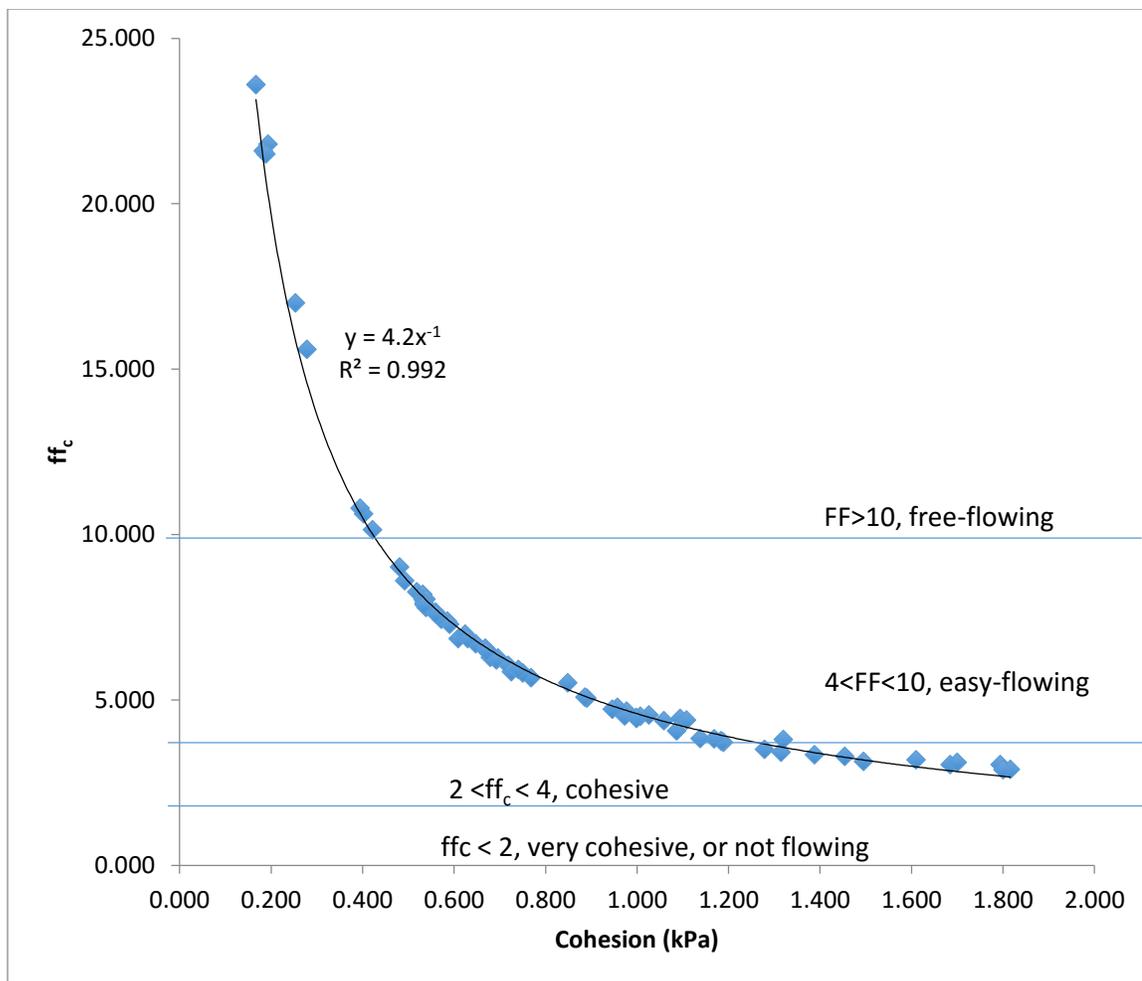
Figure 5-2 A power relation was observed between ffc and cohesion (kPa). Using this correlation, the classification of flow behavior based on ffc can be extended to cohesion.
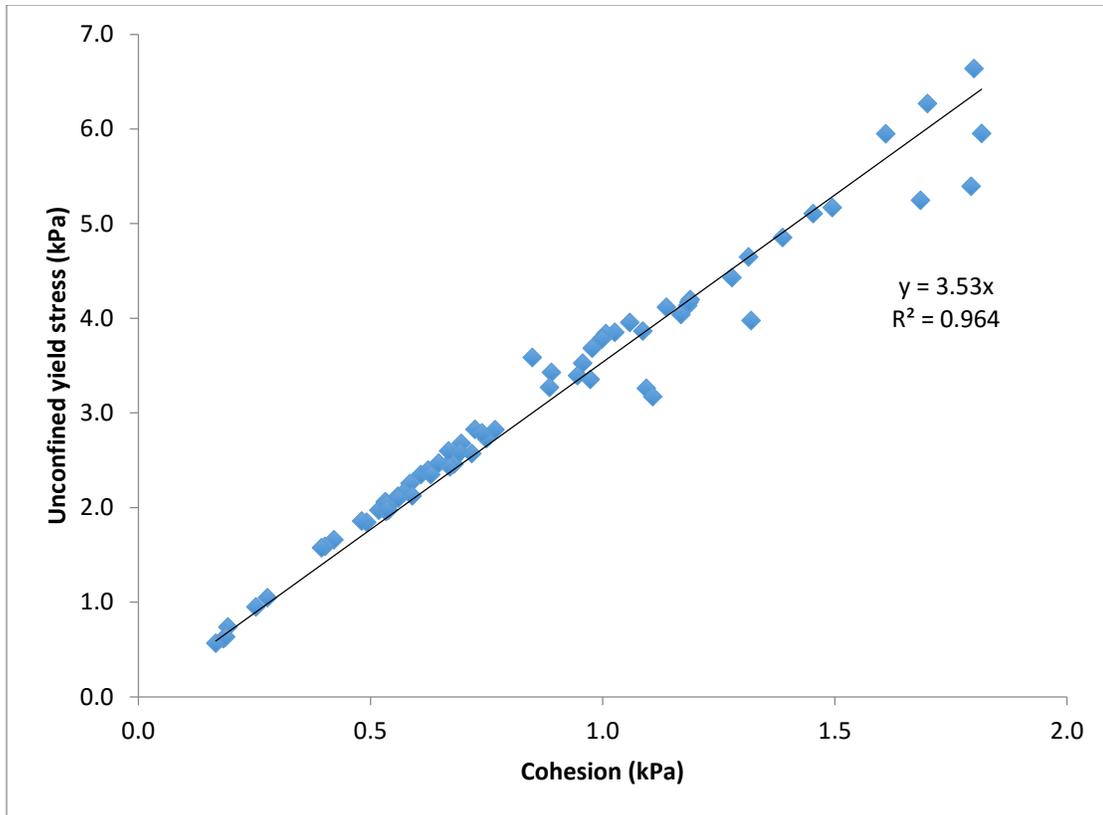
Figure 5-3 A linear relation between unconfined yield strength and cohesion explains the power correlation between flow factor and cohesion.
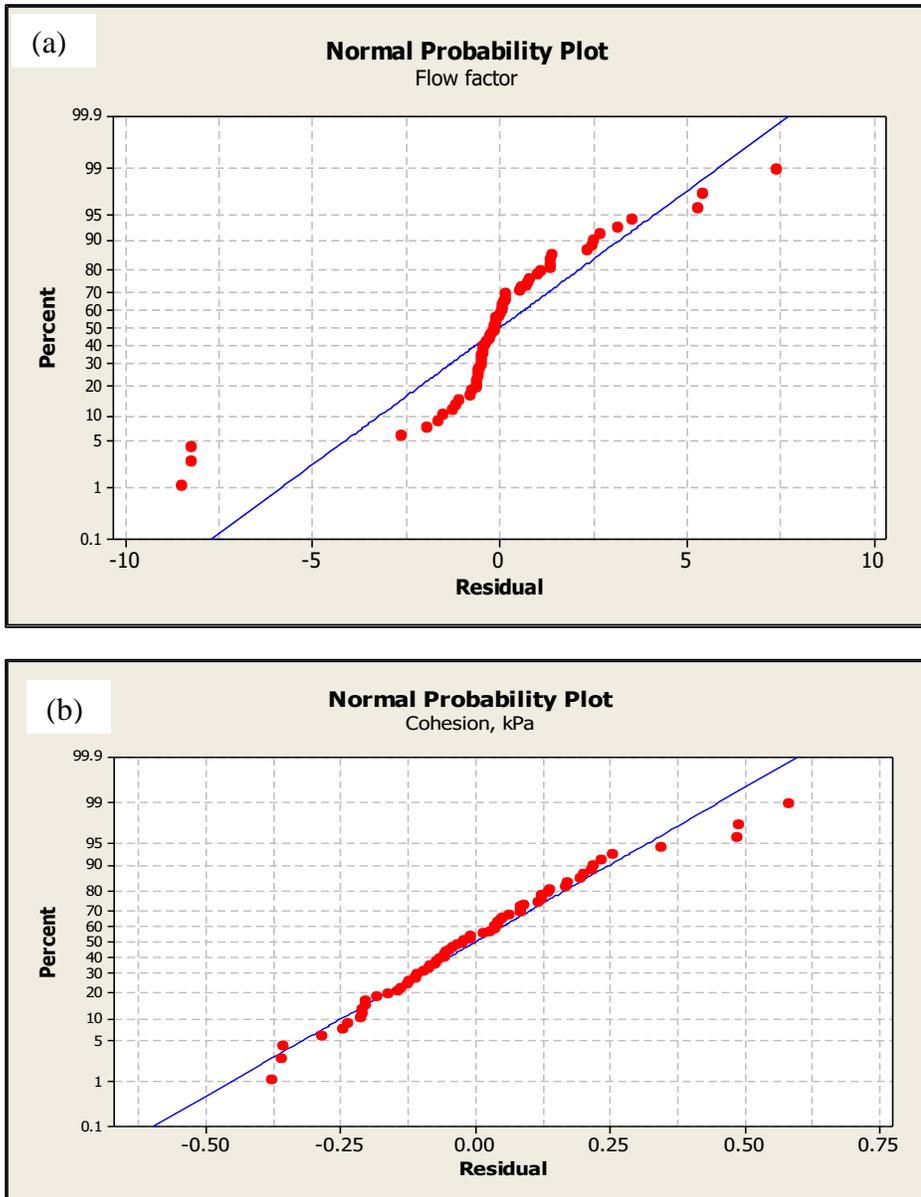
Figure 5-4 Normality plot to test normality assumption for analysis of variance (ANOVA) for (a) flow factor and (b) cohesion. Regression model for cohesion was more robust than flow factor.
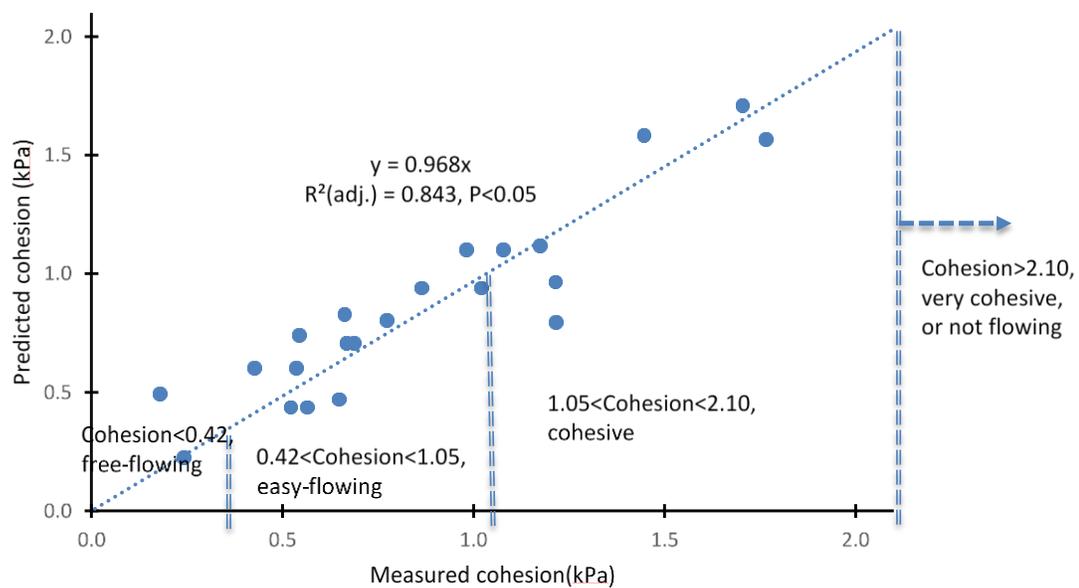
Figure 5-5 Comparison of experimental and estimated cohesion with classification criteria shown. .
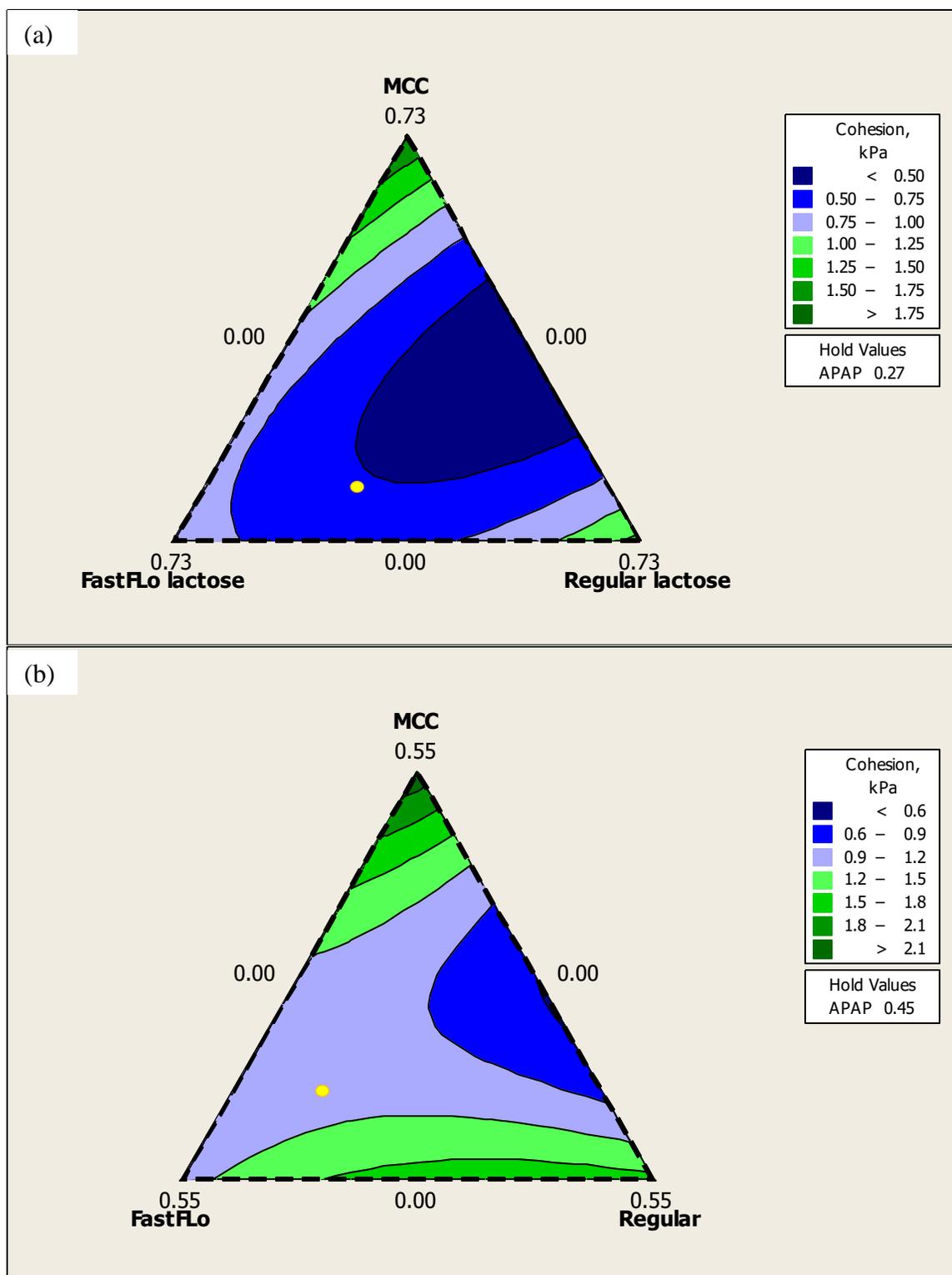
Figure 5-6 Contour plots of cohesion (kPa) at (a) APAP=27%, (b) APAP=45%. Highlighted dots were selected for external validation.

5.6. Tables for Chapter 5

Table 5-1 Particle size information for materials used in this study.

Material	Mean(µm)	d10(µm)	d50(µm)	d90(µm)
Semifine Acetaminophen	48.9	5.6	32.6	122.7
MCC	141.2	34.0	120.8	244.1
Fast-Flow Lactose	114.5	54.2	113.3	174.6
Regular Lactose	71.9	10.3	63.5	157.7

Table 5-2 Design points for mixture model: formulation for each blend was shown.

Blend	APAP (x_1)	MCC (x_2)	FastFlo lactose (x_3)	Regular lactose (x_4)
1	1.000	0.000	0.000	0.000
2	0.000	0.000	0.000	1.000
3	0.000	1.000	0.000	0.000
4	0.000	0.000	1.000	0.000
5	0.090	0.000	0.000	0.910
6	0.090	0.000	0.910	0.000
7	0.450	0.000	0.550	0.000
8	0.450	0.000	0.000	0.550
9	0.090	0.400	0.510	0.000
10	0.090	0.400	0.000	0.510
11	0.450	0.410	0.000	0.140
12	0.450	0.410	0.140	0.000
13	0.270	0.000	0.365	0.365
14	0.270	0.000	0.365	0.365
15	0.270	0.200	0.265	0.265
16	0.270	0.200	0.265	0.265
17	0.450	0.200	0.175	0.175
18	0.450	0.200	0.175	0.175
19	0.270	0.400	0.165	0.165
20	0.270	0.400	0.165	0.165
21	0.270	0.200	0.530	0.000
22	0.270	0.200	0.000	0.530

Table 5-3 Four models were fitted to flow factor and cohesion respectively. Adjusted R-squared were used for model selection.

Model	Terms included	Number of Terms	Flow factor R-sq adj.	Cohesion R-sq adj.
1	Linear	4	35.24%	34.71%
2	Linear, Quadratic	10	48.62%	54.88%
3	Linear, Quadratic, Special cubic	14	74.54%	64.87%
4	Linear, Quadratic, Special cubic, Full cubic	18	96.75%	96.14%

Table 5-4 Analysis of variance of the mixture design. Effect of linear terms was not evaluated due to nature of mixture design. Terms of APAP*FastFLo and MCC*FastFlo*Regular were not included after model reduction due to insignificance effect.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	**13**	**9.476**	**9.476**	**0.861**	**18.68**	**0.00**
Linear	**3**	**4.574**	**3.934**	**1.311**	**28.43**	**0.00**
Quadratic	**5**	**2.092**	**2.296**	**0.459**	**9.96**	**0.00**
APAP*MCC	1	0.029	1.744	1.744	37.81	0.00
APAP*Regular	1	0.089	0.339	0.339	7.34	0.009
MCC*FastFLo	1	0.283	0.779	0.778	16.88	0.00
MCC*Regular	1	0.355	0.671	0.671	14.56	0.00
FastFLo*Regular	1	1.337	0.296	0.296	6.41	0.014
Special cubic	**3**	**2.809**	**2.809**	**0.936**	**20.30**	**0.00**
APAP*MCC*FastFLo	1	0.224	0.805	0.805	17.46	0.00
APAP*MCC*Regular	1	2.404	2.566	2.566	55.65	0.00
APAP*FastFLo*Regular	1	0.180	0.180	0.180	3.91	0.053
Residual	**54**	**2.490**	**2.49**	**0.046**		
Total	**65**	**11.966**				
R-Sq = 79.19%, R-Sq(adj) = 74.95%						

Table 5-5 Two points were selected from contour plots as an example of model application for flow performance estimation during formulation design.

Formulation	APAP	MCC	FastFlo Lactose	Lactose Regular	Predicted Cohesion (kPa)	Measured Cohesion (kPa)
1	0.27	0.11	0.35	0.27	0.53	0.59
2	0.45	0.12	0.32	0.11	1.06	1.10

6. Controlled Shear System and Resonant Acoustic Mixing Effects on Lubrication and Flow Properties of Pharmaceutical Blends

6.1. Introduction

Achieving desired powder flowability is critical for powder handling and processing. In the pharmaceutical industry, the active pharmaceutical ingredient, which is often cohesive, is mixed with excipient materials through a series of unit operations before the delivery of final solid dosage forms [7]. Lubricants, such as magnesium stearate (MgSt), or glidants such as Silicon dioxide (SiO_2) are commonly included during formulation development. MgSt interacts with other materials during lubrication process to improve the flowability of the blends, and therefore to ensure reproducible final products with acceptable weight variation and content uniformity [177, 178]. The mechanism of lubrication has been explained by the formation of lubricant layers on the large particles, reducing frictional forces, cohesive forces or both between particles [179, 180]. Studies have shown that lubrication significantly affects the density and compactability of blends [181, 182]. Properties of the final solid dosage forms, such as tensile strength and *in vitro* drug release, are also heavily dependent on the lubrication process [82, 183, 184].

Controlling the extent of shear, or energy, that the blend experiences during lubrication is necessary to achieve the desired lubricity and flowability and to avoid over-lubrication [185]. Sufficient energy during mixing is necessary to homogenize cohesive ingredients by de-lumping agglomerates and redistributing them within the blend matrix. However, excessive energy can lead to electrostatic buildup and over-lubrication [186]. Studies have shown that tablet properties made from over-lubricated blends can be adversely affected [187-189]. Lubrication can take place in numerous unit operations that are

intrinsically associated with shear mixing, such as blending and feeding in a tablet press [190]. Two process variables are important for lubrication: the total energy per unit mass, and the rate of energy being applied. Although the rate is difficult to measure inside a blender, several studies have shown that these variables can be indirectly correlated to operating conditions, such as mixing time, fill level and rotation speed, and are used for process scale-up [191-193].

In recent studies, two laboratory scale devices have been used substantially to investigate the effect of lubrication variables on blend and tablet properties: the controlled shear system and the Resonant Acoustic Mixer (RAM). The controlled shear system, also known as the modified Couette shear cell, provides a controlled and uniform shear environment to the blends by equally spaced interlocking pins creating a homogeneous shear filed. Both total shear, quantified by total revolutions, and shear rate, quantified by the rotational speed, can be controlled. The principles by which the controlled shear system works are explained in more detail elsewhere [12, 185]. Studies show that increasing total shear improves the powder flowability and decreases tablet hardness. The shear rate was found to have smaller effect than total shear on the blend and tablet properties.[12, 186]   The RAM uses low frequency and high intensity acoustic energy to induce mixing and allows for sufficient mixing for small-scale blends [194]. The total energy during mixing can be easily assessed by controlling the energy input rate and the mixing time. Our previous work discussed the blending performance observed in a laboratory-scale resonant acoustic mixer (LabRAM) and investigated the effect of process parameters on the material properties of blends and tablet performance. Studies suggest that improved flowability and increased wettability of the blend can be obtained

with increasing energy input rate and blending time. The tableting performance and tablet properties were found to be correlated to the total energy. [195, 196].

Although both devices have been well studied, to our knowledge, there is no published work that systematically compares and correlates the two devices. Based on our understanding of the shear mixing mechanism, it is expected that the observed lubrication using one device will agree with that from the other. However, questions that remain unanswered include: 1) Are the lubrication process of the two devices, controlled by different operation parameters, processed at different scales, and measured in different units, comparable? 2) Can final blends with the same flow properties be made using different device by adjusting process parameters? 3) More importantly, when comparing two devices that have different mixing principles, what properties can be correlated and what properties cannot? Considering how common and critical lubrication is, it is important to answer these questions using a science-based process and product development approach [11].

The objective of this chapter is to systematically compare the lubrication effects of the controlled shear system and the resonant acoustic mixer on the flow properties of pharmaceutical blends. Importantly, understanding the effects of lubrication variables on blend flow properties supplements the mixture model, introduced in Chapter 5, to predict blend cohesion by considering both formulation and processing variables. A model formulation consisting of a drug substance, a filler and a lubricant was used. Two factorial experimental designs were carried out to fully characterize the effect of the total shear in the controlled shear system and the total energy in the LabRAM, and the lubrication rate on the blend flow properties. Due to the complexity of powders, many

measurements have been developed to elucidate their flow properties. The flow properties in this study were characterized by density, compressibility and flow indices extracted from the shear cell tests [18].

6.2. Materials and Methods

6.2.1. Materials

The blend consisted of semi-fine acetaminophen (Mallinckrodt Inc. Raleigh, North Carolina) as the active pharmaceutical ingredient in the formulation, lactose monohydrate NF (Foremost Farms, Rothschild, Wisconsin) as the filler, and magnesium stearate NF (Mallinckrodt Inc. St Louis, Missouri) as the lubricant. The particle size information of the materials is listed in **Table 6-1.** Particle size was determined using a laser-diffraction (LS-13320) analyzer with a Tornado Dry Powder System (Beckmann-Coulter, Brea, California).

6.2.2. Blending

The formulation in this study consisted of 90% w/w lactose, 9% w/w acetaminophen and 1% w/w magnesium stearate (MgSt). Prior to lubrication, 900 g of lactose and 90 g of semi-fine acetaminophen were mixed in a 1.87-L V-blender (Patterson Kelley, East Stroudsburg, PA) at 15 rpm for 15 minutes. 10 g of MgSt was then added to the pre-blend and mixed further for 2 minutes. The blend with MgSt without further lubrication, either in the controlled shear system or in the LabRAM, is referred to as "0Rev" blend in this study. The above steps were repeated five times so that in total 5 kg of 0Rev blends were prepared for the lubrication experiments. The flow chart of the experiment methods is shown in **Figure 6-1**.

To investigate the controlled shear system, a full factorial design covering three levels of shear rate and five levels of total shear was conducted. The variables to characterize the controlled shear system are listed in **Table 6-2**. A total of 250 g of blend was prepared each condition. For the LabRAM study, a full factorial design including three levels of intensity and four levels of total energy was performed. The experimental design is listed in **Table 6-3**. A total of 100 g of blend for each condition was prepared.

In addition, a control group of each device was prepared, which is referred to as the "reference blends." The reference blend was prepared by firstly mixing 90 g of acetaminophen and 900 g of lactose in the 1.87-L V-blender at 15 rpm for 17 minutes. The reference blends of the controlled shear system were then subjected to the five different shear levels from the experimental design (80, 160, 320, 640, and 1280 revolutions) at shear rate of 80 rpm. The reference blends for the LabRAM were subjected to the four different energy levels from the predefined experimental design (2000, 5000, 10000, and 50000 J/kg) at 60% intensity (10.4 Watts).

6.2.3. Blend Characterization

The bulk and tapped densities were measured using the standard procedure (ASTM Standard D7481-09).[197] The bulk density was measured using a 100-mL graduated cylinder. The cylinder was filled at 60 mL and the mass was weighed. An automatic tapping machine (Model No. AT.4.110.60, Quantachrome Instruments, Boynton Beach, Florida) was used for the tap density measurement. Three replications were performed for each experiment condition. In addition to particle size analysis, and density measurements, blend flow properties were characterized by compressibility test and shear cell test, which were discussed in details in Chapter 3 and Chapter 4.

6.3. Results and discussion

The effect of lubrication in the controlled shear system (CSS) and the resonant acoustic mixer (RAM) on the material and flow properties of the model blend under investigation was characterized and compared. Powder blend properties, namely particle size, bulk density, compressibility, cohesion, flow function coefficient (ffc), and the angle of internal friction, were compared. Selected powder flow property measurements were also correlated to one another to directly compare the effect of shearing (energy input rate) and total shear (energy) of the two systems under investigation. Statistical analysis was carried out to elucidate the effect of the process parameters on the final properties measured.

6.3.1.   The effect of lubrication on flow properties

Figure 6-2a shows the particle size distribution of the lubricated blends from the controlled shear system at shear rate of 80 rpm. **Figure 6-2b** shows the particle size distribution of the lubricated blends from the LabRAM at intensity of 60% (10.4 Watts). Blends from other lubrication rates had similar results and therefore are not shown here. Results suggested that the particle size distribution did not change significantly compared to the 0Rev blend. During mixing in lubrication, adhesion forces and possibly electrostatic effects may result in larger agglomerated particles. These particles could also break down to smaller particles during mixing. From the particle size measurements, the distribution did not show a marked difference, indicating that the two mechanisms were acting simultaneously, and neither was dominating.

The bulk density of the lubricated blends increased as the total shear or energy increased (**Figure 6-3**). Results of tapped density also showed increasing trends and therefore are

not shown here. The observed densification effect was independent of the lubrication device and the lubrication rate, and was heavily dependent on the presence of MgSt. For the reference blend, the bulk density decreased as the total shear or energy increased. Typically, a powder with a strong structural strength (i.e., a cohesive powder) will resist collapse and will have a low bulk density, while a structurally weak powder will collapse easily and have a high bulk density [198]. Therefore, although the particle size distribution was not changing as illustrated earlier, the lubrication process impacted the strength of interparticle forces

The compressibility measurements confirmed the change of interparticle forces as shown in **Figure 6-4**. The compressibility decreased as the total shear or energy increased. Notably, the results of the reference blends suggest that further mixing the pre-blend without adding lubricant significantly increased the compressibility of the powders. **Figure 6-5** shows the same findings for the cohesion from the shear cell test. Interestingly, the 0Rev blend, which had MgSt mixed for 2 min during experiment after pre-blending, was significantly more cohesive than the reference blend at zero revolution or energy ($p < 0.05$). Typically, the lubrication process reduces interparticle cohesive forces and improves the powder flowability. However, the findings indicate that insufficient shear or energy during lubrication may result in more cohesive materials. It is therefore important during process design and development stages to consider sufficient lubrication time, the appropriate time to add the lubricant, and the mixing order of the lubricant [199].

Figure 6-6 shows the effect of lubrication in both devices on the flow function coefficient. Based on the flow classification criteria, the lubrication process in the study

improved the flowability of the blend from relatively cohesive (ffc = 4.74) to free flowing (ffc > 10)[175]. Cohesive forces and frictional resistance are the two main interaction forces between pharmaceutical particles. As discussed earlier, as MgSt forms layers on the surface of large particles, the capillary cohesion between particles are disrupted. In addition, as particles are coated by lubricant, they will create smoother surface area with fewer sharp edges [200]. The frictional forces between particles will thus be smaller. This is confirmed by the results of the angle of internal friction as shown in **Figure 6-7**. Surprisingly, the angle of internal friction of the 0 Rev blend is similar to the reference blend at zero shear or zero energy ($p < 0.05$), implying that the decreased flowability due to insufficient lubrication is dominated by the increased cohesive forces rather than the frictional forces.

Lubricated blends increased in bulk density due to a reduction in capillary cohesion and Van der Waals forces by possible coating of the lubricant onto other "carrier" particles. MgSt is a dry and soft material that deforms under load and accommodates surface velocity differences by adhering to surfaces and shearing in the bulk medium. When sufficient energy during lubrication is provided, MgSt can coalesce and transfer a thin lubricating film to achieve enhanced lubrication performance.[201] For the unlubricated blends, further mixing of the cohesive drug (acetaminophen) might be causing increased capillary cohesion, Van der Waals forces, and electrostatic charging.[185] Experimental results agree with the two competing regimes during lubrication process, one where the cohesive materials dominate to increase the interparticle forces, and another where the lubricant is dispersive enough in the bulk medium to provide lubricating capabilities.[202]

6.3.2.   Statistical analysis

Statistical analysis, using analysis of variance (ANOVA) and omega-squared statistics, were used to test the effect of total shear or energy, the lubrication rate, and their interaction on the blend flow properties. The p-value for each factor and the two-way interaction was calculated using ANOVA. Typically, a factor with a p-value of less than 0.05 was considered statistically significant. Since the statistical significance of an effect depends on both the effect size and the sample size, knowing the magnitude of an effect conveys information on practical significance [203]. Calculation of ANOVA and effect size using omega-squared statistics can be referred to Chapter 2.

The ANOVA and omega squared results for the controlled shear system and for the LabRAM are tabulated in **Tables 6-4** and 6-**5**, respectively. Results show that shear rate in both devices had minimal effect on the flow properties. On the other hand, the effect of total shear, or total energy, during lubrication was significant ($p < 0.05$) and had much larger magnitude ($\omega^2 > 0.14$). A previous study showed that the rotation rate during convective mixing had little influence on the mixing process [204]. During lubrication, MgSt acquires sufficient energy to disrupt its adhesive forces and to be dispersed. Hence, the most important factor determining the overall extent of lubrication is the total amount of mechanical energy dissipated in the blend, rather than the rate of the energy being applied. Since the effect of shear rate is small, it can be adjusted during process scale-up development to accommodate the lubrication time, if necessary.

6.3.3.   Correlation between flow properties

Figure 6-8 shows that the lubricated blends from both devices fit into one linear correlation ($p < 0.05$) between the cohesion and the compressibility. Both flow indices

reflect cohesive forces between particles, and thus have a good correlation as lubrication weakens the cohesive forces. **Figure 6-9** shows two relatively parallel lines, indicating the correlation between the bulk density and the angle of internal friction. A previous study showed that the angle of internal friction is a material property that is insensitive to the change of normal stress applied to the material [135]. The results indicate that as lubrication processes densified the blend, the frictional forces changed in a reasonably similar rate in both devices. However, the correlations between the bulk density and the cohesion of the two devices were significantly different ($p < 0.05$) as shown in **Figure 6-10**. In other words, as the controlled shear system and the LabRAM work on different mixing principles, the frictional forces during lubrication can be well correlated, rather than the cohesive forces. Interestingly, mixing without adding MgSt significantly changed cohesive forces, while the frictional forces remained relatively constant, as can be observed from the reference blends in **Figure 6-9** and **Figure 6-10**. Therefore, the experimental results suggest that although the two devices had comparable lubrication effects on the overall blend flowability, the changes of the interparticle forces were not identical.

6.4. Conclusion

In this study, the lubrication effect of the controlled shear system and the laboratory scale Resonant Acoustic Mixer (LabRAM) was compared. Both devices enable direct control of the total shear, or energy, and the lubrication rate. A model formulation was selected to examine the effects of the process parameters on the blend flow properties. The lubricated blends were characterized by density measurement, compressibility test, and the shear cell test. Experimental results support the previous knowledge that the

lubrication significantly improves blend flowability. Reference blends were prepared as control sets to illustrate the effect of mixing on the blend flow properties with presence of MgSt. Results show that mixing without MgSt created more cohesive blends. In addition, early stage of lubrication made the blends more cohesive due to insufficient energy applied to disperse the MgSt and increased cohesive forces between particles.

Analysis of Variance (ANOVA) and effect size test using omega-squared were used to assess the relative magnitude of the process parameters. The statistical analysis suggests that the shear rate, which is controlled by the rotation rate in the controlled shear system and the intensity in the LabRAM, has minimal effect on the blend flow properties. In other words, the most important factor determining the overall extent of lubrication is the total energy transferred to the blend, rather than the rate of the energy being applied. The non-significant effect of the shear rate can be used during process design and development to reduce the lubrication time and to increase lubrication efficiency.

This study also represents an initial effort to demonstrate a scientific approach to compare lubrication processes, which have different mixing mechanisms and are different in scales, in a reproducible manner. Full factorial experimental design covering the operational space was conducted. Rigorous statistical analysis was performed to investigate the effect of lubrication parameters. Different flow characterization techniques were used to provide insight into the change of interparticle forces during lubrication. Results suggest that although the two devices had comparable lubrication effect on the overall blend flowability, the changes of the interparticle forces were not identical. The frictional forces can be well correlated with the bulk density of lubricated blends from both devices, while the correlations between the bulk density and cohesive

forces were significantly different. In other words, while the power relation between the flow function coefficient and the dimensionless cohesion is valid for both raw materials (Chapter 3) and powder mixtures (Chapter 5), the results from this chapter suggest that the correlation between different flow indices may be dependent on processing parameters for intermediate blends. Future work is needed on different formulations to test the robustness of the results and to generalize across available lubrication equipment.

6.5. Figures for Chapter 6
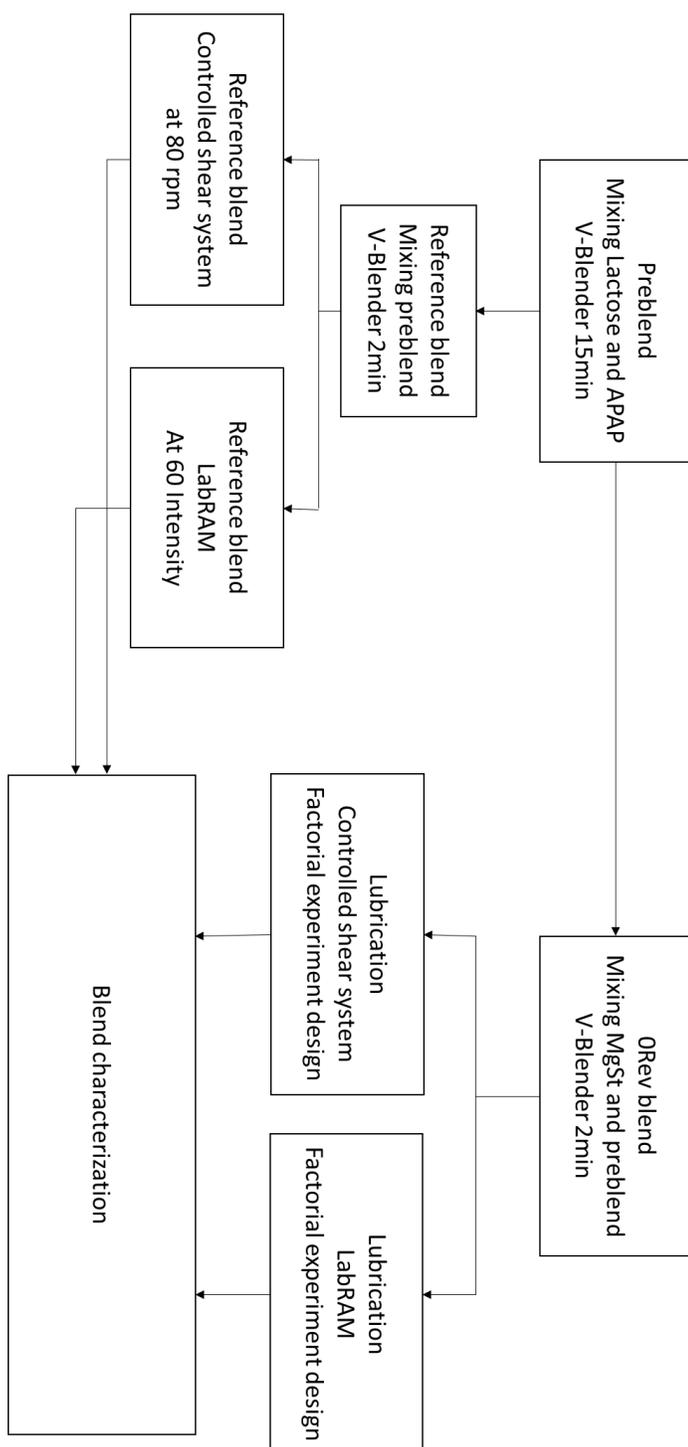
Figure 6-1 Experimental procedure in this study.
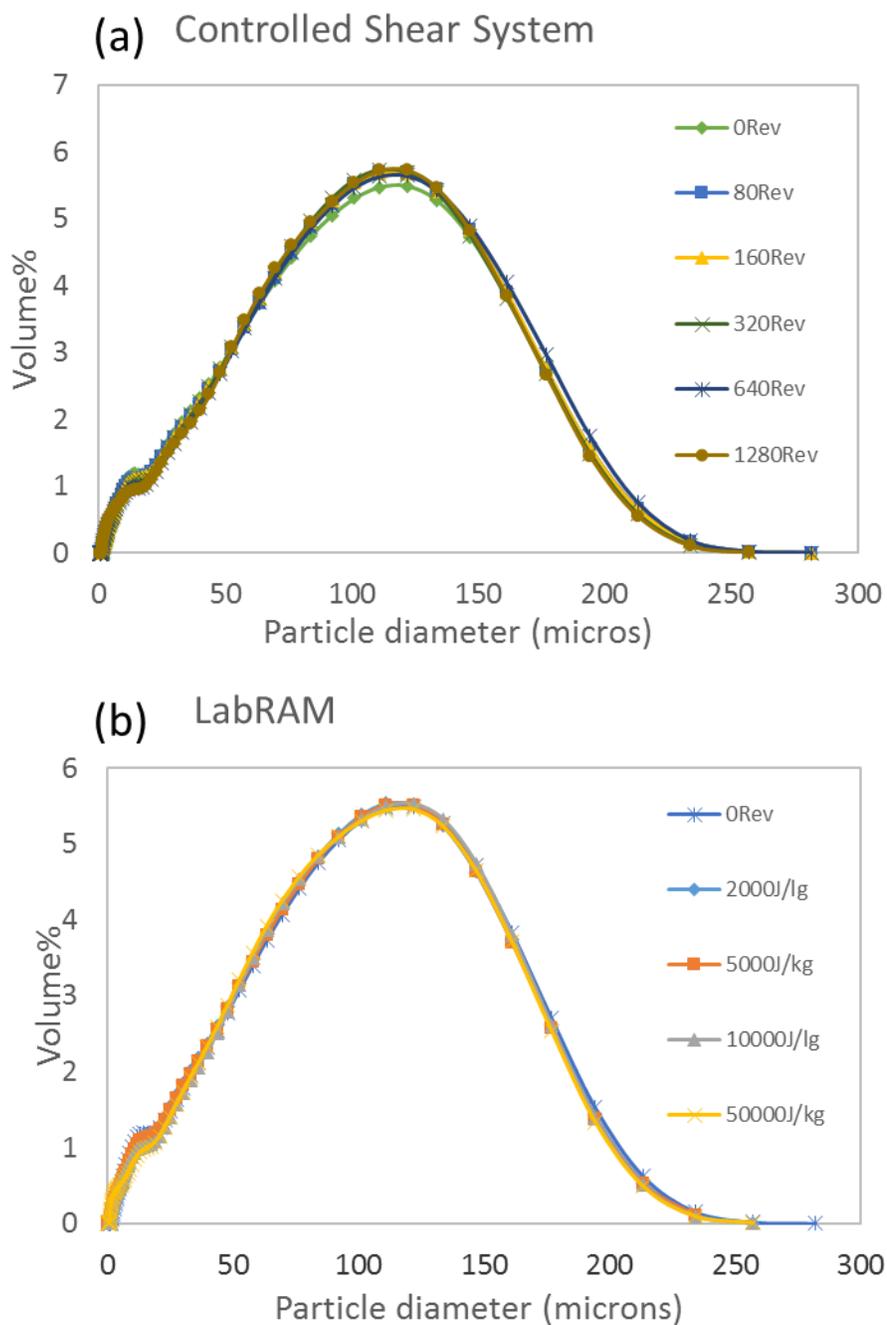
Figure 6-2 Particle size distribution from (a) controlled shear system at 80rpm and (b) LabRAM at 80% Intensity (22.0 Watts).
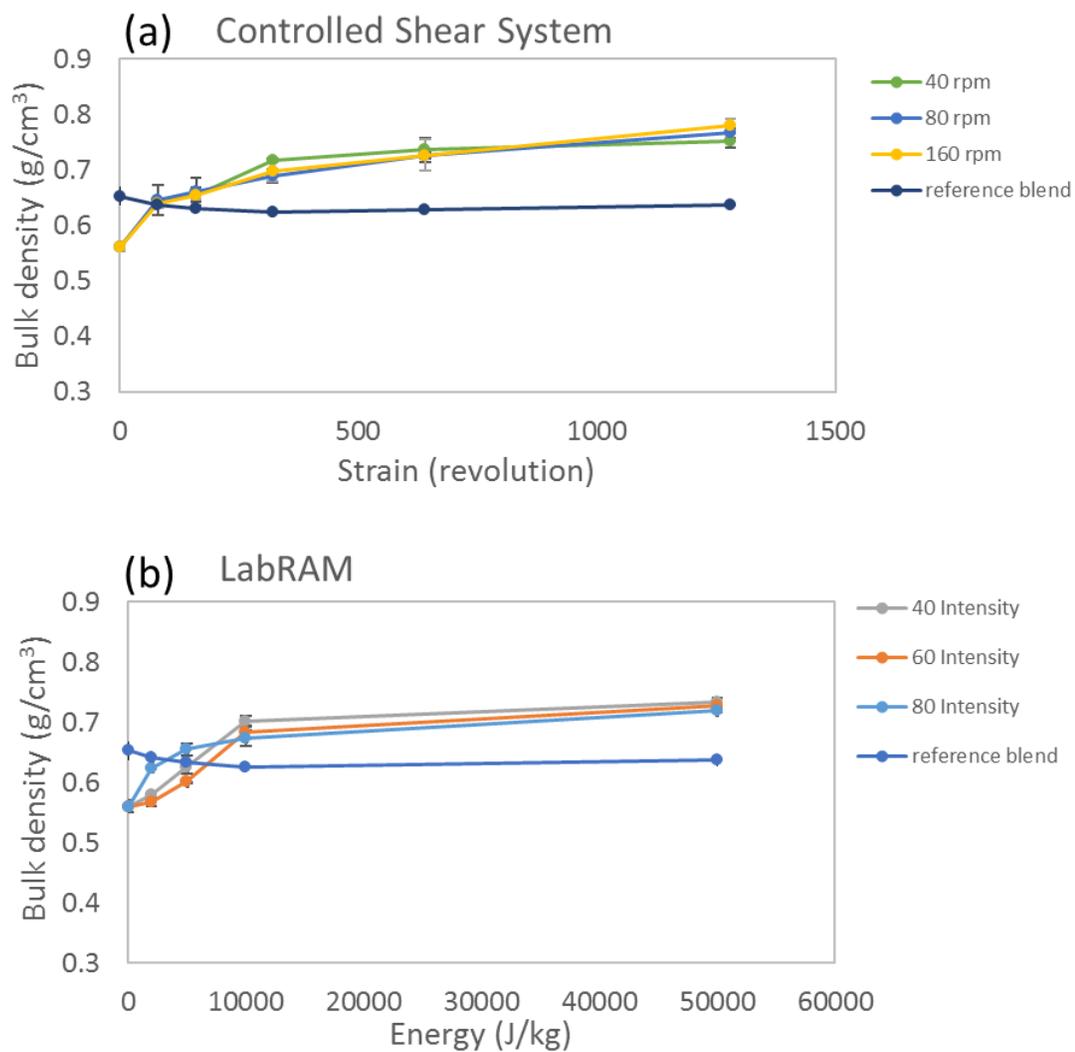
Figure 6-3 Effect of lubrication on blend bulk density in (a) controlled shear system and (b) LabRAM.
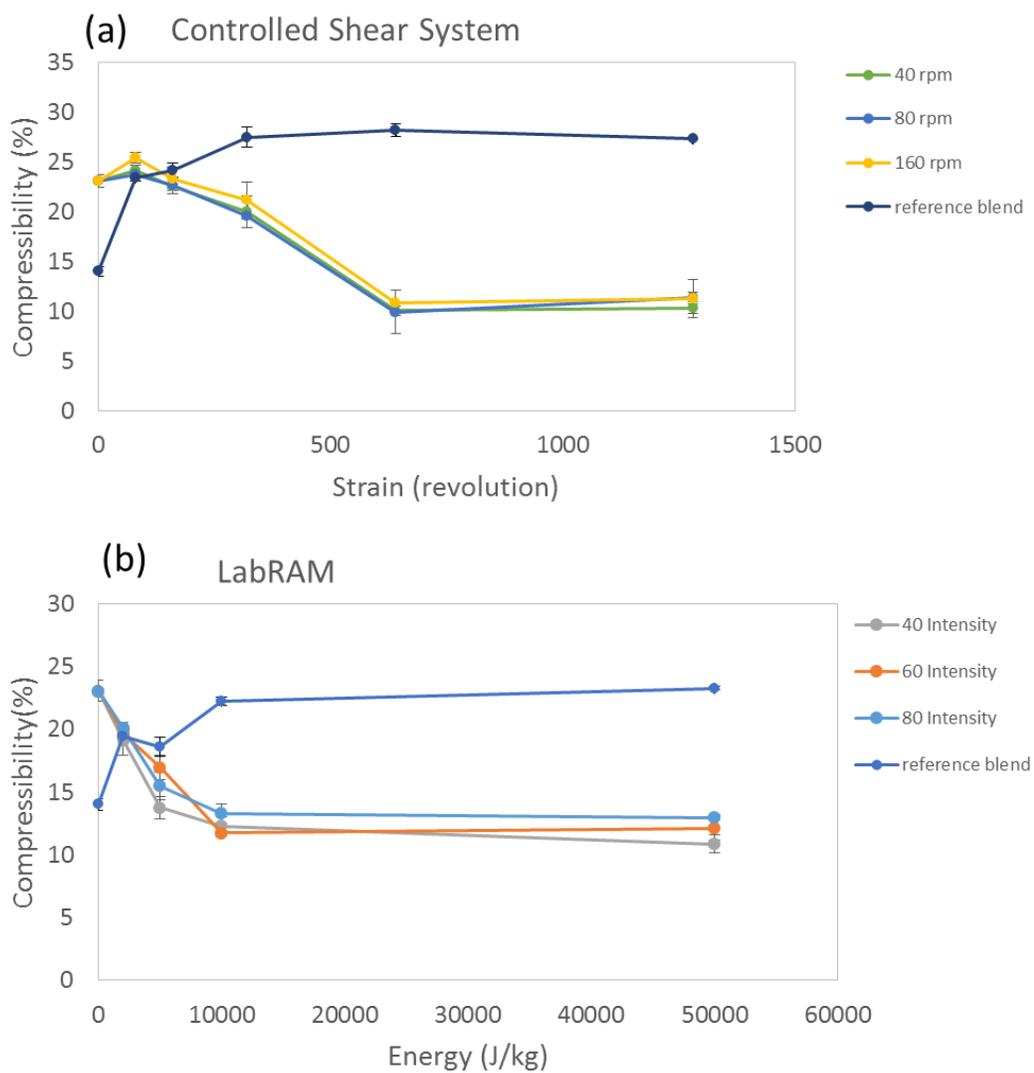
Figure 6-4 Effect of lubrication on blend compressibility in (a) controlled shear system and (b) LabRAM
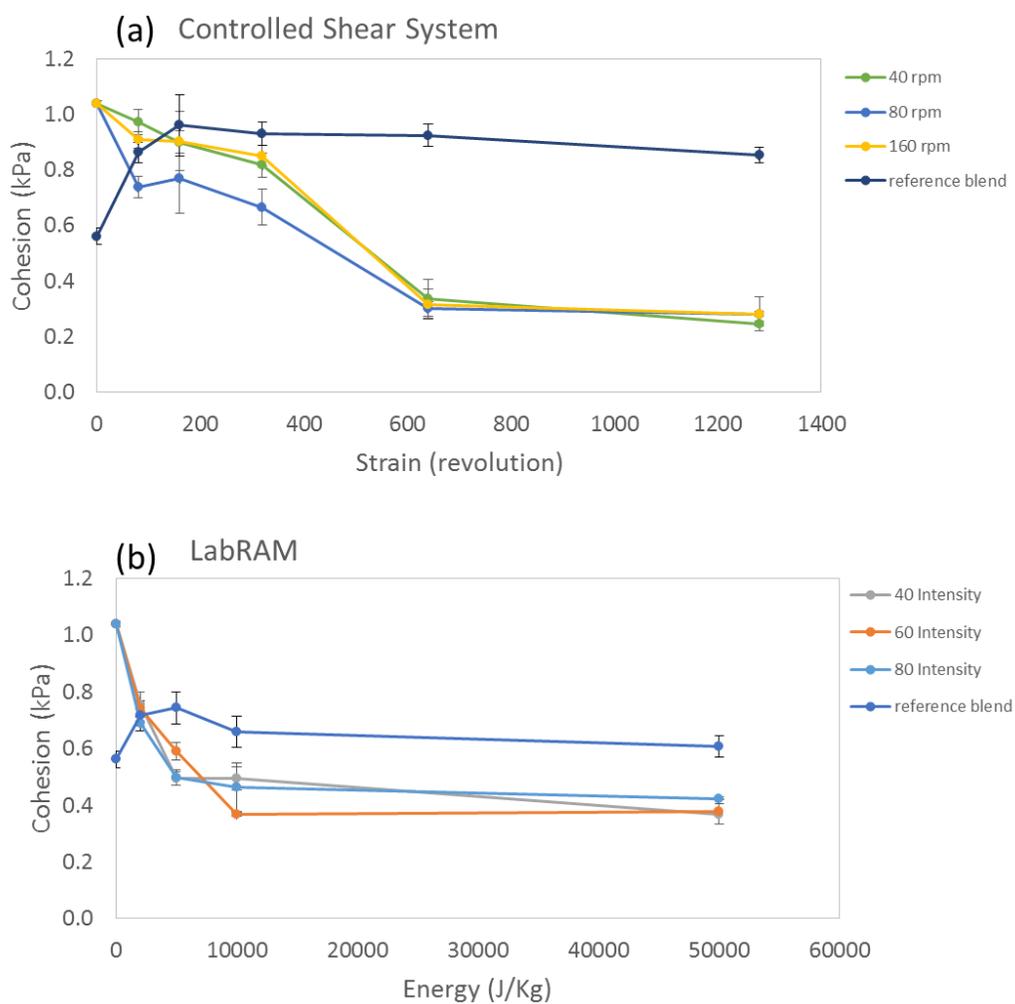
Figure 6-5 Effect of lubrication on blend cohesion in (a) controlled shear system and (b) LabRAM
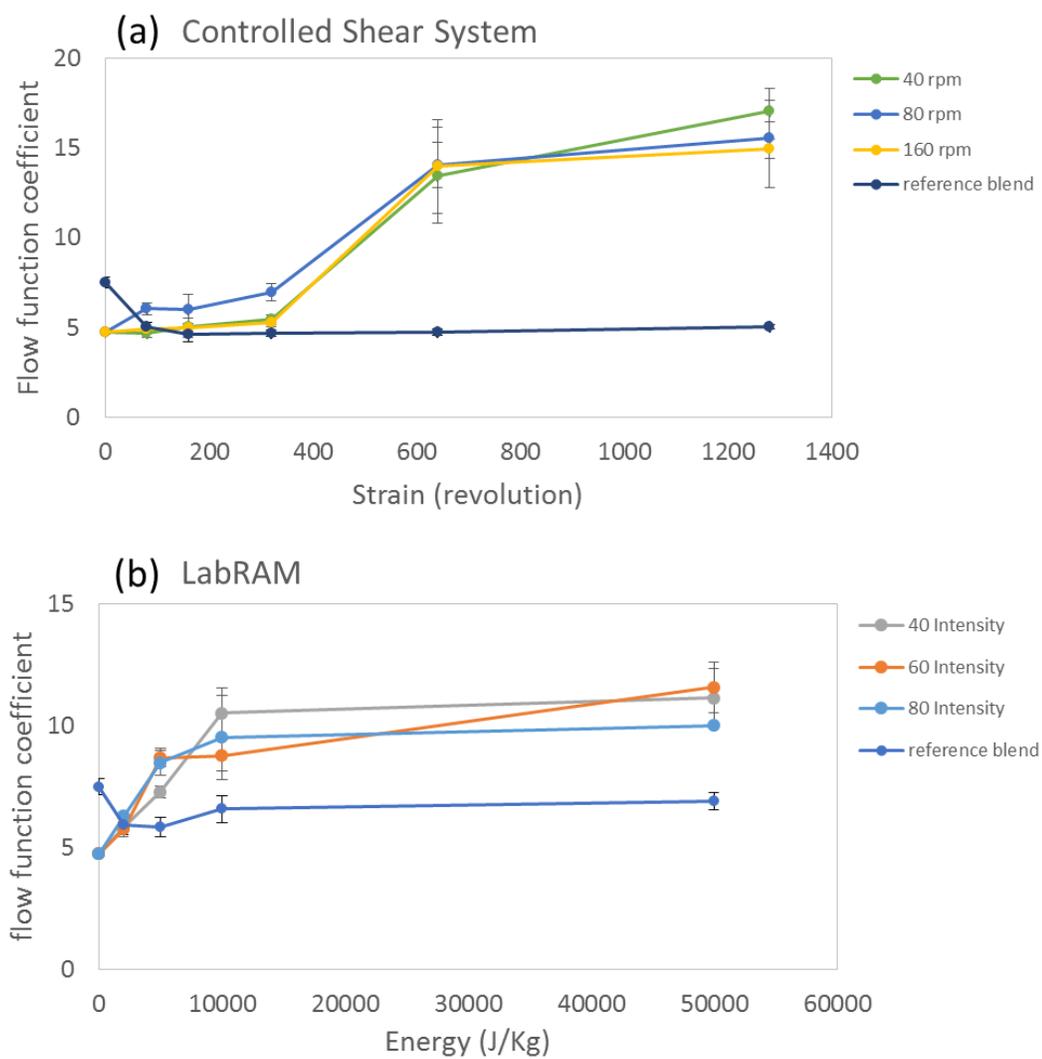
Figure 6-6 Effect of lubrication on blend flow function coefficient in (a) controlled shear system and (b) LabRAM
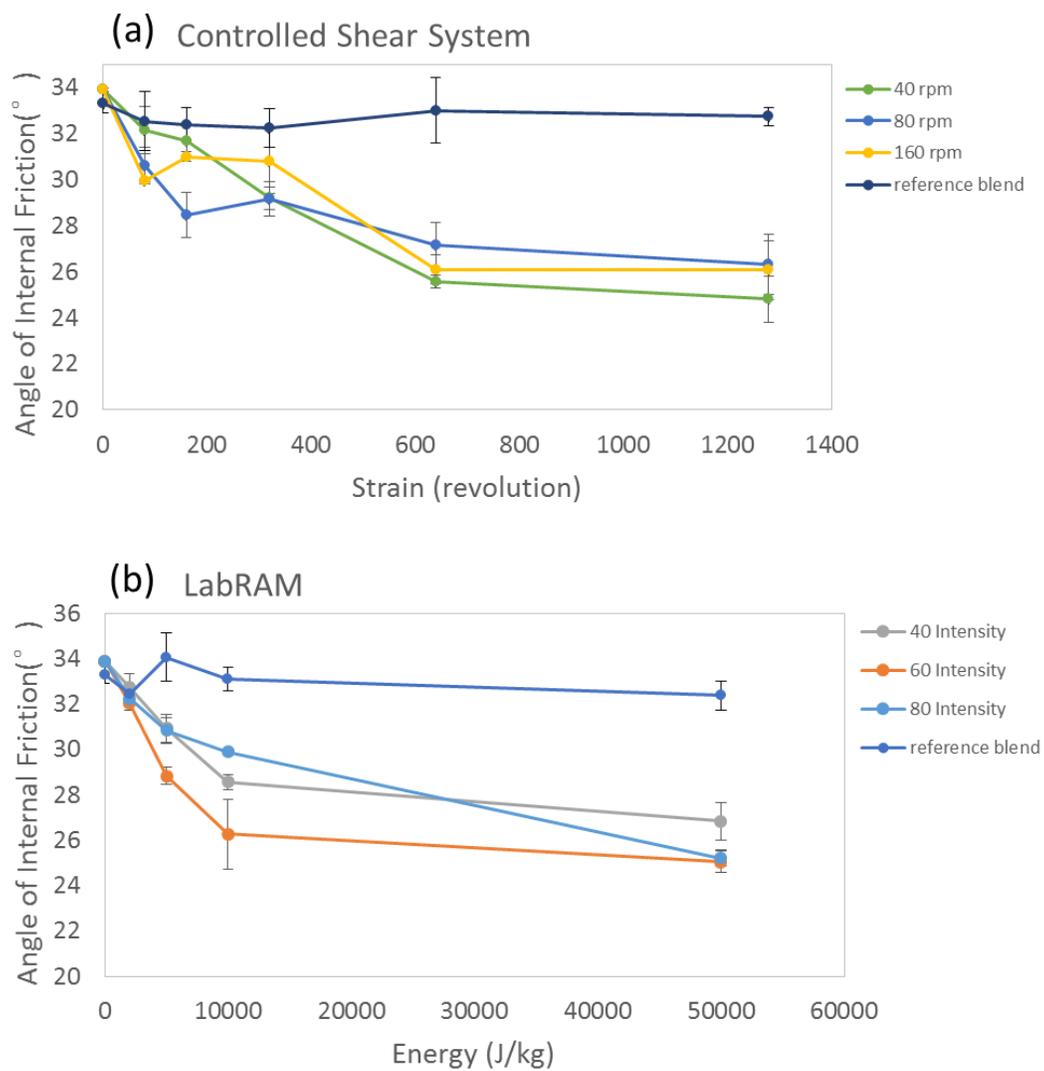
Figure 6-7 The effect of lubrication on blend angle of internal friction in (a) controlled shear system and (b) LabRAM
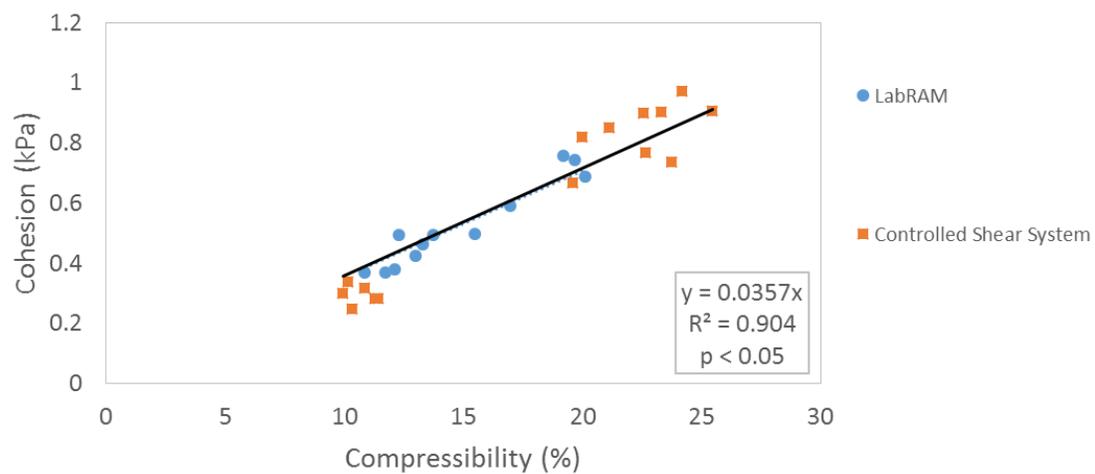
Figure 6-8 The correlation between the cohesion and compressibility.
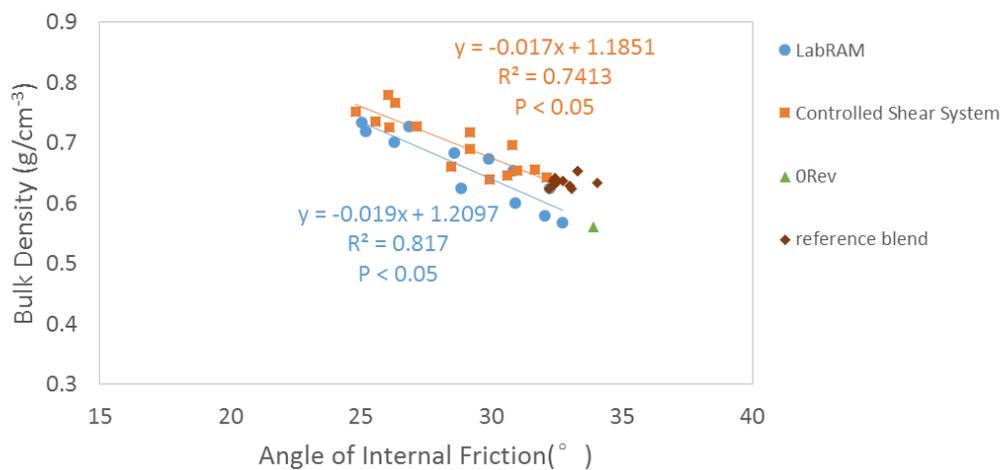
Figure 6-9 The correlation between the bulk density and the angle of internal friction.
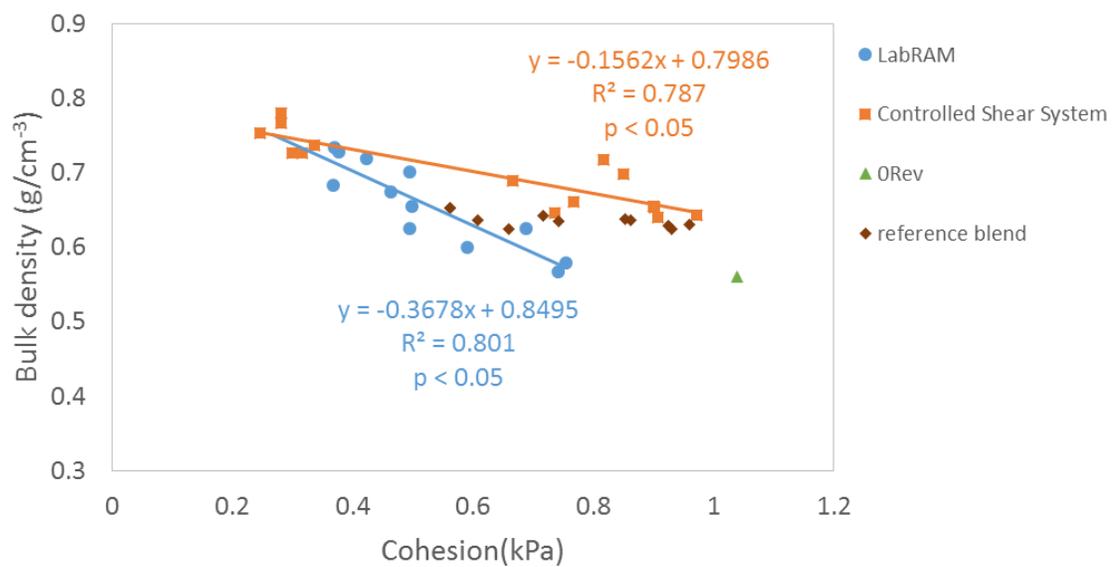
Figure 6-10 Comparison of the correlation between the bulk density and the cohesion of the two devices.

6.6. Tables for Chapter 6

Table 6-1 Particle size distribution of the raw materials used in the study.

Material	Mean (μm)	d10 (μm)	d50 (μm)	d90 (μm)
Lactose	71.9	10.3	63.5	157.7
Semifine Acetaminophen	48.9	5.6	32.6	122.7
MgSt	8.8	2.1	7.8	16.6

Table 6-2 Experimental design for the controlled shear system.

Total Revolutions	40 rpm	80 rpm	160 rpm
80	√	√	√
160	√	√	√
320	√	√	√
640	√	√	√
1280	√	√	√

Table 6-3. Experimental design for the LabRAM.

Energy (J/kg)	40% Intensity (3.7 Watts)	60% Intensity (10.4 Watts)	80% Intensity (22.0 Watts)
2000	√	√	√
5000	√	√	√
10000	√	√	√
50000	√	√	√

Table 6-4 Statistical analysis, using analysis of variance (ANOVA) and omega-square (ω^2), of the effect of total shear (revolutions) and shear rate on the powder flow properties for the controlled shear system.

	Source	DF	Seq. SS	Adj. SS	Adj. MS	F	p	ω^2
Bulk Density (g/cm³)	Shear rate (rpm)	2	0.0001	0.0001	0.0000	0.0800	0.9240	-0.0065
	Shear (revolutions)	4	0.0938	0.0938	0.0234	61.4500	0.0000	0.8522
	Shear rate*Shear	8	0.0026	0.0026	0.0003	0.8500	0.5710	-0.0044
	Error	30	0.0114	0.0114	0.0004			
	Total	44	0.1079					
S = 0.0195347 R-Sq = 89.39% R-Sq(adj) = 84.44%								
Compressibility (CPS%)	Shear rate (rpm)	2	36.9100	36.9100	18.4500	11.6900	0.0000	0.0170
	Shear (revolutions)	4	1885.4600	1885.4600	471.3600	298.7000	0.0000	0.9458
	Shear rate*Shear	8	15.5600	15.5600	1.9500	1.2300	0.3150	0.0015
	Error	30	47.3400	47.3400	1.5800			
	Total	44	1985.2700					
S = 1.25621 R-Sq = 97.62% R-Sq(adj) = 96.50%								
Cohesion (kPa)	Shear rate (rpm)	2	0.1063	0.1063	0.0531	10.1000	0.0000	0.0268
	Shear (revolutions)	4	3.2255	3.2255	0.8064	153.2100	0.0000	0.8960
	Shear rate*Shear	8	0.0816	0.0816	0.0102	1.9400	0.0910	0.0110
	Error	30	0.1579	0.1579	0.0053			
	Total	44	3.5712					
S = 0.0725486 R-Sq = 95.58% R-Sq(adj) = 93.52%								
Flow function coefficient	Shear rate (rpm)	2	6.2130	6.2130	3.1060	1.2200	0.3090	0.0011
	Shear (revolutions)	4	963.5860	963.5860	240.8970	94.6100	0.0000	0.8991
	Shear rate*Shear	8	11.6600	11.6600	1.4570	0.5700	0.7920	-0.0082
	Error	30	76.3850	76.3850	2.5460			
	Total	44	1057.8440					
S = 1.59567 R-Sq = 92.78% R-Sq(adj) = 89.41%								
Angle of Internal Friction (°)	Shear rate (rpm)	2	1.0910	1.4710	0.7360	0.5900	0.5590	-0.0037
	Shear (revolutions)	4	199.8970	196.5670	49.1420	39.6600	0.0000	0.7117
	Shear rate*Shear	8	33.5560	33.5560	4.1950	3.3800	0.0080	0.0878
	Error	30	33.4580	33.4580	1.2390			
	Total	44	268.0030					
S = 1.11319 R-Sq = 87.52% R-Sq(adj) = 81.04%								

Table 6-5 statistical analysis, using analysis of variance (ANOVA) and omega-square (ω^2), of the effect of total energy and lubrication rate (intensity) on the powder flow properties for the LabRAM..

	Source	DF	Seq. SS	Adj. SS	Adj. MS	F	p	ω^2
Bulk Density (g/cm³)	Shear rate (Intensity, %)	2	0.0033	0.0033	0.0017	17.9800	0.0000	0.0276
	Shear (Energy, J/kg)	3	0.0995	0.0995	0.0332	360.9900	0.0000	0.8766
	Shear rate*Shear	6	0.0081	0.0081	0.0013	14.6800	0.0000	0.0666
	Error	24	0.0022	0.0022	0.0001			
	Total	35	0.1131					
S = 0.00958338 R-Sq = 98.05% R-Sq(adj) = 97.16%								
Compressibility (CPS%)	Shear rate (Intensity, %)	2	13.7340	13.7340	6.8670	7.9100	0.0020	0.0311
	Shear (Energy, J/kg)	3	336.7070	336.7070	112.2360	129.3500	0.0000	0.8661
	Shear rate*Shear	6	13.6020	13.6020	2.2670	2.6100	0.0430	0.0218
	Error	24	20.8240	20.8240	0.8680			
	Total	35	384.8670					
S = 0.931481 R-Sq = 94.59% R-Sq(adj) = 92.11%								
Cohesion (kPa)	Shear rate (Intensity, %)	2	0.0007	0.0007	0.0004	0.1700	0.8410	-0.0048
	Shear (Energy, J/kg)	3	0.5989	0.5989	0.1996	97.9300	0.0000	0.8385
	Shear rate*Shear	6	0.0564	0.0564	0.0094	4.6100	0.0030	0.0625
	Error	24	0.0489	0.0489	0.0020			
	Total	35	0.7049					
S = 0.0451481 R-Sq = 93.06% R-Sq(adj) = 89.88%								
Flow function coefficient	Shear rate (Intensity, %)	2	0.1150	0.1150	0.0580	0.0700	0.9350	-0.0102
	Shear (Energy, J/kg)	3	121.8290	121.8290	40.6100	47.7800	0.0000	0.7669
	Shear rate*Shear	6	12.3470	12.3470	2.0580	2.4200	0.0570	0.0466
	Error	24	20.3970	20.3970	0.8500			
	Total	35	154.6880					
S = 0.921876 R-Sq = 86.81% R-Sq(adj) = 80.77%								
Angle of Internal Friction(°)	Shear rate (Intensity, %)	2	21.1050	21.1050	10.5530	17.3100	0.0000	0.0745
	Shear (Energy, J/kg)	3	216.2790	216.2790	72.0930	118.2900	0.0000	0.8034
	Shear rate*Shear	6	14.2990	14.2990	2.3830	3.9100	0.0070	0.0399
	Error	24	14.6270	14.6270	0.6090			
	Total	35	266.3100					
S = 0.780669 R-Sq = 94.51% R-Sq(adj) = 91.99%								

7. Statistical Comparison of Dissolution Profiles

7.1. Introduction

The ability to compare in-vitro dissolution profiles and demonstrate similarity for tablets and other oral products is extremely important to the pharmaceutical industry. Consider three typical situations: (i) Formulation and Optimization decisions: During product development, for products where dissolution performance is a critical quality attribute, both the product formulation and the manufacturing process are optimized based on achieving specific dissolution targets. (ii) Equivalence decisions: During generic product development, and also when implementing post-approval process or formulation changes, similarity of *in vitro* dissolution profiles between the reference product and its generic or modified version are one of the key requirements for regulatory approval decisions. (iii) Product compliance and release decisions: During routine manufacturing, dissolution outcomes are very often one of the criteria used to make product release decisions. [190, 205, 206]

The first two decision scenarios described above are typically based either on pairwise comparisons or multiple comparisons of average dissolution profiles, while the third one usually involves the comparison of an average profile to one or more discrete specifications (i.e, NMT 40% dissolution after one hour, NLT 90% dissolution after eight hours, etc.). However, the data is typically composed of multiple (6 or 12) dissolution profiles for individual tablets. From a theoretical perspective, to compare multiple curves (whether dissolution profiles or anything else), there are three quantitative approaches: ANOVA-based, model-dependent, and model-independent methods. ANOVA-based

methods maintain dissolution data in their original form [207], and test the statistical significance of observed effects by quantifying and comparing sources of variability, often estimated based on experimental design. On the other hand, both model-dependent and model-independent methods describe a dissolution curve by calculated parameters, and thus transform comparison of curves into comparison of parameter values.[208] The difference between the last two approaches is the additional need for selecting a mathematical model and adopting a curve-fitting procedure in model-dependent methods.

While all three approaches are clearly rooted in practice and consistent with current regulations, current industrial practice rarely uses any of them, instead relying on the computation of f_1 and f_2, the so-called similarity and difference indexes. In our opinion, in spite of their popularity, these indexes have dubious meaning, unknown reliability, and also poor sensitivity in a number of important situations.

This chapter focuses on the first of the three decision scenarios described above, the "optimization" scenario, characterized by the need to compare dissolution profiles obtained for multiple experimental conditions. We introduce several alternative methods for comparing dissolution profiles that not only are significantly more rigorous and meaningful than f2, but also enable the user to assess the reliability of results using standard statistical metrics, at the expense of a very small increase in calculation complexity. The other two decision scenarios, i.e., the "equivalence between two conditions", and the "compliance with specification" conditions, will be addressed in future publications.

However, before proceeding with a further discussion of statistical method significance, it is useful to consider the different meanings of the term "significance" in the context of pharmaceutical product development.

- *Statistical Significance,* as a means of assessing the reliability of a scientific result, is a measurement of the probability that the observed result actually exists, (or complementary, that an effect believed to be absent is indeed not present). Statistical significance, and thus, statistical reliability, depends on four factors – how much data has been collected, how the experimental space has been sampled, how much variability is present in the data, and how large are measurement errors relative to the magnitude of the effect.

- *Clinical Significance*, in the context of pharmaceutical development, product equivalence, product release, and quality control, is related to a completely different question: whether the effect is large enough to have a meaningful impact on a patient's health. Critically, Statistical Significance must not be confused with Clinical Significance: given enough data, even very small effects can be quantified in a statistically significant manner, but that does not make them clinically relevant.

- Finally, *Regulatory Significance,* ideally, should be the answer to yet another question – whether the observed effects, known with a given degree of reliability, are important enough to warrant regulatory action (or inaction)

The proposed methodology aims to improve all three types of significance – to increase statistical significance of dissolution profile comparison, so that effects can be measured

more accurately and reliably, so that clinical significance can be ascertained with increased confidence, and regulatory decisions can be made more objectively.

7.1.1. Current approaches for dissolution profile comparison

The most common current practice to compare dissolution profiles is to use the "similarity factor" f_2, which is a model-independent approach to define dissolution similarity based on the mean-squared difference between a pair of profiles [209], usually estimated as the average of twelve tablets (although comparisons based on the average of six tablets, or even as little as three tablets, are not uncommon in product development). This method is one of the approaches described in the latest FDA's Guidance for Industry regarding dissolution testing [210-212]. In the f_2 protocol, the average dissolution values of test and reference profiles are used. According to the guidance, an f_2 value greater than 50 is required for curves to be considered similar, and roughly corresponds to differences between profiles being smaller than about 10%. Another model-independent index derived from average absolute differences between individual time points of the average profiles is termed as the "difference factor" f_1, with value less than 15 presumably indicating similarity. Other methods mentioned in the current FDA Guidance for immediate release dosage forms include a multivariate confidence region procedure and a model-dependent method with no more than three parameters, where both similarity decisions are based on whether the "confidence region" is within the limits of a so-defined "similarity region".[210-212]

7.1.2. Disadvantages of the f_2 method for pharmaceutical product development

Because of its simplicity and its adoption by various regulatory agencies, the f_2 method has gained considerable popularity and is widely used to guide similarity decisions.[213-215] However, this method has several easily noted shortcomings. First, f_2 is limited to pairwise comparison. When there are N groups of dissolution profiles, N*(N-1)/2 f_2 values need to be calculated. Moreover, f_2 is generally used to compare two dissolution profiles such as that of reference and test batches, pre-change batch and post-change batch. However, in cases when similarity is investigated among groups of dissolution profiles without an obvious reference, the f_2 matrix cannot provide a direct interpretation.

In addition, and perhaps much more importantly, the similarity factor does not consider the degree of variability among individual measurements. Although there is a recommendation in the guidance for the coefficient of the variability (RSD) between individual values at specific time points to be less than 10% (except at early time points, where the recommendation is that the RSD should not exceed 20%), the use of the average alone fails to consider the measurement variation and sampling error.[216] When variability exceeds the rule from the guidance, f_2 cannot be applied and other statistical models should be considered.[217]

Also importantly, the outcomes of the f_2 test can depend on the number and spacing of time points selected to perform the test, allowing for another source of ambiguity.[218] FDA addresses this issue by suggesting using no more than one time point corresponding to dissolution above 85%, but as seen below, this recommendation by itself does not resolve the problem, and does not even apply to situations where the relevant and important comparisons take place at early stages of the dissolution process.

Finally, the lower similarity limit of 50 in f_2 method is calculated based on an average 10% difference at all sampling times. However, it is standard statistical practice in statistical data analysis to use "within group" variance components to evaluate the reliability of any measurement. It is also very common to compare "within group" to "between groups" variance components, informed by the size of the sample, to ascertain the statistical significance of the observed effect (in fact, this latest comparison is the root concept of all ANOVA methods). Remarkably (and not in a positive way), the f_2 method ignores the sample size (other than a recommendation by FDA to use at least 12 tablets, which as mentioned, is often not followed), and wastes the known within-group variability information, providing neither an estimate of the reproducibility of the measurement nor an estimate of its significance.[219]

The lack of statistical justification of the similarity factor has been reported in the literature. Liu *et al.* used simulated dissolution profiles to generate distribution statistics of f_2 and suggested that the similarity factor is too liberal in concluding similarity between dissolution profiles.[220] As observed by Hsu *et al.*, the f_2 formula simply reflects the overall average difference and neglects the difference of dissolution pattern between the test and reference.[221] Shah *et al.* simulated the confidence interval of f_2 with a Bootstrap method, and found that the commonly used similarity factor formula can generate biased and conservative estimates.[222] Further investigation by Ma *et al.* showed that in order to produce the acceptable 50-100 range, the distribution of f_2 would be unnecessarily complicated.[223] Later simulation work from Ocana *et al.* confirmed the insufficiency of statistical justification for 50-value threshold.[224] The similarity factor may be indeed easy to calculate and be intuitively appealing for empirical

comparison of two dissolution profiles, but it is not a good statistical estimator of similarity.

7.1.3.   Alternatives to f_2 index in the previous literature

The limitations of f_2 have inspired researchers to propose other methods to conduct pairwise dissolution comparisons. For example, the concept of similarity factor S_d was introduced in combination of f_2 to evaluate selected variables on dissolution in a factorial experiment.[225, 226] Seo *et al.* used a four metrics along with f_2 to compare dissolution profiles.[227] Saranadasa *et al.* proposed a multivariate test to compare two dissolution profiles with the assumption of multivariate normality of the data.[228] Interestingly, Maggio *et al.* proposed a method based on principal component analysis with the establishment of a confidence region (PCA-CR) after an outlier detection step using Hotelling's test to avoid high variability in the data.[229, 230] The advantage of PCA is its ability to examine multivariate data and reveal hidden trends in fewer dimensions, thus simplifying interpretation of results.[231] For dissolution profiles with missing elements, Adams *et al.* used expectation-maximization algorithm in combination with PCA, and constructed confidence limits by bootstrap technique.[232, 233] Comparison methods based on nonparametric Permutation Test theory were also described as viable alternatives.[234] Another method based on cluster analysis from D. Enachescu demonstrated the capability to visualize data and reflect variability within the individual dissolution curves.[235] In addition, the use of time series models,[236] fractional dissolution rate functions,[237] and Lagrange multipliers[238] have also appeared in the literature to assist pairwise comparison.

As diverse as these methods are, they do not address several critical features of multi-group testing. For example, pairwise comparison is of limited value when one needs to investigate several factors with multiple levels (which is the typical situation when optimizing a formulation or a process). In such a case, one needs to examine not only be the similarity between dissolution profiles, but also the relative effects of multiple factors and their interactions, usually from designed experiment. Moreover, it is useful to ascertain whether the observed effects are independent of sample size, which usually requires an estimation of effect size (besides the typical calculation of p-value). Importantly, to maximize usefulness, a method should be as easy to apply as possible.

To summarize, the shortcomings of the f_2 method are inconsistent with the enormous reliance currently placed on dissolution testing to make quality and regulatory decisions, particularly in situations where dissolution is a critical product attribute and patient risk and manufacturer failure risk are high. Industry and regulatory agencies need a flexible and reliable analysis framework that can be adapted to the different types of comparisons required by the multiple scenarios that use dissolution results to make decisions.

The purpose of this chapter is to introduce such a framework. To that end, we introduce two methodologies for the reliable and meaningful comparison of multiple groups of dissolution profiles: (i) MANOVA with repeated measures and (ii) modified level-shape PCA analysis. In Section 3, we use a designed case study to compare the introduced methods with results obtained using f_2. In performing this comparison, we propose the use of the omega-squared statistic to quantify effect size and to promote the application of a test with increased statistical significance as a tool to ascertain clinical significance.

In the final section, we provide a step-by-step illustration of the statistical procedures through a demonstrative case study.

7.2. Methods

Methods of effect size test, multivariate ANOVA (MANOVA) and modified-PCA were detailed in Chapter 2. Importantly, both the modified-PCA and the MANOVA repeated measures methods examine level and shape independently, and their conclusions can be readily compared. The treatment effects in MANOVA repeated measures correspond to the level effects in modified-PCA, and Time*Treatment interaction effects in MANOVA are in fact shape effects in modified-PCA. Both methods can be used to perform both pairwise and multiple comparisons, and both can examine multiple effects and their interactions. Moreover, for both methods, as seen below, reliability of measurements, statistical significance of results, and effect size can be readily determined. In our opinion, neither method is uniformly superior.

Dissolution data from a case study is used here for the purpose of illustration. The formulations were designed with an enteric coating, which was supposed to prevent the release of the drug under acidic conditions. The dissolution test measured the amount of drug released under different conditions. Time points of interest extended up to 240 min (**Figure 7-1**). The case study aimed to study the effect of drug strength (level 1 and level 2), tablet stability time (level 1- level 4), and dissolution testing condition (level 1 and level 2) on release profiles by a 2*4*2 full factorial design. In total 16 conditions were generated, and for each condition, 6 tablets were sampled for dissolution test. Further technical details, which are unavailable due to confidentiality requirements of the product owner, are not needed to gain the benefit of the example. Importantly, this example is

also relevant to other situations where only small amounts of drug are supposed to be released, such as abuse resistant formulations, dose dumping prevention, etc.

7.3. Results and discussion

Since there are 16 conditions tested (6 tablets each), 16 averaged profiles were calculated. The 16 profiles, shown in **Fig 7-1**, are all found to be "similar" ($f_2 > 50$) when tested using the similarity factor method. In other words, all (15*14/2) pairwise interactions gave a value of ($f_2 > 50$).

However, visual examination readily indicates that the profiles are not equivalent. The differences between profiles, and moreover, the effects of the independent variables, can be readily detected by MANOVA and modified PCA. The information with nominal variables described in **Table 7-1** is sufficient to implement a full analysis to find target information: 1) Do the 16 groups have statistically significant differences? 2) Does all three variables have a statistically significant effect on dissolution? Do the variables have statistically significant interactions? What is the relative magnitude of each effect and each interaction? 3) Which groups are (statistically) significantly different among the 16? Which groups can be considered to be similar to each other?

7.3.1.   Test effect of treatment by MANOVA repeated measures

To test the treatment effects, commercial software JMP 10 (SAS Institute Inc.) was used for MANOVA repeated measures analysis. Full results of the F test are listed in **Table 7-2**. This analysis immediately shows that strength, stability time, and testing condition all have significant ($p < 0.05$) effects on dissolution profiles. In other words, the observed differences between the 16 groups of responses under investigation are statistically

significant. Also, one of the two-way interactions (strength*stability time) and the three-way interaction were statistically significant. In other words, all three main effects, one of the two-way interactions, and the three-way interaction all had statistically significant effect in the total amount of drug released.

The analysis can then be shifted to the "within subject" effects, where the effect of time is of course statistically significant, and where the presence of significant interactions of the time factor with other factors indicates that differences in profile shape are also statistically significant. Interactions with time imply that profiles are not parallel in shape.[239]

7.3.2.   Modified-PCA for level and shape analysis

Before M-PCA is performed, the level for each profile can firstly be calculated according to equation (1). The formula averages the percentage released over 240 min and directly reflects an overall performance (total amount released) without considering the time trend (i.e., the shape).

In order to calculate shape score, average profiles are calculated for each group. The simplest procedure to organize average profiles is as follows: the data is arranged in a spreadsheet as in **Table 7-3**, followed by calculation of row mean ($y_{i.}$) and column mean ($y_{.j}$) respectively, as well as grand mean ($y_{..}$).

The residual matrix is then constructed from the average profiles. The element in residual matrix can be calculated according to equation (2). The constructed residual matrix is shown in **Table 7-4** as an example. Residuals are subtracted from the mean effect, and thus are representative of differences in shape. The advantage of using average profiles

for modified principal component analysis is to better capture the "between group" variability with minimization of experimental variations. The "within group" variability is still kept in the residual matrix calculated from individual profiles as shown in subsequent step.

Principal component analysis is then performed on the residual matrix (M-PCA), instead of original dissolution data, based on average profiles. Special attention needs to be paid to the eigenvectors (**e**) that explain most of the variance in shape. As shown in **Table 7-5**, the first four components account for more than 99.5% of the variance. Further observation shows the first component claims 94%, and thus can be used alone to represent shape. Column vector (**e**) is shown in **Table 7-6**.

Finally, a residual matrix based on individual profiles is constructed. The matrix is composed of 96 row vectors (y_i) and each vector records individual release percentage at 10 time points. Based on eigenvectors from M-PCA on average profiles, the shape score of individual profiles can be calculated from the residual matrix of individual profiles. Once this is done, a shape score is calculated for each individual profile. The shape score is the dot product $\mathbf{e} \cdot \mathbf{y_i}$. Since only one vector is sufficient in this case study to describe shape, only one score needs to be calculated for each profile In other words, the calculated value is the score of individual profile after an orthogonal transformation of residual matrix. (N.B. – more complex cases might require multiple eigenvectors, and therefore, multiple scores per profile).

Once the level and shape factors are calculated, ANOVA methods can be used to test the effect of treatment on level and shape. By now, level and shape factors have been

calculated separately for each individual profile, and comparing dissolution profiles is transformed into comparing the two descriptive values (level value and shape score). **Table 7-7** and **Table 7-8** show ANOVA results for level and shape. Since tablets and associated profiles represent sampling error, the degrees of freedom (DoF) of the error term come from pooling the two and three factor interaction to produce a valid test of significance. Pooling does not produce additional DoF, and the total DoF is still 15. High R-squared values confirm that the reduced model explains most variability. Minitab 16 (Six Sigma Academy Module, Minitab Inc.) was used.

 ANOVA shows that product strength and testing condition both had significant ($p<0.05$) effects on both level and shape of dissolution profiles. A further graphical presentation of the main effect suggests the trend and magnitude of these effects. For example, in **Figure 7-2**, change from strength 1 to strength 2 will decrease the level of dissolution profiles (suggesting a formulation effect on the release rate), while change from testing condition 1 to testing condition 2 will increase the level (indicating a sensitivity with respect to testing conditions). On the other hand, in shape analysis, change from strength 1 to strength 2 will decrease the shape score while change of testing condition will result in an increase of the shape score (**Figure 7-3**).

 The values of the shape score are less direct for interpretation. For illustration purposes, two profiles with the shape scores of 2 and -2 are selected from the 96 profiles to provide a visual comparison. From **Figure 7-4**, it is clear that the shapes are quite different. The release percentage of the profile with shape score of 2 after subtraction of level effect has a wider range than the profile with score of -2. Profile shape is frequently affected by the drug release mechanism, and therefore, ability to isolate profile shape observations are

useful to understand how process and formulation parameters might affect the drug release mechanisms of the product.

Moreover, it is also clear that dissolution profiles with higher values of the level score also has higher values of the shape score, meaning, reasonably, that the same factors that enhance drug release affect the shape of the profile. **Figure 7-5** demonstrates a general linear relationship between level and shape score from all 96 dissolution profiles. Such an effect can occur, for example, in situations where the product total weight is kept constant as the product strength is increased, by decreasing the concentration of a second ingredient. In such a situation, the change in drug content can also lead to a change in the mechanism controlling drug release.

3.1 Effect size test

As mentioned, the Omega squared index is an unbiased estimate of the proportion of variance in the population that is explained by a given treatment (or by an interaction). The index can be calculated using equation (3). In general, large effect display ω^2 values above 0.14.[240] Recall that two-way and three-way interaction effects have been pooled to estimate experimental error. Table 9 shows the effect size of main effects from ANOVA table. Strength and testing condition can be categorized as large effects both for level and shape, suggesting that the significance reported in ANOVA table is independent of sample size, and that the statistical significance from p-value is also indicating practical significance in terms of effect size. It is also clear that the strength and testing condition factors are much larger than the stability time factor, indicating that these two parameters contribute much more to both the amount of drug released, and the time pattern of drug release, than the stability time.

 Effect size indexes, while very useful to ascertain relative importance of multiple statistically significant factors, are generally normalized by the total system variability, and therefore are not a stand-alone indication of clinical or regulatory significance. An extrinsic standard of maximum acceptable variability (e.g., "difference between profiles should not exceed 15%"), ideally determined based on knowledge of therapeutic characteristics of the product, is needed for the determination of clinical or regulatory significance.

3.4 Post-hoc analysis

Tukey's test was applied for post hoc analysis of the level and shape factors using Minitab 16. The method first calculates group means, and then compares group means with 0.95 confidence interval. The null hypothesis is that pairs of groups under investigation are from the same population. The test statistic in Tukey's method is a generalization of the t-test in order to avoid increasing type I error for multiple comparisons:

$$q_s = |Y_1 - Y_2| / SE \quad (4)$$

Y_1, Y_2 represent two group means, and SE is standard error of the data. If q_s is larger than the $q_{critical}$, the null hypothesis is rejected and two groups are concluded to be significantly different.[241]

Table 7-10 and **Table 7-11** have classified similar groups with the same letter based on values of level and shape factors, respectively. Groups that do not share a letter suggest

significant differences. A combination of the two gives 3 statistically similar groups: Group 1, 3, 5, 7, 10, 12, 14; Group 4, 6, 8; and Group 11, 13, 15 (**Figure 7-6**).

7.4. Conclusions

One of the most important product quality attributes, dissolution performance, is currently analyzed using a non-rigorous, and often inadequate method introduced 17 years ago. While commonly used similarity factor f_2 method provides a simple method, based on calculation of a single value, to perform pairwise profile comparisons, other approaches are better suited for many situations of interest.

In this chapter, rigorous yet easy to use methodologies are introduced for the determination of statistical significance of observed multivariate effects in dissolution testing. MANOVA with repeated measures and modified PCA methods in combination offer detailed and valuable information. A step-by-step procedure including effect size test and Tukey's method was demonstrated using commercially available software packages and actual dissolution data. The statistical assumptions are the usual assumptions of statistical tests of significance, which is normally distributed experimental variation with homogeneous variance.

The main point in this chapter is that analysis of dissolution profiles is dependent on the target information and needs to consider the applicable decision scenario and also the intrinsic data structure. Critically, if the purpose is to optimize a product or a process, the method used to analyze the data should consider the multivariate nature of the information, the self-correlated intrinsic nature of dissolution profiles, and the availability of within-group variability as a means for determining reliability and significance.

Confronted with such a situation, the pharmaceutical scientist can first run MANOVA with repeated measures and obtain a quick assessment of the significance of between and within subject effects. This analysis, combined with effect size calculations, allows the skilled artisan to quickly establish the significance and reproducibility of the dissolution method *as a whole*, and to discriminate factors that are important from those that are not.

Importantly, since Omega-squared tests the magnitude of effects relative to the total variability in the data set, in order to make a final assessment regarding the practical (or clinical) significance of observed effects, an extrinsic standard of total acceptable variation is needed. Such a factor is, necessarily, product dependent, and should be different for drugs with narrower or wider therapeutic indexes, toxicity, and side effect issues, etc.

Subsequent to MANOVA, modified PCA analysis enables to dissect dissolution profiles to separate "level" (how fast is the product dissolving) from shape (are drug release mechanisms changing?). Interestingly, a strong correlation between level and shape was found in the case study discussed here, indicating that sometimes level and shape are both affected by the same factors, i.e., they are covariate.

Using M-PCA, curve comparison becomes a straightforward comparison of a small number of representative, model-independent values, the level and shape scores. Although specific calculations vary, the level in modified PCA method corresponds loosely to between-subject effect in MANOVA repeated measures, and shape is interpreted as within-subject effect from interactions with time. PCA makes the analysis more straightforward, at the expense of a small additional computational effort.

A final post-hoc analysis using Tukey's method identified groups of similar release behavior. The similarity from Tukey's method is more sensitive to profile differences than the threshold of 50 in f_2.

To conclude, there is a difference between statistical significance and clinical/regulatory significance. Statistical significance implies that observed difference is larger than random variation. Commonly reported p-value measures statistical significance, which is dependent not only on the size of the effect, but on the amount and noisiness of the data. Clinical significance, on the other hand, is determined by whether the observed differences are likely to have an effect on a patient. Statistical significance must be present before it is meaningful to assess practical significance, but given enough data, some statistically significant difference may be of no practical importance. A clear illustration of this issue is the fact that, for a fixed amount of data, as measurement error decreases, statistical significance increases, without changing the true size of effects or their putative clinical/regulatory significance.

However, this does not mean that we should continue using a less robust method to mask uncertainty. On the contrary, only through the efforts of providing better methodology and narrowing the gap between methodology and implementation can we continue to move towards scientifically grounded regulatory decisions. It is hoped that the proposed statistically based methods in this chapter with ease of practice could facilitate understanding the nature of dissolution curves and more appropriate ways to analyze them. To further demonstrate the application of the proposed methodology, Chapter 8 connects M-PCA to chemometrics and multi-linear regression, to enable nondestructive prediction of tablet dissolution profiles.

7.5. Figures for Chapter 7
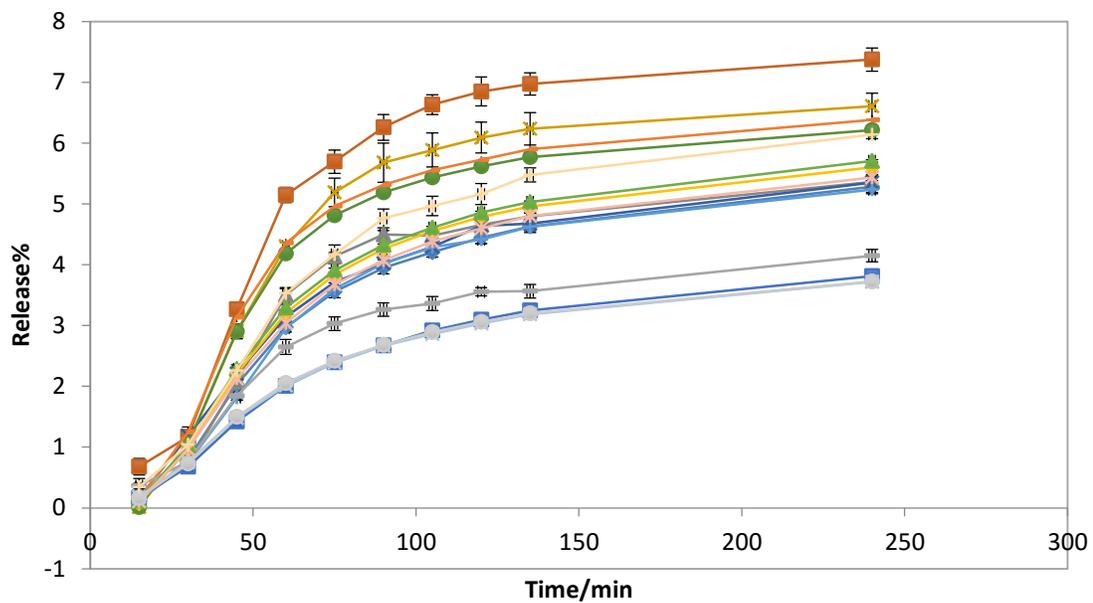
Figure 7-1 Dissolution profiles from 16 conditions. Each is an average of 6 profiles with error bars shown.
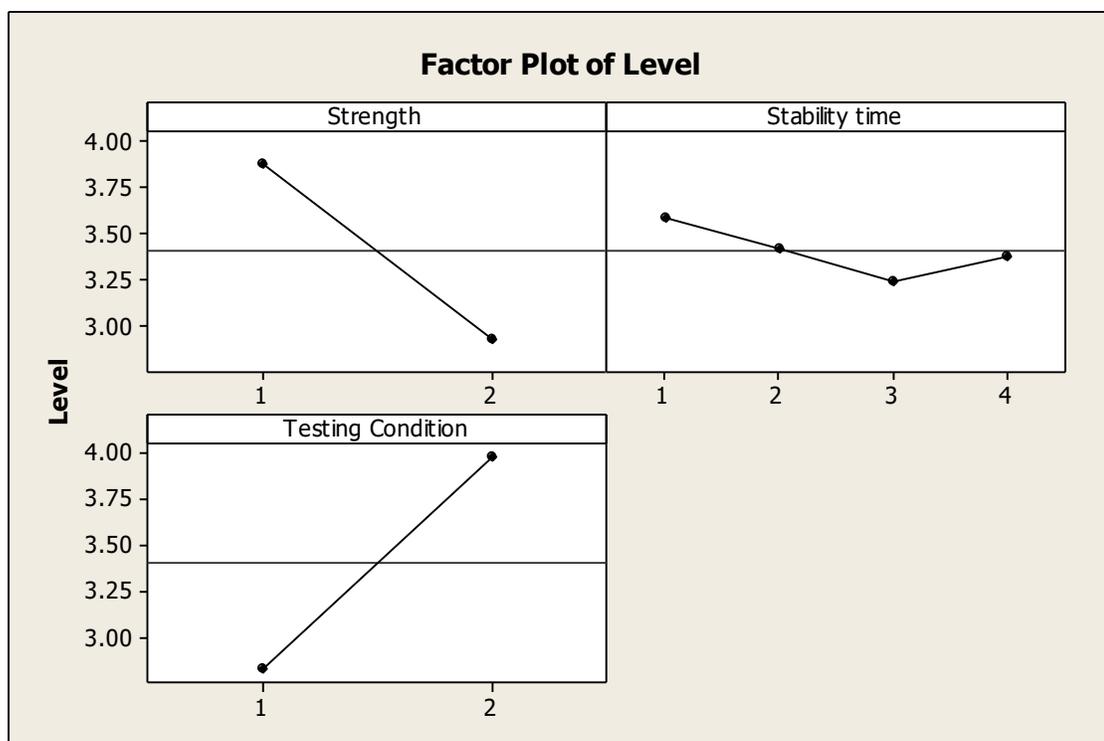
Figure 7-2 Main effect plot of level
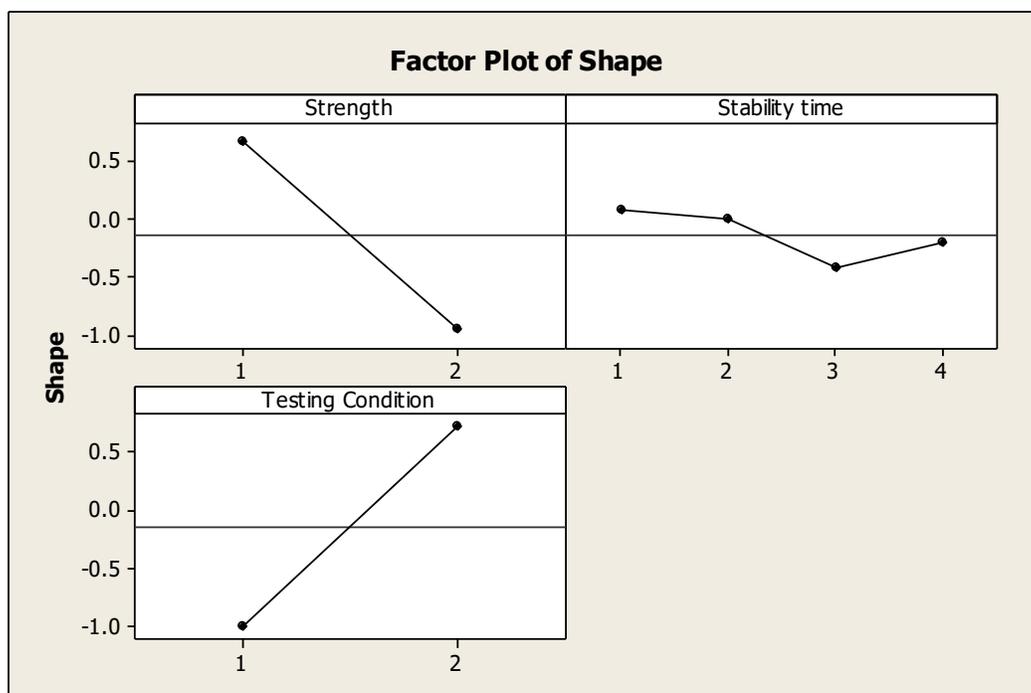
Figure 7-3 Main effect plot of level
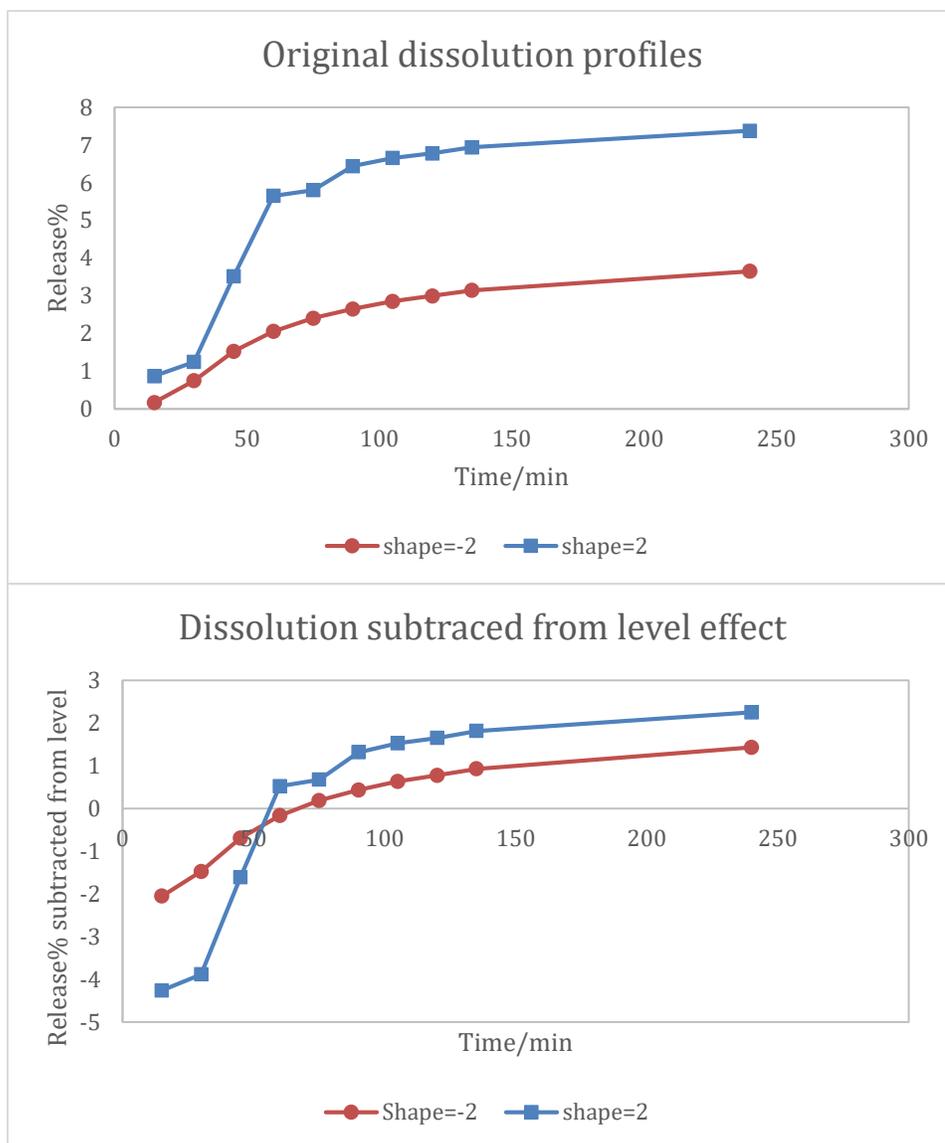
Figure 7-4 Dissolution profiles from shape value of -2 and value of 2.
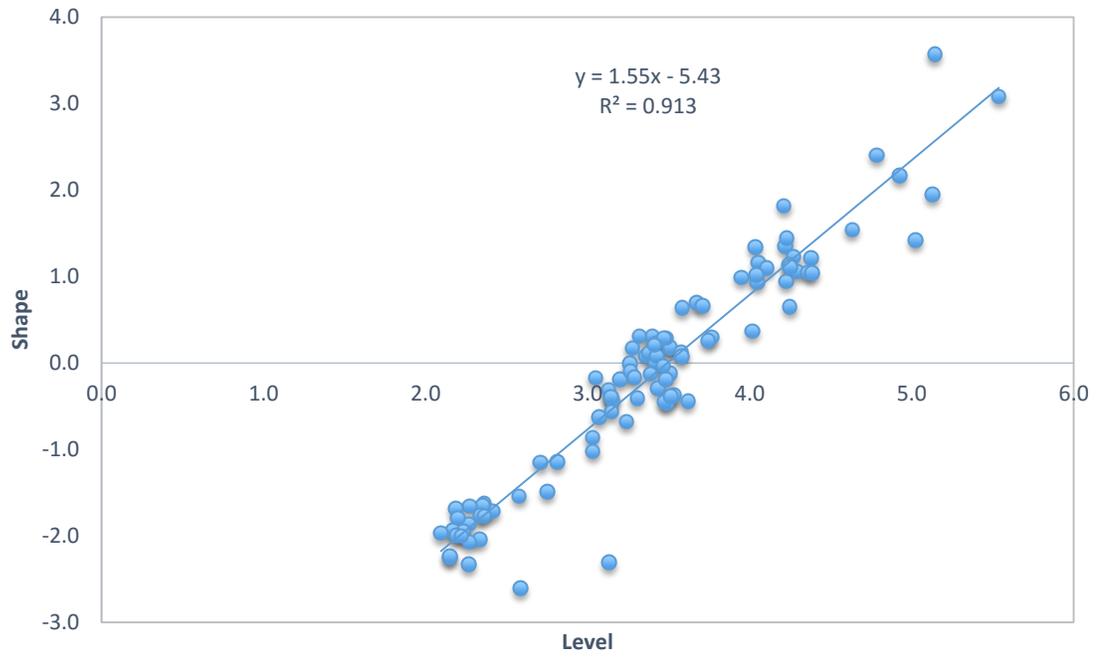
Figure 7-5 Correlation between level and shape.
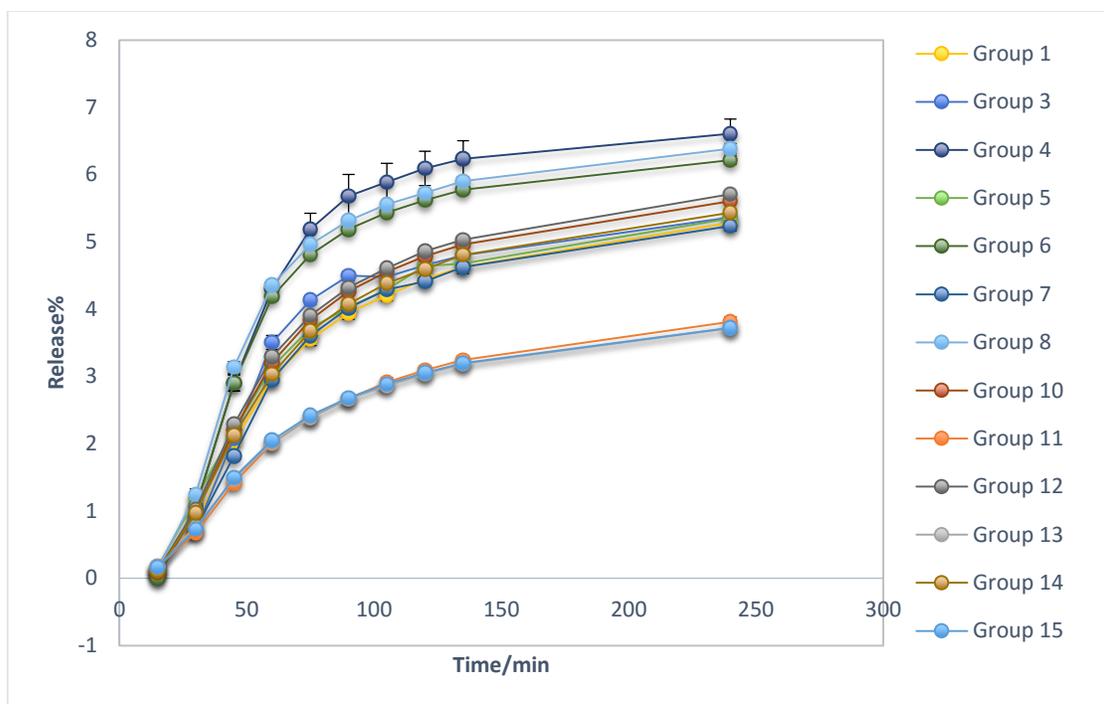
Figure 7-6 Based on level and shape scores, 13 among the 16 groups can be classified into three different categories: Group 1, 3, 5, 7, 10, 12, 14; Group 4, 6, 8; and Group 11, 13, 15. Groups in the same category are considered similar both in terms level and shape

7.6. Tables for Chapter 7

Table 7-1 Design of experiment for the case study presented in this chapter

Strength	Stability Time	Testing Condition	Group No.
1	1	1	1
		2	2
	2	1	3
		2	4
	3	1	5
		2	6
	4	1	7
		2	8
2	1	1	9
		2	10
	2	1	11
		2	12
	3	1	13
		2	14
	4	1	15
		2	16

Table 7-2 MANOVA repeated measures results

Factor		F value	DoF	Prob.>F
Between Subject	Strength	324.105	1	0*
	Stability time	12.517	3	0*
	Testing condition	509.289	1	0*
	Strength*Stability time	4.116	3	0.046*
	Strength*Testing condition	0.0308	1	0.861
	Stability time*Testing condition	0.986	3	0.324
	Strength*Stability time*Testing condition	25.288	3	0*
Within Subject	Time	174.785	9	0*
	Time*Strength	79.115	9	0*
	Time*Stability time	3.989	27	0*
	Time*Testing condition	129.109	9	0*
	Time*Strength*Stability time	2.238	27	0.030*
	Time*Strength*Testing condition	16.403	9	0*
	Time*Stability time*Testing condition	1.914	27	0.066
	Time*Strength*Stability time*Testing condition	2.9614	27	0.005*

* denotes statistical significance

Table 7-3 Arrangement of average profiles for each group

Group No.	15min	30min	45min	…	j	240min	Average
1							
2							
3							
…							
i					y_{ij}		$y_{i.}$
16							
Average					$y_{.j}$		$y_{..}$

Table 7-4 Example of residual matrix

Group	15min	30min	45min	60min	75min	90min	105min	120min	135min	240min
1	0.185	0.084	0.026	-0.106	-0.092	-0.085	-0.069	-0.030	-0.003	0.090
2	-1.047	-1.351	-0.546	0.261	0.241	0.421	0.551	0.564	0.532	0.373
3	-0.023	-0.239	-0.162	0.188	0.249	0.226	-0.032	-0.063	-0.074	-0.069
4	-1.104	-0.863	-0.299	0.029	0.344	0.448	0.414	0.414	0.403	0.214
5	0.103	0.327	0.058	-0.077	-0.073	-0.151	-0.119	0.018	-0.099	0.014
6	-0.830	-0.574	-0.027	0.186	0.243	0.233	0.239	0.216	0.216	0.099
7	0.228	0.038	-0.176	-0.106	-0.028	0.004	0.029	-0.050	0.006	0.055
8	-0.858	-0.539	0.056	0.206	0.240	0.208	0.209	0.173	0.191	0.114
9	0.991	0.591	0.382	0.113	-0.072	-0.228	-0.370	-0.378	-0.526	-0.503
10	-0.089	0.004	-0.049	-0.122	-0.055	-0.016	0.032	0.060	0.077	0.159
11	1.212	0.920	0.371	-0.114	-0.296	-0.407	-0.402	-0.429	-0.432	-0.424
12	-0.121	-0.005	-0.040	-0.101	-0.058	-0.028	0.016	0.063	0.079	0.194
13	1.233	1.010	0.440	-0.093	-0.294	-0.395	-0.443	-0.473	-0.478	-0.507
14	0.067	0.131	-0.003	-0.165	-0.096	-0.080	-0.012	-0.005	0.047	0.117
15	1.203	0.972	0.450	-0.069	-0.270	-0.396	-0.426	-0.466	-0.480	-0.517
16	-0.163	-0.265	-0.335	-0.134	-0.064	0.144	0.104	0.095	0.255	0.361

Table 7-5 Eigenvalues from principal component analysis

Number	Eigenvalue	Percent	Cum.Percent
1	1.5506	94.039	94.039
2	0.055	3.333	97.372
3	0.0297	1.801	99.197
4	0.0072	0.439	99.612

Table 7-6 The eigenvector from principal component analysis that most of variance in residual matrix.

15min	-0.623
30min	-0.527
45min	-0.207
60min	0.060
75min	0.147
90min	0.212
105min	0.233
120min	0.238
135min	0.251
240min	0.217

Table 7-7 ANOVA table for Level

Source	DF	SS	MS	F	P
Strength	1	3.8321	3.8321	72.41	0*
Stability time	3	0.2133	0.0711	1.34	0.315
Testing condition	1	5.8291	5.8291	110.15	0*
Error	10	0.5292	0.0529		
Total	15	10.4037			
S=0.23, R-Sq = 94.91%, R-sq(adj.)=92.37%, R-sq(pred)=91.23%.					

* denotes statistical significance

Table 7-8 ANOVA table for shape

Source	DF	SS Adj.	MS Adj.	F	P
Strength	1	10.8772	10.7882	74.91	0*
Stability time	3	0.9051	0.3017	2.09	0.165
Testing condition	1	14.7866	14.7866	102.67	0*
Error	10	1.4402	0.1440		
Total	15	27.9201			
S=0.38, R-Sq = 94.84%, R-sq(adj.)=92.26%, R-sq(pred) =91.07%.					

* denotes statistical significance

Table 7-9 Effect size test using omega-squared

Response	Effect	Omega-squared
Level	Strength	0.361
	Stability time	0.005
	Testing condition	0.552
Shape	Strength	0.379
	Stability time	0.017
	Testing condition	0.522

Table 7-10 Post-hoc analysis based on level

Group	N	Mean	A	B	C	D	E	F
1	6	3.20				D		
2	6	5.01	A					
3	6	3.49			C	D		
4	6	4.40		B				
5	6	3.34				D		
6	6	4.12		B				
7	6	3.21				D		
8	6	4.27		B				
9	6	2.64					E	
10	6	3.48			C	D		
11	6	2.27						F
12	6	3.51			C	D		
13	6	2.23						F
14	6	3.28				D		
15	6	2.24						F
16	6	3.76			C			

Table 7-11 Post-hoc analysis based on shape.

No.	N	Mean	Grouping				
1	6	-0.21				D	
2	6	2.09					E
3	6	0.13			C	D	
4	6	1.69		B			E
5	6	-0.34				D	
6	6	1.10		B	C		E
7	6	-0.23				D	
8	6	1.06		B	C		
9	6	-1.61	A				
10	6	-0.01				D	
11	6	-1.87	A				
12	6	0.05				D	
13	6	-1.97	A				
14	6	-0.46				D	
15	6	-1.93	A				
16	6	0.24			C	D	

8. Predicting drug *in-vitro* release profiles for pharmaceutical continuous manufacturing processes

This chapter describes joint work with Pallavi Pawar and Golshid Keyvan.

8.1. Introduction

In the last decade, advanced pharmaceutical manufacturing technology, including continuous manufacturing, has been embraced by industry and by regulators.[242] Long ago adopted from other manufacturing industries, such as automobile and semiconductor, advanced manufacturing processes require distributed control systems to enable process control and achieve real-time product quality assurance.[243, 244] Current real time release testing (RTRt) initiative encourages online/inline process monitoring and real-time quality control in a robust and reliable manner.[11] In other words, RTRt is real-time measurement or prediction of product CQAs during the process, without requiring time consuming off-line assays. Importantly, the ability to achieve RTRt builds on process understanding, and appropriate use of process analytical technology (PAT) tools.[245]

In the pharmaceutical industry, drug *in vitro* dissolution profiles, often measured from dissolution testing, are a critical quality attributes (CQAs) of finished products.[11] As discussed previously in Chapter 7, dissolution profiles are commonly used for formulation optimization, equivalence decisions, product compliance, and product release decisions. For decades, regulatory agencies and the pharmaceutical industry have placed significant emphasis on dissolution testing for approval of new products, bio-equivalence studies for generic products, and for approval of existing products post minor changes in

the processing or formulation.[246, 247] Additionally, dissolution testing is also useful during process development to identify critical process parameters in different unit operations.[248-251]

Despite its widespread use in the pharmaceutical industry, dissolution testing has several shortcomings. The test itself is time-consuming, expensive, and requires extensive media and analytical instrument preparation,[82] which makes it difficult to include dissolution testing as part of the RTRt strategy for process control. Furthermore, as it is intrinsically a destructive test, the sample used in dissolution testing cannot be used in other quality measurements. It is thus difficult to conduct gauge repeatability and reliability study for dissolution testing to separate sampling error from overall experimental error.[252]

Near infrared spectroscopy (NIR), being a non-destructive and fast PAT tool, has been extensively investigated in terms of the ability to evaluate the drug release from the final products. Zannikos *et al.* found that NIR spectra was able to capture the effect of humidity in tablet dissolution profiles.[253] Blanco *et al.* combined NIR with partial least squares regression (PLSR) to predict dissolution profiles considering the effects of compaction force and API concentration.[254] Tabasi *et al.* used similar multivariate analysis to predict dissolution profiles at different time points with different tablet coating grades.[255] Previous studies also showed that NIR spectroscopy in combination with appropriate chemometric modeling can be used to predict dissolution considering effect of medium pH, surfactants and crystallization inhibitors, and shear during mixing.[81, 82, 256]

Although attention has been paid to predicting tablet dissolution profiles using non-destructive PAT tools, very few cases considered longitudinal nature of dissolution profiles. Specifically, many of the previous efforts are either predicting drug release at one or multiple time-points, or relying on fitting an empirical kinetic equation to dissolution profiles. However, as explained in Chapter 7, dissolution profiles are, intrinsically, repeated measurements datasets and thus have intrinsic auto-correlation. Single time-point prediction is often not representative to capture overall profile variability. Furthermore, models focused on predicting multiple, or even all time-points neglect autocorrelation and significantly reduce prediction power by building up Type-I error. Fitting model equations to dissolution profiles, on the other hand, is able to transform predicting a full curve into predicting one or two descriptive values. However, since it is a model-dependent approach, its application is limited when a master equation cannot be found to best fit a large number of dissolution profiles available.

The objective of this chapter is to propose a real time release testing method for online dissolution profile prediction. The proposed method includes at-line NIR spectroscopy implementation, process experimental design, statistical analysis of dissolution profile, and multivariate predictive modeling. Particularly, in the case presented here, tablets were manufactured using a continuous manufacturing processes. A fractional factorial experimental design was used including formulation and process variables. M-PCA, a methodology proposed in Chapter 7 that takes into account auto-correlation in dissolution profiles, was used for analysis. Principal component regression models were developed to enable predicting power.

8.2. Materials and method

8.2.1. Materials

The ingredients in this study were semi-fine acetaminophen (APAP) (Mallinckrodt Inc., Raleigh, NC) as the active pharmaceutical ingredient, lactose monohydrate NF (Foremost Farms, Rothschild, Wisconsin) as the filler excipient, and magnesium stearate NF (MgSt) (Mallinckrodt, St. Louis, Missouri) as the lubricant.

8.2.2. Methods

8.2.2.1. Continuous manufacturing direct compression

Tablets were made in a continuous direct compaction (CMDC) line developed at the Engineering Research Center for Structured Organic Particulate Systems (ERC-SOPS). A schematic illustration of the CMDC line is shown in **Figure 8-1.** The manufacturing line consisted of a KT20 and a KT35 gravimetric feeders (Coperion K-Tron, Sewell, NJ) used for feeding APAP and Lactose respectively. MgSt was fed through a MT12 feeder (Coperion K-Tron, Sewell, NJ). Tablet production rate was kept constant at around 20 kg/hr. The feed rate in the feeders was adjusted to obtain the desired concentrations of individual raw materials. The APAP and the Lactose were fed into a S197 Comil (Quadro Engineering, Waterloo, Canada), which was used to delump the active and the excipient before entering continuous mixer.

Ingredients were blended in a Glatt GCG 70 mixer with 24 blades arranged in the "1/3 forward + 1/3 alternate + 1/3 forward" blade configuration (Glatt group, Binzen, Germany). In this blade configuration, only the middle 1/3 (or blades 9-16) are alternated forward 45° and backward 45°. Previous study concluded that the "1/3 forward" configuration yielded improved mixing performance without sacrificing throughput.[193]

After passing through the mixer, the blend passed through a chute into a tablet rotary press. A 36-station Kikusui Libra2 tablet press (KIKUSUI America Inc., Santa Clara, CA) was used in a single layer configuration. The press was fitted with a type B flat tooling to make tablets of 10 mm diameter. The fill depth was adjusted to obtain 350 mg tablets. The press was operated at a variable feed-frame speed but a constant turret speed of 20 rpm for the 20 kg/h flow rate.

In this study, the target drug concentration and process parameters were 9% APAP, compaction force of 24 kN, blender speed at 200 rpm and feed frame speed of 25 rpm. The aim of this study is to develop predictive models to for tablet dissolution profiles with target settings.

8.2.2.2. *Experimental design*

Four variables were included in the experimental design: API concentration (low 5%, medium 9%, high 13%), blender speed (150 rpm, 200 rpm, 250 rpm), feed frame speed (20 rpm, 25 rpm, 30 rpm), and compaction force (8 kN, 15 kN, 24 kN). The variables and their coded factors are tabulated in **Table 8-1**. The experimental design consisted of a 3^{4-1} fractional factorial design with additional three center point replications, which leads to a total of 30 conditions. All conditioned in the experimental design are detailed in **Table 8-2.** For each condition, six tablets were analyzed for dissolution tests, resulting into a total of 180 (30*6) tablets under investigation in the study.

8.2.2.3. *NIR transmission spectroscopy*

A Bruker Optics Multipurpose Analyzer (MPA) FT-NIR spectrometer (Bruker Corporation, Billerica, Massachusetts) was used to collect tablet spectra in transmission

mode at line. In transmittance mode, the light interacts with the tablet, as opposed to tablet surface in reflectance mode. The spectral range chosen was from 12500 to 5800 cm^{-1} with a resolution of 32 cm^{-1}. A background and a tablet spectrum were obtained by averaging a total of 256 background scans and sample scans respectively.

8.2.2.4. *Dissolution testing*

A total of 180 dissolution profiles of the tablets obtained from the continuous line described previously were used to build a calibration model. A VK 7010 dissolution apparatus (Varian Inc., Santa Clara, CA) fitted with USP paddles was used to study the drug release. The paddle rotational speed was 50 rpm. The dissolution medium used was phosphate buffer at pH 5.8 and the media temperature was maintained at 37.0 ± 0.5^{o}C. Six tablets were each placed in the dissolution apparatus chambers containing 900 ml of medium, simultaneously at the start of the experiment. A peristaltic pump VK 810 (Varian Inc., Santa Clara, CA) was used to pump out aliquots of the dissolution medium at 3-minute time intervals. The medium was then filtered through 35 μm full flow filters prior to detection using a UV spectrophotometer (Varian Inc., Santa Clara, CA) at a wavelength of 243 nm. Absorbance values for each tablet were converted to the percent of drug released at each analysis time. For each tablet, a dissolution profile was obtained recording the change of drug release percent over time.

8.2.2.5. *Chemometrics*

The data obtained from the NIR spectrometer was analyzed using Unscrambler X 10.2 (Camo Inc., Oslo, Norway). The data was baseline corrected, where the lowest absorbance value in the spectrum was subtracted from the absorbance values at all the other wavenumbers. This was followed by a Savitzky-Golay first derivative, fitted to a

second order polynomial with a total of 11 smoothing points. This pretreatment removed the baseline offset between samples and improved the resolution of the peaks that carried the chemical information associated with the concentration of the API. Principal Component Analysis (PCA) was then performed on pre-treated data. Calculation of PCA can be found in Chapter 2.

8.2.2.6. *Dissolution profile analysis*

A previously developed methodology in Chapter 7, M-PCA, was used in this chapter for dissolution profile prediction. A detailed description of the M-PCA can be found in Chapter 2. In brief, the methodology looks at a dissolution curve in terms of its level and shape separately. The advantage of this approach is that it takes into account the auto-correlation in the dissolution profile as time series data. The level is extracted based on clustering analysis, and is thus intrinsically associated with similarity/dissimilarity values. In addition, it is a model-independent approach for dissolution profile analysis. In other words, this method does not rely on fitting dissolution profiles using any empirical equations. Software JMP 10 (SAS institute, Inc., Cary. NC) was used for M-PCA analysis.

8.2.2.7. *Prediction models*

A principal component regression (PCR) model was developed. The structure of PCR model was discussed in more details in Chapter 2. The model was based on the assumption that, in this experimental design, NIR spectroscopy captures variability that contributes to dissolution profile changes. Scores of principal components from chemometric analysis were used as PCR model input, while response variables were level and shape values obtained from dissolution profile analysis.

8.3. Results and discussion

8.3.1. Chemometric analysis

As dissolution profiles were potentially affected by four variables in this study, the aim of NIR data pre-treatment was to improve the resolution that helps to distinguish contributions of each variable. A baseline offset was observed to be present in the NIR absorbance data. The offset was removed by subtracting the lowest intensity in the spectra from all the other variables for each sample. This treatment was useful in removing the effect of scattering on the NIR signal.[257] The effect of compaction force was evidently shown on the baseline corrected NIR absorbance spectra (**Figure 8-2**). It was observed that the absorbance decreased with an increase in the compaction force. In addition to baseline correction, a Savitzky-Golay first derivative (second order polynomial, 11-point smoothing) treatment was performed to separate the drug concentration effects. [258] The second order polynomial and an 11- point smoothing enhanced the signal-to-noise ratio without compromising on the resolution of the peaks.

PCA on the pre-treated NIR data yielded three principal components accounting for 99% of the total variability in the data. The first principal component (PC1) explained 85% of total variance in the data. As shown in **Figure 8-3**, three clusters were observed in the scores plot along the PC1 axis, suggesting PC1 represents tablet variability contributed by compaction force. In terms of the second principal component, as shown in **Figure 8-4**, the effect of API concentration in the formulation can be distinguished along the axis of PC2. In addition to the first two PCs, the third principal component was also considered in this study. Although PC3 only accounted for 2% variability in the dataset, it was included as it potentially related to amount of shear during mixing based on previous

studies.[82] As discussed in Chapter 6, the shear experienced by the powder blends leads to the coating of MgSt on other particles, affecting the particle bonding, and hence the internal structure of the tablets.[259] The shear experienced by the powder in the continuous line is known to be affected by the blender speed and the feed frame rotation rate and is also known to have an impact on the micro- mixing on MgSt, which in turn affects tablet dissolution.

The scores of PC1, PC2, and PC3 for all 180 tablets in this study were obtained. To leverage tablet-to-tablet variability, for each processing condition, the average scores of each PC was calculated. The average scores were then used for principal component regression.

8.3.2. Principal component regression (PCR)

PCR was used to examine the relationship between the eigenvalues obtained from the chemometric analysis and the dissolution parameters, level and shape, obtained using M-PCA. As described previously, the level using M-PCA was calculated by using the average of individual dissolution profile across all the time points. The shape was calculated by performing a PCA on the residual matrix after the level was subtracted. The principal components were based on the covariance matrix. The first two PCs explained 98.7% of the variability of the dataset, and thus two shape values, corresponding to "shape I" and "shape II" were calculated. PCR models were developed to predict level, shape I, and shape II separately. The equations can be expressed as:

Level = 74.12 - 81.66 * PC1 + 103.41 * PC2 + 30.17 * PC3 + 377.84 * PC1*PC2 +

  1762.93 * PC1*PC3 - 653.278 * PC2*PC3 - 38489.1 * PC1*PC2*PC3

Shape I = 3.57 - 330.25 * PC1 + 254.11 * PC2 + 176.14 * PC3 - 1050.08 * PC1*PC2 +

$$6994.8 * PC1*PC3 - 2452.04 * PC2*PC3 - 109821 * PC1*PC2*PC3$$

$$\text{Shape II} = -0.81 + 65.23 * PC1 - 100.16 * PC2 + 15.17 * PC3 - 566.01*$$

$$PC1*PC2 - 1989.88 * PC1*PC3 + 1310.25 * PC2*PC3 + 41747.1 * PC1*PC2*PC3$$

The equation for each of the parameters was calculated using standard least squares algorithm. The equation provided a relatively good fit with R^2 values of 0.80, 0.82 and 0.73 for level, shape I and shape II predictions respectively ($p<0.001$). The parity plots are shown in **Figure 8-5.** Both internal validation and external validation were further conducted to evaluate predicting power of the model.

8.3.3. PCR model internal validation

Ten tablets from the 180 tablets were used in the internal validation set. These tablets were randomly picked from these 180 tablets and covered varying conditions with respect to the four input variables. The internal validation set was projected into the scores plot space as shown in **Figure 8-6**. All the ten tablets were within the 95% confidence interval as marked by the Hotelling's T^2 ellipse in the figure. Similar to cross-validation, performed in Chapter 4, the PCA followed by the PCR was repeated for the remaining 170 tablets. The dissolution profiles of the ten tablets were predicted using the regression models developed.

f_1 (difference factor) and f_2 (similarity factor) indices were used to compare the measured dissolution profiles and the predicted dissolution profiles. As mentioned in Chapter 7, using f_1 and f_2 analysis is a common approach recommended by FDA's dissolution testing Guidance for Industry. It is a model- independent approach to define dissolution similarity based on the mean-squared difference between a pair of profiles. A predicted dissolution profile is considered similar to the reference profile when f_1 value is less than

15 and f_2 value between the two profiles is more than 50. **Table 8-3** tabulates the f_1 and f_2 values for the ten tablets. The predicted results were similar to experimental results for all ten tablets for internal validation. This is remarkable, considering that the typical use of f_1 and f_2 is for comparison of profiles averaged for 12 tablets, not individual tablets, that are subjected to significantly more experimental variability.

8.3.4. Prediction results for tablets manufactured from target process space

In order to further examine robustness of the prediction model, additional tablets were made at the target design space, namely drug concentration of 9%, compaction force of 24 kN, blender speed of 200 rpm, and feed frame speed of 25 rpm. Six tablets were sampled and used as external validation set for prediction. The NIR spectrum of the six tablets were projected to the calibration tablets. The scores of first three PCs were obtained. The scores of the six tablets were used as inputs of the PCR model. Once the predicted level, shape I, and shape II for each tablet was predicted, the predicted dissolution profiles were constructed. The tablets were subjected to dissolution tests to verify the prediction results. **Figure 8-7** displays the reference and the predicted dissolution profiles for individual tablets. The values of similarity factor and difference factor are also shown. For all six tablets, the prediction results were in good agreement with the experimental results.

The results confirmed the hypothesis that NIR spectroscopy is able capture critical changes in formulation and process settings that contribute to variations in tablet dissolution profiles. The proposed method is able to predict dissolution profiles using NIR spectroscopy in combination with appropriate multivariate analysis models. Important to notice, the way that M-PCA works is that it finds the most similar profile in

the calibration set to the predicted profile (both level and shape). In other words, it does not create a new profile based on the calibration sets. The precision of the prediction requires the calibration sets to be fully comprehensive and diversified for the combination. Developed models can be maintained augmented by additional measurement performed during process scale-up, validation, and commercial manufacturing to capture process knowledge and increase predictability.

8.4. Conclusions

This chapter addressed a challenge in implementing real-time release methods for advanced tablet manufacturing process: the need to non-destructively predict dissolution profiles. A quality-by-design approach was used to predict tablet dissolution profiles using NIR spectroscopy in combination with appropriate multivariate analysis methods. Designed experiments were used to understand the impacts of process changes. A PCA algorithm was used to extract the effect of process and formulation variation from the Near IR data of tablets. It was observed that changes in API concentration and process parameters can be reflected in NIR spectrum using appropriate pre-treatment processing. The established principal component regression models were able to predict the dissolution profile of individual tablets based on its NIR spectrum. The prediction power of the models was confirmed by ten internal validation samples and additional six external samples. The prediction results were in a good agreement with experimental results.

Only one type of formulation was examined in this paper. This formulation has relatively simple dissolution behavior, determined primarily by three parameters: tablet porosity, amount of API, and extent of shear experienced by the blend. For such a system, a small

number of degrees of freedom are needed to predict dissolution behavior. More complex formulations where drug release might depend on amounts of other ingredients (such as controlled-release polymers and pH modifiers), it might require a larger number of factors, and perhaps a more extensive calibration set. However, to the extent that the relevant parameters can be tested non-destructively, the methods introduced here are likely to be effective. Importantly, once the ability to predict dissolution profiles reliably is properly demonstrated, meeting the remaining requirement for closed-loop quality control, real-time quality assurance, and real-time release should be relatively straightforward.

8.5. Figures for Chapter 8
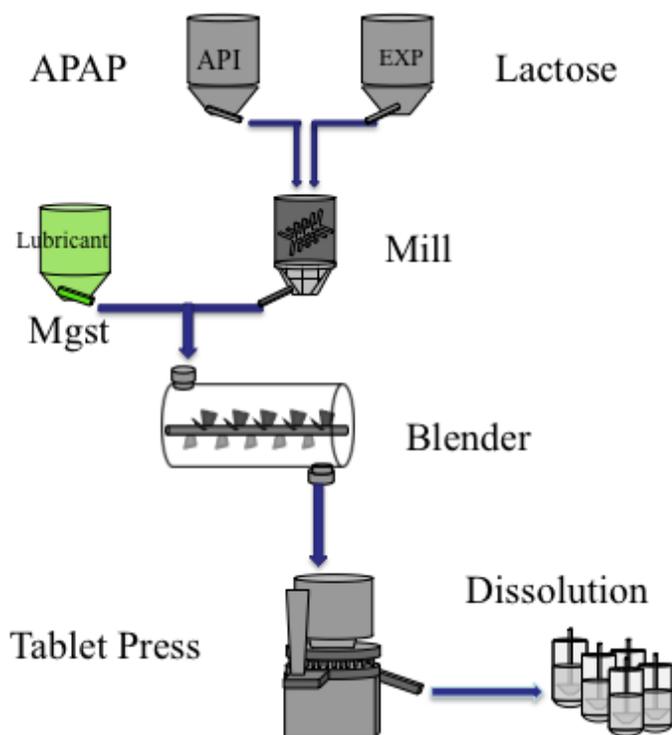
Figure 8-1 Schematic illustration of continuous manufacturing direct compression line. (Picture source: Sebastian Escotet, Fernando Muzzio)
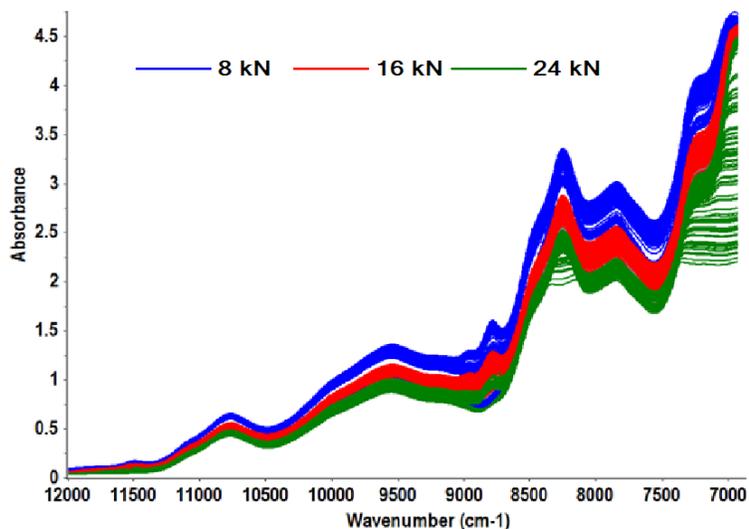
Figure 8-2 NIR spectrum after baseline correction. The absorbance value was observed to decrease as compaction force increases.
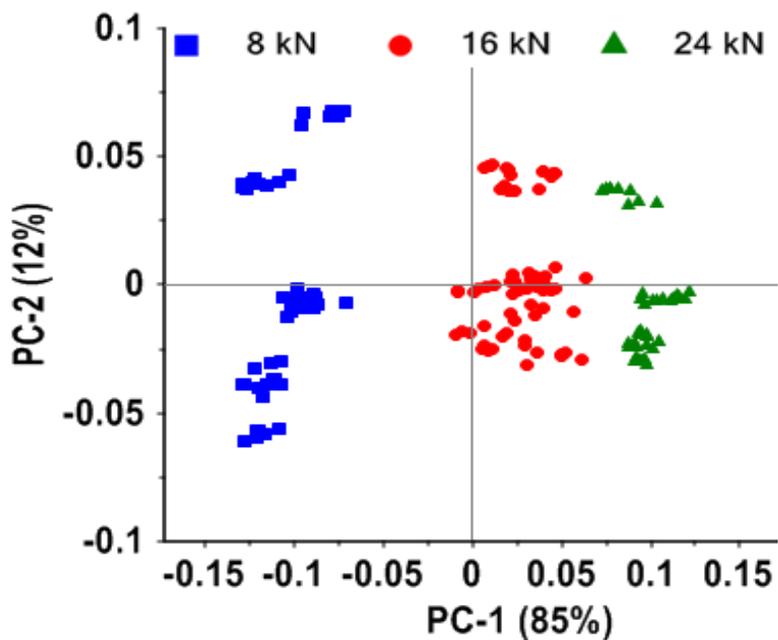
Figure 8-3 A scatter plot of the first two principal components (PCs). PC1 is shown to account for variations contributed by compaction force.
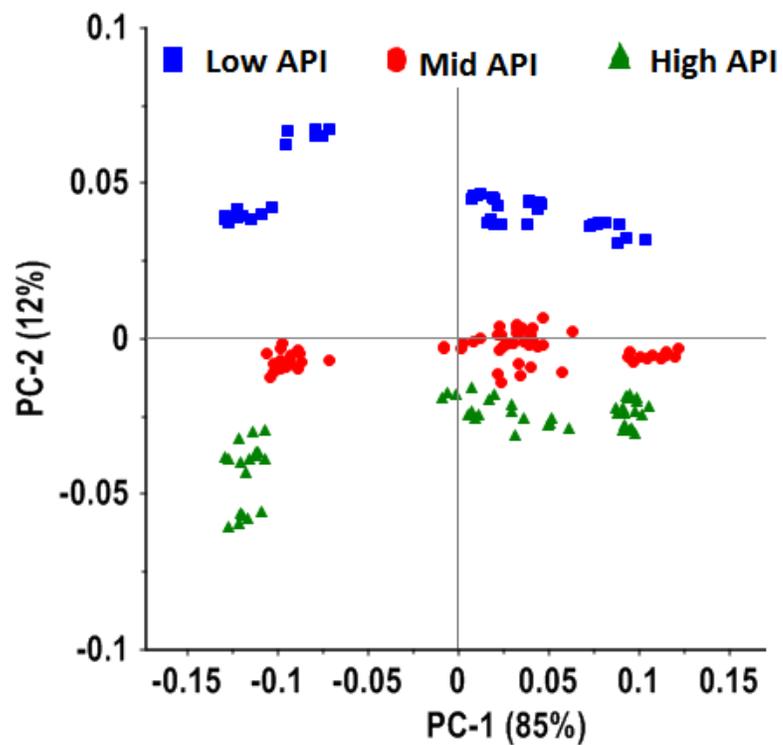
Figure 8-4 A scatter plot of the first two principal components. PC2 is shown to represent variations caused by API concentration in the formulation.
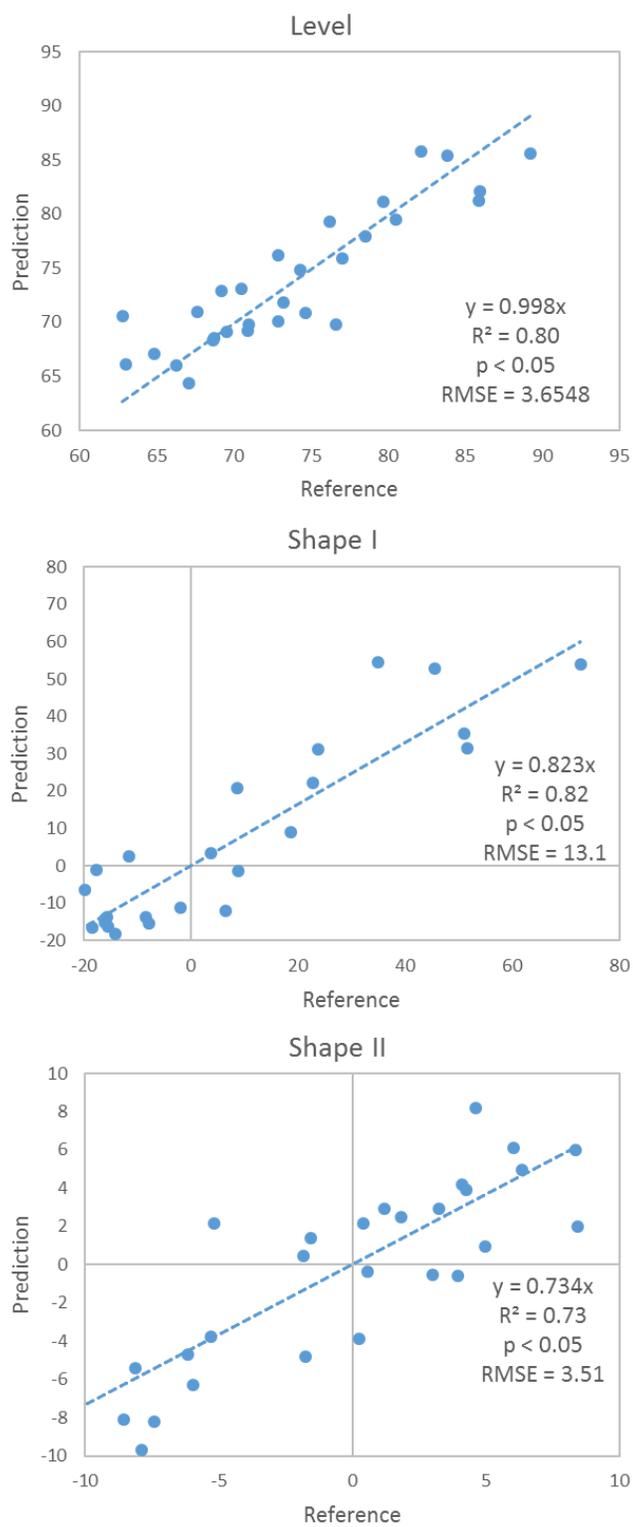
Figure 8-5 Predicted versus reference parity plot for level and shape values.
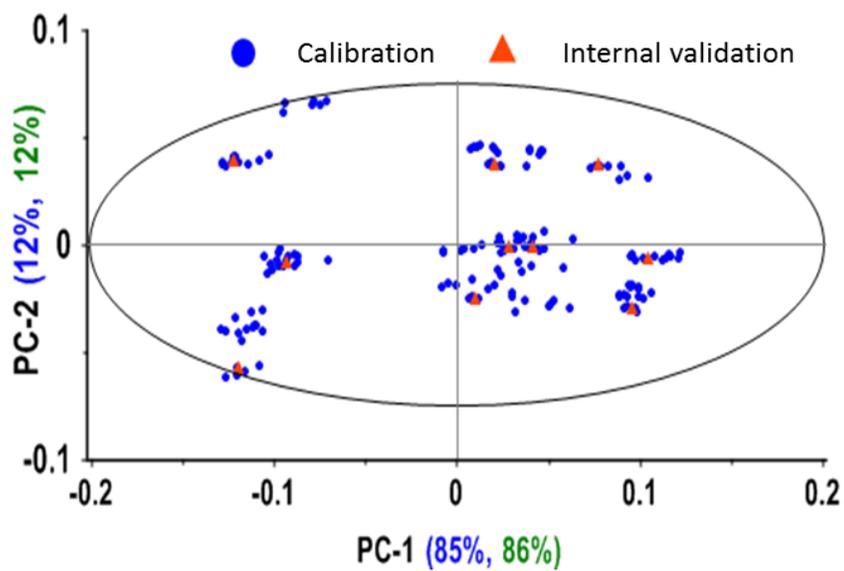
Figure 8-6 Projection of ten tablets used for internal validation. The projected tablets were shown to have varying process parameters, and were within the 95% Hotelling's ellipse.
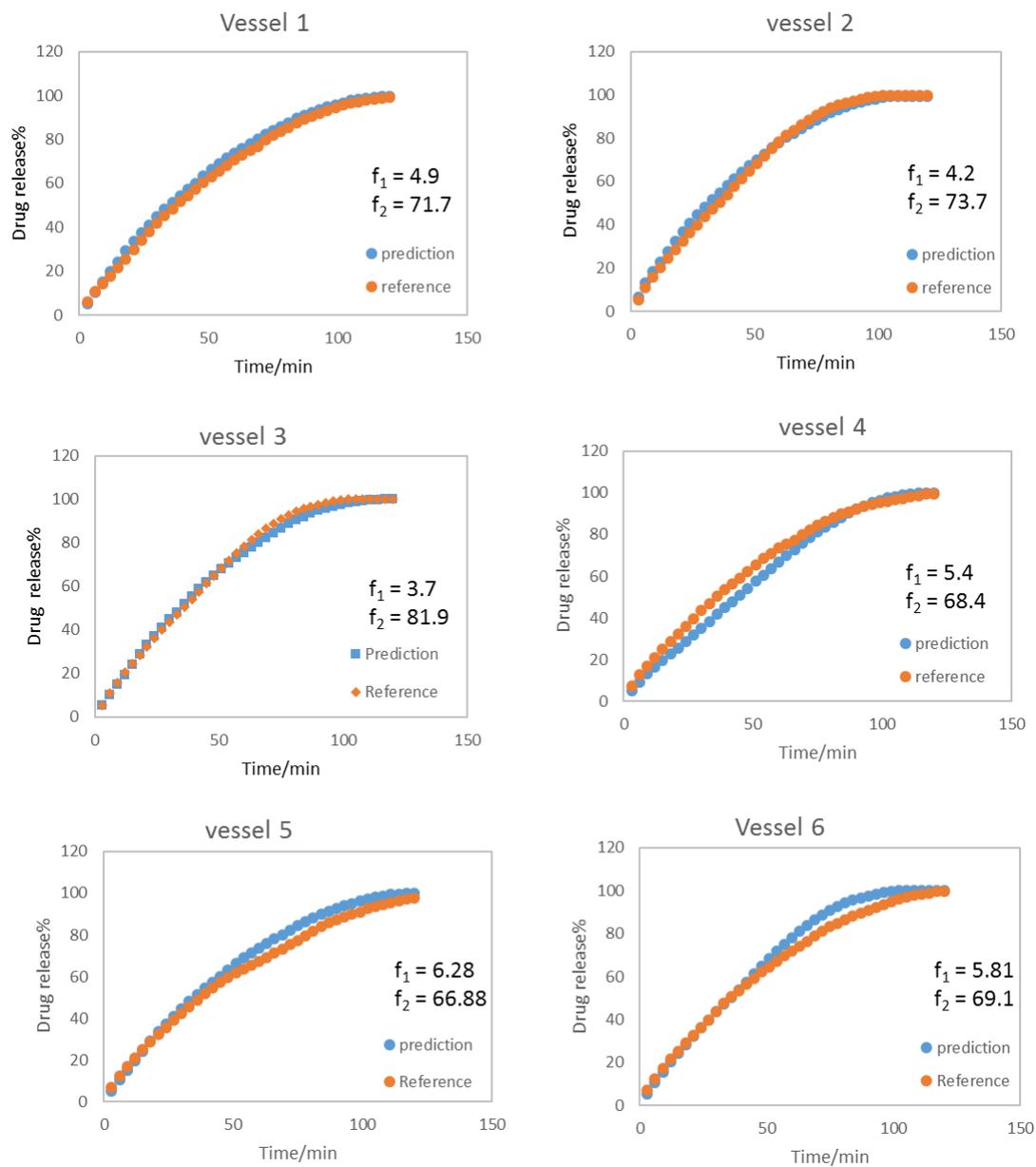
Figure 8-7 Comparison between experimental results (reference) and prediction results for six external validation tablets.

8.6. Tables for Chapter 8

Table 8-1 Design space using coded factors

	A	B	C	D
coded level	API Concentration (%)	blender (RPM)	feed frame (RPM)	compaction force (kN)
0	5	150	20	8
1	9	200	25	16
2	13	250	30	24

Table 8-2 Fractional factorial design in this study using coded factors. A total of 30 experiments were included.

	A	B	C	D
1	0	0	0	0
2	0	0	1	1
3	0	0	2	2
4	0	1	0	1
5	0	1	1	2
6	0	1	2	0
7	0	2	0	2
8	0	2	1	0
9	0	2	2	1
10	1	0	0	1
11	1	0	1	2
12	1	0	2	0
13	1	1	0	2
14	1	1	1	0
15	1	1	2	1
16	1	2	0	0
17	1	2	1	1
18	1	2	2	2
19	2	0	0	2
20	2	0	1	0
21	2	0	2	1
22	2	1	0	0
23	2	1	1	1
24	2	1	2	2
25	2	2	0	1
26	2	2	1	2
27	2	2	2	0

28	1	1	1	1
29	1	1	1	1
30	1	1	1	1

Table 8-3 Similarity and different factor between predicted results and experimental results for model internal validation.

Internal validation tablet	Predicted vs. Measured	
	f_1	f_2
1	9.83	55.13
2	5.34	59
3	1.62	64.35
4	5.75	60.72
5	2.89	78.66
6	1.4	89.87
7	1.08	94.37
8	1.16	93.49
9	2.99	68.56
10	3.92	72.80

9. Conclusions and recommendations

The work presented in this dissertation focused on using multivariate analysis tools to improve understanding of pharmaceutical powder-based processing, and to optimize analytical methods for characterizing material properties and final product quality. Throughout the case studies examined here, the ability of predicting process performance, critical material properties, and critical quality attribute was achieved. Flow properties of raw materials and their effects on feeder performance were examined. An analytical method to analyze data from shear cell data was introduced. Commonly used techniques to characterize powder flow properties were investigated and used to predict gravimetric feeder performance. Mixing effects on flow properties of intermediate blends were also discussed. A mixture model was developed to predict cohesion of powder mixtures. Drug *in vitro* release profile was studied using a statistically reliable approach. A toolbox of statistical methods for dissolution profile comparisons was proposed. In addition, a nondestructive method for predicting dissolution profiles was developed. This chapter summarizes work presented in the dissertation and outlines recommendations for future work.

9.1. Conclusions

The first specific aim focused on developing methods for analyzing flow properties of raw materials, and using the methods to predict process performance. Chapter 3 of the dissertation augmented the shear cell data analysis through understanding the relationships of different measurements. A dimensionless cohesion C^* was defined. The ratio between the flow function coefficient and C^* was considered to be a testing equipment characteristic, which can also be used to describe the relationship between

applied normal stress and the characteristic state when materials are critically consolidated. In addition, a material characteristic line was generated to collapse all the yield loci of a material into a single line that is independent of the initial consolidation stress. Identifying material characteristic lines enables comparison of materials measured under different initial consolidations stresses.

Chapter 4 explored the possibility to correlate raw material flow properties with feeder performance. In addition to shear cell tests, other commonly used characterization techniques were included in the study. It was shown that material flow properties affect feeder performance, and therefore, selection of screw that achieves best feeding performance is dependent on material flow properties. Two approaches were discussed: principal component analysis followed by similarity scoring (PCA-SS), and partial least squares regression (PLSR). This chapters showed that by characterizing flow properties of a material, its initial feed factor, or maximal capacity, of each screw can be predicted, which enables to quickly identify operation ranges. Using PCA-SS method, placebo material with similar flow behavior can be identified. When a material is not available in large amounts, placebo materials can be used for process development. Additionally, instead of selecting optimal tooling based on trial-and-error, by developing predictive PLSR models, feeding performance using different screws can be directly quantified, predicted and compared.

The second specific aim looked into mixing powders and predicting the effect of the mixing process on flow properties of intermediate blends. In Chapter 5, a model was developed to predict the flow regime of pharmaceutical blends based on a four-component mixture design. A special cubic model was established using 22 calibration

blends including four-component powder mixtures and raw materials. ANOVA was used for model selection and optimization. The work presented showed an effective approach to characterize flow properties of powder mixture using statistical methods. Using a quality-by-design approach, the cohesion of blend mixtures can be predicted based on the concentration of each ingredient. The model can be used to screen through a large number of formulations to avoid cohesive formulations with little materials or manpower cost. In addition, the methods to analyze shear cell data, introduced in Chapter 3, were validated using powder mixtures. Importantly, the prediction model can be extended by understanding effect of mixing on blend flow properties in different mixing systems.

Chapter 6 compared mixing effect in the controlled shear system (CSS) and in a laboratory scale Resonant Acoustic Mixer (LabRAM). Experimental results suggested that mixing without MgSt created more cohesive blends. Using ANOVA and effect size test, lubrication rate in both mixing systems had minimal effects on blend flow properties. Results also showed that although the two devices had comparable lubrication effect on the overall blend flowability, the changes of the interparticle forces were not identical. The frictional forces can be well correlated with the bulk density of lubricated blends, while the correlations between the bulk density and cohesive forces were significantly different.

The third specific aim developed statistically justified methodologies to analyze, compare, and predict drug *in vitro* release profiles considering both formulation and process variables. In Chapter 7, rigorous, yet easy to use methodologies are introduced for the determination of statistical significance of observed multivariate effects in dissolution testing. MANOVA with repeated measures and modified PCA (M-PCA)

methods in combination were able to offer detailed and valuable information. The proposed methods consider the multivariate nature of the information, the self-correlated intrinsic nature of dissolution profiles, and the availability of within-group variability as a means for determining reliability and significance. Chapter 8 further demonstrated application of M-PCA for dissolution profile prediction. A quality-by-design approach was used to predict tablet dissolution profiles using NIR spectroscopy in combination with appropriate multivariate analysis methods. Experimental results showed that that changes in API concentration and process parameters can be reflected in NIR spectrum using appropriate pre-treatment processing. The established principal component regression models were validated as robust and reliable, and were able to predict the dissolution profile of individual tablets based on its NIR spectrum. The work presented improved real-time release strategy for advanced tablet manufacturing process by achieving predictive capability for nondestructive dissolution testing.

9.2. Recommendations for future work

Based on the work presented in this dissertation, several potential areas are discussed here for future study.

9.2.1. Application of PCA-SS method for measurement reduction

In Chapter 4, PCA-SS approach was proposed and used to identify materials of similar flow properties. The materials with matching flow properties can be used as placebo, or surrogate materials for formulation and process development. For example, **Figure 9-1** used PCA to find surrogate materials with matching flow properties for API 4. This approach is especially helpful when the amount of drug substance is limited. Another potential application of PCA-SS is to reduce dimensionality of characterization

techniques. As discussed previously, flow properties of materials do not all vary independently. Many of the parameters are often highly correlated, suggesting that variables of interest vary in a much lower dimensional space than the number of measurements. Identifying a reduced set of measurements, based on lower dimensional space, is useful to characterize a material using minimal cost without compromising maximal information that can be obtained. Each eigenvector of the PCA provides information on contribution of each property to classifying materials. In other words, based on loading plots, the number of measurements can be reduced by removing material properties that have little contribution to the total variability of the dataset. The selection of the reduced set of measurements can be based on amount of testing time, amount of materials required, or availability of testing equipment. The prioritized selection can be archived by assignment different weights to each flow property before PCA is performed. Similarity scoring can be used to justify similarity of material scores between full and reduced characterized set. Specifically, by conducting iterations that include all possible reduction combinations, the reduced set that generates linearly correlated similarity scores with full characterization set can be selected as the optimal reduced set of measurements. The reduced measurements can then be used to find material surrogates with matching properties, and to predict process performance.

9.2.2.   Blend wettability measurement using droplet penetration technique

Blend wettability, as an intermediate material property, reflects process changes in pharmaceutical manufacturing and helps to guide process scale-up and transfer. As blend wettability affects tablet dissolution profiles, it is useful to develop a predictive correlation between the blend wettability measurements and tablet dissolution results.

For loose powders, techniques have been available to indirectly determine the contact angle based on liquid penetration process. Efforts have been devoted to developing a method based on image analysis of droplet penetration process on porous powder bed to measure its contact angle. A schematic experimental set-up is shown in Figure **9-2.** In a preliminary case study, pharmaceutical blends subjected to different level of shear, described in Chapter 6, were prepared. **Figure 9-2** shows that recorded droplet penetration time was able to reflect effect of process changes on blend wettability. Using video processing, it is possible to obtain penetration profiles that describe change of droplet volume over time. Changes in penetration profiles can be correlated to changes in the dissolution profiles of tablets. As both penetration profiles and dissolution profiles are repeated measurements, as explained in Chapter 2, M-PCA can be used to develop the predictive correlation, which can be used in combination with models developed in Chapter 8 for dissolution profile prediction.

9.3. Figures for Chapter 9
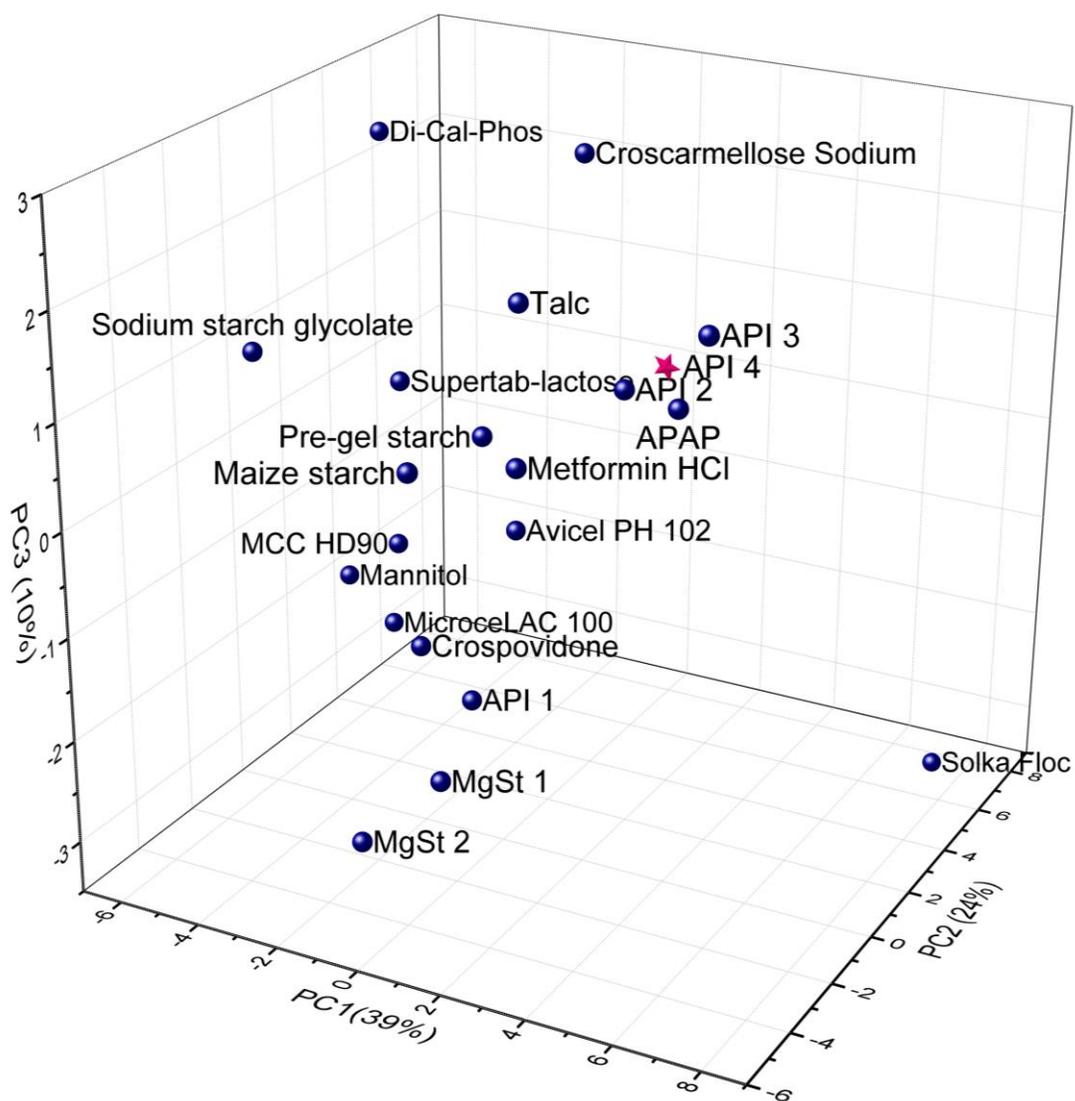

Figure 9-1 Using PCA to find surrogate materials with matching flow properties for API 4 during early formulation development when amount of materials is limited.
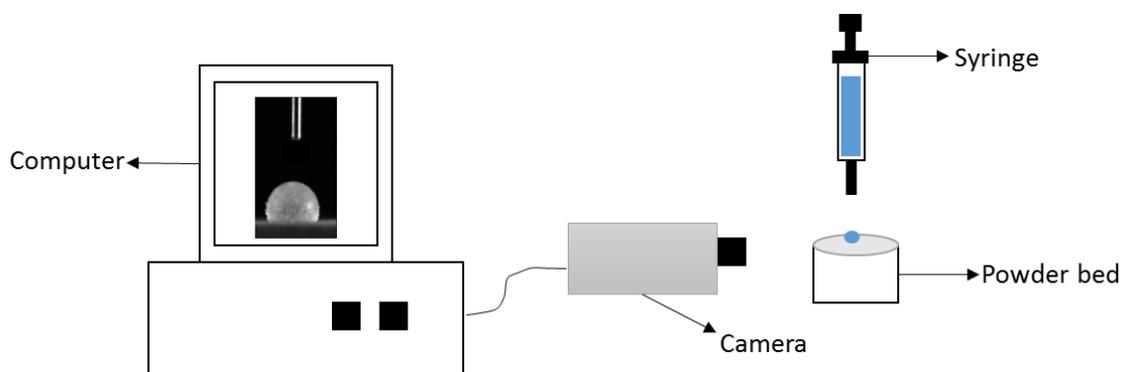
**Figure 9-2 A schematic illustration for droplet penetration method set-up.**
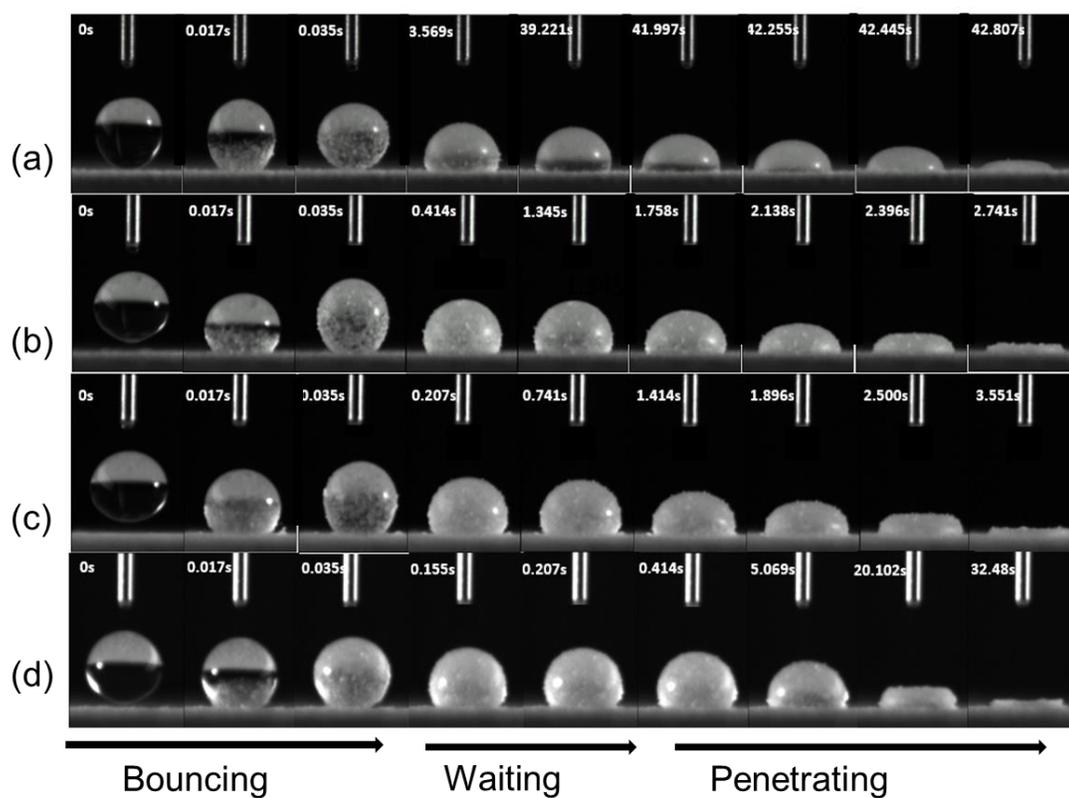
Figure 9-3 Water droplet penetration process for blends with (a)0Rev shear, (b) 160Rev shear, (c) 640Rev shear, and (d) 1280Rev shear.

Reference

1. Folland, S., A.C. Goodman, and M. Stano, *The economics of health and health care*. Vol. 6. 2007: Pearson Prentice Hall New Jersey.
2. Unit, T.E.I., *World industry outlook: Healthcare and pharmaceuticals.* May 2014.
3. Shaw, F.E., et al., *The Patient Protection and Affordable Care Act: opportunities for prevention and public health.* The Lancet, 2014. **384**(9937): p. 75-82.
4. Danzon, P.M., A.W. Mulcahy, and A.K. Towse, *Pharmaceutical pricing in emerging markets: effects of income, competition, and procurement.* Health economics, 2015. **24**(2): p. 238-252.
5. Scherer, F.M., *The pharmaceutical industry.* Handbook of health economics, 2000. **1**: p. 1297-1336.
6. *Advanced manufacturing: A snapshot of priority technology areas across the federal government*, W.H.O.o.S.a. Technology, Editor. April, 2016.
7. Muzzio, F.J., T. Shinbrot, and B.J. Glasser, *Powder technology in the pharmaceutical industry: the need to catch up fast.* Powder Technology, 2002. **124**(1): p. 1-7.
8. Schaber, S.D., et al., *Economic analysis of integrated continuous and batch pharmaceutical manufacturing: a case study.* Industrial & Engineering Chemistry Research, 2011. **50**(17): p. 10083-10092.
9. CDER, T.P.t.a.m.s.w.g., *Innovation and continuous improvement in pharmaceutical manufacturing*, F.a.D. Administration, Editor. 2004.
10. Drennen III, J.K., *Quality by Design—What Does it Really Mean?* Journal of Pharmaceutical Innovation, 2007. **2**(3-4): p. 65-66.
11. Lawrence, X.Y., *Pharmaceutical quality by design: product and process development, understanding, and control.* Pharmaceutical research, 2008. **25**(4): p. 781-791.
12. Llusa, M., et al., *Measuring the hydrophobicity of lubricated blends of pharmaceutical excipients.* Powder Technology, 2010. **198**(1): p. 101-107.
13. Gad, S.C., *Pharmaceutical manufacturing handbook: production and processes*. Vol. 5. 2008: John Wiley & Sons.
14. Amidon, G.E., *Data driven formulation development using material sparing methods*, in *4th Annual Garnet E. Peck Symposium*. 2006: Ann Arbor, MI.
15. Matero, S., et al., *Towards better process understanding: chemometrics and multivariate measurements in manufacturing of solid dosage forms.* Journal of pharmaceutical sciences, 2013. **102**(5): p. 1385-1403.
16. FDA, P., *guidance for industry–a framework for innovative pharmaceutical development, manufacturing and quality assurance, September 2004*. 2012.
17. Simon, L.L., et al., *Assessment of recent process analytical technology (PAT) trends: a multiauthor review.* Organic Process Research & Development, 2015. **19**(1): p. 3-62.
18. Freeman, R., *Measuring the flow properties of consolidated, conditioned and aerated powders—a comparative study using a powder rheometer and a rotational shear cell.* Powder Technology, 2007. **174**(1): p. 25-33.
19. Fu, X., et al., *Effect of particle shape and size on flow properties of lactose powders.* Particuology, 2012. **10**(2): p. 203-208.
20. Althaus, T.O. and E.J. Windhab, *Characterization of wet powder flowability by shear cell measurements and compaction curves.* Powder technology, 2012. **215**: p. 59-65.
21. Anand, O., et al., *Dissolution testing for generic drugs: an FDA perspective.* The AAPS journal, 2011. **13**(3): p. 328-335.
22. LeBlond, D., et al., *In vitro dissolution curve comparisons: a critique of current practice.* Dissolut Technol, http://www.qualitydigest.com/inside/quality-insider-article/problems-risk-priority-numbers.html, 2015.

23.   Pestieau, A., et al., *Towards a real time release approach for manufacturing tablets using NIR spectroscopy.* Journal of pharmaceutical and biomedical analysis, 2014. **98**: p. 60-67.

24.   Solvang, S. and P. Finholt, *Effect of tablet processing and formulation factors on dissolution rate of the active ingredient in human gastric juice.* Journal of pharmaceutical sciences, 1970. **59**(1): p. 49-52.

25.   Huang, J., C. Goolcharran, and K. Ghosh, *A Quality by Design approach to investigate tablet dissolution shift upon accelerated stability by multivariate methods.* European Journal of Pharmaceutics and Biopharmaceutics, 2011. **78**(1): p. 141-150.

26.   Bates, L., *Bulk Solids Characterisation-The Need for industrial Education in Bulk Technology.* Bulk Solids Handling, 2006. **26**(7): p. 464-475.

27.   Ennis, B.J., J. Green, and R. Davies, *The legacy of neglect in the US.* Chemical engineering progress, 1994. **90**(4): p. 32-43.

28.   Mosteller, F. and J.W. Tukey, *Data analysis and regression: a second course in statistics.* Addison-Wesley Series in Behavioral Science: Quantitative Methods, 1977.

29.   Hoerl, R.W. and R. Snee, *Statistical thinking and methods in quality improvement: a look to the future.* Quality Engineering, 2010. **22**(3): p. 119-129.

30.   De Beer, T., et al., *Optimization of a pharmaceutical freeze-dried product and its process using an experimental design approach and innovative process analyzers.* Talanta, 2011. **83**(5): p. 1623-1633.

31.   Goutte, F., et al., *Power of experimental design studies for the validation of pharmaceutical processes: case study of a multilayer tablet manufacturing process.* Drug development and industrial pharmacy, 2002. **28**(7): p. 841-848.

32.   Ragonese, R., et al., *The use of the Box–Behnken experimental design in the optimisation and robustness testing of a capillary electrophoresis method for the analysis of ethambutol hydrochloride in a pharmaceutical formulation.* Journal of pharmaceutical and biomedical analysis, 2002. **27**(6): p. 995-1007.

33.   Clarke, F.C., et al., *Determination of the information depth and sample size for the analysis of pharmaceutical materials using reflectance near-infrared microscopy.* Applied spectroscopy, 2002. **56**(11): p. 1475-1483.

34.   Diletti, E., D. Hauschke, and V. Steinijans, *Sample size determination for bioequivalence assessment by means of confidence intervals.* International journal of clinical pharmacology, therapy, and toxicology, 1991. **29**(1): p. 1-8.

35.   Lopes, J.A., et al., *Chemometrics in bioprocess engineering: process analytical technology (PAT) applications.* Chemometrics and Intelligent Laboratory Systems, 2004. **74**(2): p. 269-275.

36.   De Beer, T., et al., *In-line and real-time process monitoring of a freeze drying process using Raman and NIR spectroscopy as complementary process analytical technology (PAT) tools.* Journal of pharmaceutical sciences, 2009. **98**(9): p. 3430-3446.

37.   Misra, N., C. Sullivan, and P. Cullen, *Process Analytical Technology (PAT) and Multivariate Methods for Downstream Processes.* Current Biochemical Engineering, 2015. **2**(1): p. 4-16.

38.   Rajalahti, T. and O.M. Kvalheim, *Multivariate data analysis in pharmaceutics: a tutorial review.* International journal of pharmaceutics, 2011. **417**(1): p. 280-290.

39.   Dunn, O.J. and V.A. Clark, *Applied statistics: analysis of variance and regression.* 1974.

40.   Kotsiantis, S.B., I. Zaharakis, and P. Pintelas, *Supervised machine learning: A review of classification techniques.* 2007.

41.   Sullivan, L.M., K.A. Dukes, and E. Losina, *Tutorial in biostatistics. An introduction to hierarchical linear modelling.* Stat Med, 1999. **18**(7): p. 855-888.

42.   Montgomery, D.C., *Design and analysis of experiments.* 2008: John Wiley & Sons.

43.   St, L. and S. Wold, *Analysis of variance (ANOVA).* Chemometrics and intelligent laboratory systems, 1989. **6**(4): p. 259-272.

44. Nakagawa, S. and I.C. Cuthill, *Effect size, confidence interval and statistical significance: a practical guide for biologists.* Biological Reviews, 2007. **82**(4): p. 591-605.

45. Olejnik, S. and J. Algina, *Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs.* Psychological Methods, 2003. **8**(4): p. 434-447.

46. Geoffrey Keppel, T.D.W., *Effect size, power, and sample size*, in *Design and Analysis: A Researcher's Handbook*. 2004, Prentice Hall.

47. Rouanet, H. and D. Lepine, *Comparison between treatments in a repeated-measurement design: ANOVA and multivariate methods.* British Journal of Mathematical and Statistical Psychology, 1970. **23**(2): p. 147-163.

48. Baayen, R.H., D.J. Davidson, and D.M. Bates, *Mixed-effects modeling with crossed random effects for subjects and items.* Journal of memory and language, 2008. **59**(4): p. 390-412.

49. Raudenbush, S.W., *Hierarchical Linear Models and Experimental Design*, in *Applied Analysis of Variance in Behavioral Science*, L. Edwards, Editor. 1993, CRC Press.

50. Timm, N.H., *Multivariate Analysis of Variance of Repeated Measurements*, in *Handbook of Statistics*, P.R. Krishnaiah, Editor. 1980.

51. Weinfurt, K.P., *Repeated measures analysis: ANOVA, MANOVA, and HLM*, in *Reading and understanding MORE multivariate statistics*, L.G.G.P.R. Yarnold, Editor. 2000, American Psychological Association: Washington, DC, US. p. 317-361.

52. O'Brien, R.G. and M.K. Kaiser, *MANOVA method for analyzing repeated measures designs: An extensive primer*. 1985, American Psychological Association: US. p. 316-333.

53. Ryan, G.W. and H.R. Bernard, *Techniques to identify themes.* Field methods, 2003. **15**(1): p. 85-109.

54. Yang, J., et al., *Two-dimensional PCA: a new approach to appearance-based face representation and recognition.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2004. **26**(1): p. 131-137.

55. Yeung, K.Y. and W.L. Ruzzo, *Principal component analysis for clustering gene expression data.* Bioinformatics, 2001. **17**(9): p. 763-774.

56. Smith, L.I., *A tutorial on principal components analysis.* Cornell University, USA, 2002. **51**(52): p. 65.

57. Wold, S., K. Esbensen, and P. Geladi, *Principal Component Analysis.* Chemometrics and Intelligent Laboratory Systems, 1987. **2**(1-3): p. 37-52.

58. Snee, R.D., *On the Analysis of Response Curve Data.* Technometrics, 1972. **14**(1): p. 47-62.

59. Gollob, H.F., *Confounding of sources of variation in factor-analytic techniques.* Psychological Bulletin, 1968. **70**(5): p. 330-344.

60. Fayyad, U., G. Piatetsky-Shapiro, and P. Smyth, *From data mining to knowledge discovery in databases.* AI magazine, 1996. **17**(3): p. 37.

61. Geladi, P. and B.R. Kowalski, *Partial least-squares regression: a tutorial.* Analytica chimica acta, 1986. **185**: p. 1-17.

62. Aiken, L.S., S.G. West, and S.C. Pitts, *Multiple linear regression.* Handbook of psychology, 2003.

63. Næs, T. and H. Martens, *Principal component regression in NIR analysis: viewpoints, background details and selection of components.* Journal of Chemometrics, 1988. **2**(2): p. 155-167.

64. Aguilera, A.M., M. Escabias, and M.J. Valderrama, *Using principal components for estimating logistic regression with high-dimensional multicollinear data.* Computational Statistics & Data Analysis, 2006. **50**(8): p. 1905-1924.

65.   Lindberg, W., J.-Å. Persson, and S. Wold, *Partial least-squares method for spectrofluorimetric analysis of mixtures of humic acid and lignin sulfonate.* Analytical Chemistry, 1983. **55**(4): p. 643-648.

66.   Dijkstra, T., *Some comments on maximum likelihood and partial least squares methods.* Journal of Econometrics, 1983. **22**(1): p. 67-90.

67.   Godoy, J.L., J.R. Vega, and J.L. Marchetti, *Relationships between PCA and PLS-regression.* Chemometrics and Intelligent Laboratory Systems, 2014. **130**: p. 182-191.

68.   Maitra, S. and J. Yan, *Principle component analysis and partial least squares: Two dimension reduction techniques for regression.* Applying Multivariate Statistical Models, 2008. **79**.

69.   Yaroshchyk, P., D.L. Death, and S.J. Spencer, *Comparison of principal components regression, partial least squares regression, multi-block partial least squares regression, and serial partial least squares regression algorithms for the analysis of Fe in iron ore using LIBS.* Journal of Analytical Atomic Spectrometry, 2012. **27**(1): p. 92-98.

70.   Bylesjö, M., et al., *OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification.* Journal of Chemometrics, 2006. **20**(8-10): p. 341-351.

71.   Fujiwara, M., et al., *Paracetamol crystallization using laser backscattering and ATR-FTIR spectroscopy: metastability, agglomeration, and control.* Crystal Growth & Design, 2002. **2**(5): p. 363-370.

72.   Lewiner, F., et al., *On-line ATR FTIR measurement of supersaturation during solution crystallization processes. Calibration and applications on three solute/solvent systems.* Chemical Engineering Science, 2001. **56**(6): p. 2069-2084.

73.   Barrett, P. and B. Glennon, *Characterizing the metastable zone width and solubility curve using Lasentec FBRM and PVM.* Chemical Engineering Research and Design, 2002. **80**(7): p. 799-805.

74.   Sulub, Y., et al., *Real-time on-line blend uniformity monitoring using near-infrared reflectance spectrometry: a noninvasive off-line calibration approach.* Journal of pharmaceutical and biomedical analysis, 2009. **49**(1): p. 48-54.

75.   Ely, D., S. Chamarthy, and M.T. Carvajal, *An investigation into low dose blend uniformity and segregation determination using NIR spectroscopy.* Colloids and Surfaces A: Physicochemical and Engineering Aspects, 2006. **288**(1): p. 71-76.

76.   Sulub, Y., M. Konigsberger, and J. Cheney, *Blend uniformity end-point determination using near-infrared spectroscopy and multivariate calibration.* Journal of pharmaceutical and biomedical analysis, 2011. **55**(3): p. 429-434.

77.   Wikström, H., P.J. Marsac, and L.S. Taylor, *In-line monitoring of hydrate formation during wet granulation using Raman spectroscopy.* Journal of pharmaceutical sciences, 2005. **94**(1): p. 209-219.

78.   Alcalà, M., et al., *On-line monitoring of a granulation process by NIR spectroscopy.* Journal of pharmaceutical sciences, 2010. **99**(1): p. 336-345.

79.   Jørgensen, A., et al., *Hydrate formation during wet granulation studied by spectroscopic methods and multivariate analysis.* Pharmaceutical research, 2002. **19**(9): p. 1285-1291.

80.   Moes, J.J., et al., *Application of process analytical technology in tablet process development using NIR spectroscopy: Blend uniformity, content uniformity and coating thickness measurements.* International journal of pharmaceutics, 2008. **357**(1): p. 108-118.

81.   Freitas, M.P., et al., *Prediction of drug dissolution profiles from tablets using NIR diffuse reflectance spectroscopy: a rapid and nondestructive method.* Journal of pharmaceutical and biomedical analysis, 2005. **39**(1): p. 17-21.

82.   Hernandez, E., et al., *Prediction of dissolution profiles by non-destructive near infrared spectroscopy in tablets subjected to different levels of strain.* Journal of pharmaceutical and biomedical analysis, 2016. **117**: p. 568-576.

83. Gupta, A., et al., *Real-time near-infrared monitoring of content uniformity, moisture content, compact density, tensile strength, and young's modulus of roller compacted powder blends.* Journal of pharmaceutical sciences, 2005. **94**(7): p. 1589-1597.

84. Andersson, M., et al., *Monitoring of a film coating process for tablets using near infrared reflectance spectrometry.* Journal of pharmaceutical and biomedical analysis, 1999. **20**(1): p. 27-37.

85. Fitzgerald, A.J., B.E. Cole, and P.F. Taday, *Nondestructive analysis of tablet coating thicknesses using terahertz pulsed imaging.* Journal of pharmaceutical sciences, 2005. **94**(1): p. 177-183.

86. De Smet, K., et al., *Selectivity control by use of near-IR for a hydrogenation process.* Organic process research & development, 2005. **9**(3): p. 344-347.

87. Sena, M.M., et al., *Direct determination of diclofenac in pharmaceutical formulations containing B vitamins by using UV spectrophotometry and partial least squares regression.* Journal of pharmaceutical and biomedical analysis, 2004. **36**(4): p. 743-749.

88. Yu, W., et al., *Prediction of bulk powder flow performance using comprehensive particle size and particle shape distributions.* Journal of pharmaceutical sciences, 2011. **100**(1): p. 284-293.

89. Ge, Z. and Z. Song, *Process monitoring based on independent component analysis-principal component analysis (ICA-PCA) and similarity factors.* Industrial & Engineering Chemistry Research, 2007. **46**(7): p. 2054-2063.

90. Chen, J. and K.-C. Liu, *On-line batch process monitoring using dynamic PCA and dynamic PLS models.* Chemical Engineering Science, 2002. **57**(1): p. 63-75.

91. Gunther, J.C., J.S. Conner, and D.E. Seborg, *Fault Detection and Diagnosis in an Industrial Fed-Batch Cell Culture Process.* Biotechnology progress, 2007. **23**(4): p. 851-857.

92. Kirdar, A.O., et al., *Application of Multivariate Analysis toward Biotech Processes: Case Study of a Cell-Culture Unit Operation.* Biotechnology progress, 2007. **23**(1): p. 61-67.

93. MacGregor, J.F., et al., *Process monitoring and diagnosis by multiblock PLS methods.* AIChE Journal, 1994. **40**(5): p. 826-838.

94. Lopes, J., et al., *Multiblock PLS analysis of an industrial pharmaceutical process.* Biotechnology and bioengineering, 2002. **80**(4): p. 419-427.

95. Lekhal, A., et al., *Characterization of granular flow of wet solids in a bladed mixer.* AIChE Journal, 2006. **52**(8): p. 2757-2766.

96. Pingali, K.C. and R. Mendez, *Physicochemical behavior of pharmaceutical particles and distribution of additives in tablets due to process shear and lubricant composition.* Powder Technology, 2014. **268**: p. 1-8.

97. Mullarney, M.P., et al., *Applying dry powder coatings to pharmaceutical powders using a comil for improving powder flow and bulk density.* Powder Technology, 2011. **212**(3): p. 397-402.

98. Muzzio, F., *Solids mixing*, in *Handbook of Industrial Mixing*. 2004, John Wiley Sons Inc. p. 887-985

99. van Ommen, J.R., J.M. Valverde, and R. Pfeffer, *Fluidization of nanopowders: a review.* Journal of Nanoparticle Research, 2012. **14**(3): p. 1-29.

100. Barling, D., D.A. Morton, and K. Hapgood, *Pharmaceutical dry powder blending and scale-up: Maintaining equivalent mixing conditions using a coloured tracer powder.* Powder Technology, 2015. **270**: p. 461-469.

101. Shenoy, P., et al., *Effect of powder densities, particle size and shape on mixture quality of binary food powder mixtures.* Powder Technology, 2015. **272**: p. 165-172.

102. Dave, R.N., et al., *Magnetically mediated flow enhancement for controlled powder discharge of cohesive powders.* Powder technology, 2000. **112**(1): p. 111-125.

103. Jallo, L.J., et al., *Prediction of inter-particle adhesion force from surface energy and surface roughness.* Journal of Adhesion Science and Technology, 2011. **25**(4-5): p. 367-384.

104. Prescott, J.K. and R.A. Barnum, *On powder flowability.* Pharmaceutical technology, 2000. **24**(10): p. 60-85.

105. Pierrat, P. and H.S. Caram, *Tensile strength of wet granula materials.* Powder Technology, 1997. **91**(2): p. 83-93.

106. Goldfarb, D., H. Nakagawa, and S. Conway, *Component morphology, size, and compositional impact on pharmaceutical powder blend flowability.* Bulletin of the American Physical Society, 2014. **59**.

107. Poux, M., et al., *Powder mixing: some practical rules applied to agitated systems.* Powder Technology, 1991. **68**(3): p. 213-234.

108. Schwedes, J., *Review on testers for measuring flow properties of bulk solids.* Granular matter, 2003. **5**(1): p. 1-43.

109. Vasilenko, A., B.J. Glasser, and F.J. Muzzio, *Shear and flow behavior of pharmaceutical blends—Method comparison study.* Powder Technology, 2011. **208**(3): p. 628-636.

110. Lindberg, N.O., et al., *Flowability measurements of pharmaceutical powder mixtures with poor flow using five different techniques.* Drug development and industrial pharmacy, 2004. **30**(7): p. 785-791.

111. Pingali, K.C., *Consolidated states of dilation in granular material: Experimental investigation and comparison of rheological properties.* Advanced Powder Technology, 2015.

112. Koynov, S., et al., *Measurement of the axial dispersion coefficient of powders in a rotating cylinder: Dependence on bulk flow properties.* Powder Technology, 2016.

113. Jenike, A.W., *Storage and flow of solids, bulletin no. 123.* Bulletin of the University of Utah, 1964. **53**(26).

114. Schulze, D., *Powders and bulk solids*, in *Behaviour, Characterization, Storage and Flow. Springer*. 2008, Springer. p. 35-74.

115. Ganesan, V., K. Muthukumarappan, and K.A. Rosentrater, *Flow properties of DDGS with varying soluble and moisture contents using jenike shear testing.* Powder technology, 2008. **187**(2): p. 130-137.

116. Carson, J.W. and H. Wilms, *Development of an international standard for shear testing.* Powder technology, 2006. **167**(1): p. 1-9.

117. Standard, A., *D6773-02: Standard shear test method for bulk solids using the Schulze ring shear tester, ASTM International.*

118. Freeman, R., J. Cooke, and L. Schneider, *Measuring shear properties and normal stresses generated within a rotational shear cell for consolidated and non-consolidated powders.* Powder Technology, 2009. **190**(1): p. 65-69.

119. Vasilenko, A., et al., *Role of consolidation state in the measurement of bulk density and cohesion.* Powder Technology, 2013. **239**: p. 366-373.

120. Podczeck, F. and Y. Mia, *The influence of particle size and shape on the angle of internal friction and the flow factor of unlubricated and lubricated powders.* International Journal of Pharmaceutics, 1996. **144**(2): p. 187-194.

121. Hou, H. and C.C. Sun, *Quantifying effects of particulate properties on powder flow properties using a ring shear tester.* Journal of pharmaceutical sciences, 2008. **97**(9): p. 4030-4039.

122. Teunou, E. and J. Fitzpatrick, *Effect of storage time and consolidation on food powder flowability.* Journal of Food Engineering, 2000. **43**(2): p. 97-101.

123. Koynov, S., B. Glasser, and F. Muzzio, *Comparison of three rotational shear cell testers: powder flowability and bulk density.* Powder Technology, 2015.

124.   Taylor, M.K., et al., *Composite method to quantify powder flow as a screening method in early tablet or capsule formulation development.* AAPS PharmSciTech, 2000. **1**(3): p. 20-30.

125.   Schwedes, J., *Consolidation and flow of cohesive bulk solids.* Chemical Engineering Science, 2002. **57**(2): p. 287-294.

126.   Vegter, H. and A. Van den Boogaard, *A plane stress yield function for anisotropic sheet material by interpolation of biaxial stress states.* International Journal of Plasticity, 2006. **22**(3): p. 557-580.

127.   Falticeanu, C., et al., *New methods for assesing the frictional properties in a mass of consolidated powder.* In Proceding of Balkantrib, 2005. **5**: p. 132.

128.   Schulze, D., *Powders and bulk solids: behavior, characterization, storage and flow*. 2007: Springer.

129.   14, A.D.-. *Standard Test Method for Shear Testing of Bulk Solids Using the Jenike Shear Cell*. 2014.

130.   Wang, Y., et al., *Predicting flow behavior of pharmaceutical blends using shear cell methodology: A quality by design approach.* Powder Technology, 2016.

131.   Feise, H.J., *A review of induced anisotropy and steady-state flow in powders.* Powder technology, 1998. **98**(3): p. 191-200.

132.   Campbell, C.S., *Granular material flows–an overview.* Powder Technology, 2006. **162**(3): p. 208-229.

133.   Remy, B., B.J. Glasser, and J.G. Khinast, *The effect of mixer properties and fill level on granular flow in a bladed mixer.* AIChE journal, 2010. **56**(2): p. 336-353.

134.   Remy, B., et al., *Scale-up of agitated drying: Effect of shear stress and hydrostatic pressure on active pharmaceutical ingredient powder properties.* AIChE Journal, 2015. **61**(2): p. 407-418.

135.   Çağli, A.S., et al. *Flow property measurement using Jenike shear cell for 7 different bulk solids*. in *Proceedings of European Congress of Chemical Engineering (ECCE-6), Copenhagen*. 2007.

136.   Campbell, C.S., *Granular shear flows at the elastic limit.* Journal of fluid mechanics, 2002. **465**: p. 261-291.

137.   Jenck, J.F., F. Agterberg, and M.J. Droescher, *Products and processes for a sustainable chemical industry: a review of achievements and prospects.* Green Chemistry, 2004. **6**(11): p. 544-556.

138.   Noureddini, H., D. Harkey, and V. Medikonduru, *A continuous process for the conversion of vegetable oils into methyl esters of fatty acids.* Journal of the American Oil Chemists' Society, 1998. **75**(12): p. 1775-1783.

139.   Weinekötter, R. and L. Reh, *Continuous mixing of fine particles.* Particle & particle systems characterization, 1995. **12**(1): p. 46-53.

140.   Engisch, W., M. Ierapetritou, and F. Muzzio. *Hopper refill of loss-in-weight feeding equipment*. in *Proceedings of the 2010 AIChE Annual Meeting, Salt Lake City, UT, USA*. 2010.

141.   Johnson, N.R., *System for precisely controlling discharge rates of loss-in-weight feeder systems*. 1992, Google Patents.

142.   Hachtel, R., *Loss-in-weight feeder with discharge pressure compensator*. 2002, Google Patents.

143.   Aalto, P. and J.-P. Björklund, *Loss-in-weight feeder control*. 2002, Google Patents.

144.   Hopkins, M., *LOSS in weight feeder systems.* Measurement and Control, 2006. **39**(8): p. 237-240.

145.   Engisch, W.E. and F.J. Muzzio, *Method for characterization of loss-in-weight feeder equipment.* Powder technology, 2012. **228**: p. 395-403.

146. Engisch, W.E. and F.J. Muzzio, *Feedrate deviations caused by hopper refill of loss-in-weight feeders.* Powder Technology, 2015. **283**: p. 389-400.

147. Kehlenbeck, V. and K. Sommer, *Possibilities to improve the short-term dosing constancy of volumetric feeders.* Powder technology, 2003. **138**(1): p. 51-56.

148. Tardos, G.I. and Q. Lu, *Precision dosing of powders by vibratory and screw feeders: an experimental study.* Advanced Powder Technology, 1996. **7**(1): p. 51-58.

149. Royal, T.A. and J.W. Carson, *Fine powder flow phenomena in bins, hoppers and processing vessels*. 2000, Bulk.

150. Engisch, W. and F. Muzzio, *Using Residence Time Distributions (RTDs) to Address the Traceability of Raw Materials in Continuous Pharmaceutical Manufacturing.* Journal of Pharmaceutical Innovation, 2015: p. 1-18.

151. Fassihi, A. and I. Kanfer, *Effect of compressibility and powder flow properties on tablet weight variation.* Drug Development and Industrial Pharmacy, 1986. **12**(11-13): p. 1947-1966.

152. Owen, P. and P. Cleary, *Prediction of screw conveyor performance using the Discrete Element Method (DEM).* Powder Technology, 2009. **193**(3): p. 274-288.

153. Tim Freeman, D.M.-S., Ralf Weinekotter, *Predicting Feeder Performance from Powder Flow Measurements*, in *Powder and Bulk Solids*. 2015.

154. Koynov, S. and F.J. Muzzio, *A Quantitative Approach to Understand Raw Material Variability*, in *Process Simulation and Data Modeling in Solid Oral Drug Development and Manufacture*, R.R. Marianthi G Ierapetritou, Editor. 2016, Humana Press. p. 85-104.

155. Koynov, S., *PhD thesis. Using statistical methods to optimize powder flow measurements and to predict powder processing performance*. 2015, Rutgers University-Graduate School-New Brunswick.

156. Jain, A., K. Nandakumar, and A. Ross, *Score normalization in multimodal biometric systems.* Pattern recognition, 2005. **38**(12): p. 2270-2285.

157. Al Shalabi, L. and Z. Shaaban. *Normalization as a preprocessing engine for data mining and the approach of preference matrix*. in *Dependability of Computer Systems, 2006. DepCos-RELCOMEX'06. International Conference on*. 2006. IEEE.

158. Esbensen, K., et al., *Multivariate Analysis in Practice: A Training Package*. 1996: Camo As.

159. Ferreira, A.P., et al., *Use of similarity scoring in the development of oral solid dosage forms.* International journal of pharmaceutics, 2015.

160. Naes, T., et al., *A user friendly guide to multivariate calibration and classification*. 2002: NIR publications.

161. Wang, Y., et al., *A method to analyze shear cell data of powders measured under different initial consolidation stresses.* Powder Technology, 2016. **294**: p. 105-112.

162. Martens, H. and M. Martens, *Multivariate analysis of quality. An introduction*. 2001, IOP Publishing.

163. Engisch Jr, W.E., *PhD thesis. Loss-in-weight feeding in continuous powder manufacturing*. 2014: Rutgers University - Graduate School - New Brunswick.

164. Shinbrot, T., K. LaMarche, and B.J. Glasser, *Triboelectrification and razorbacks: geophysical patterns produced in dry grains.* Physical review letters, 2006. **96**(17): p. 178002.

165. LaMarche, K.R., et al., *Granular flow and dielectrophoresis: The effect of electrostatic forces on adhesion and flow of dielectric granular materials.* Powder Technology, 2010. **199**(2): p. 180-188.

166. Vasilenko, A.V., *Rheological Properties of Granular Materials-Critical Parameters And Mixing Rules*, in *Chemical and Biochemical Department*. 2011, Rugers, the State University of New Jersey.

167. Rios, M., *Developments in powder flow testing*. 2006.

168. Jenike, A.W., *Gravity flow of bulk solids.* Bulletin No. 108, Utah State University, 1961.

169. Fitzpatrick, J., S. Barringer, and T. Iqbal, *Flow property measurement of food powders and sensitivity of Jenike's hopper design methodology to the measured values.* Journal of Food Engineering, 2004. **61**(3): p. 399-405.

170. Teunou, E. and J. Fitzpatrick, *Effect of relative humidity and temperature on food powder flowability.* Journal of Food Engineering, 1999. **42**(2): p. 109-116.

171. Sandler, N. and D. Wilson, *Prediction of granule packing and flow behavior based on particle size and shape analysis.* Journal of pharmaceutical sciences, 2010. **99**(2): p. 958-968.

172. Cornell, J.A., *Experiments with mixtures: designs, models, and the analysis of mixture data.* Vol. 895. 2011: John Wiley & Sons.

173. Osorio, J., K. Sowrirajan, and F. Muzzio. *Effect of resonant acoustic mixing on pharmaceutical powder blends and tablets.* in *Proceedings of the Society of Engineering Science 51st Annual Technical Meeting.* 2014. West Lafayette: Purdue University Libraries Scholarly Publishing Services.

174. Vanarase A, O.J., Muzzio F, Coguill S, Lucon P. *RESONANTACOUSTIC® MIXING; UNIFORM DISTRIBUTION OF MINOR MATERIALS DURING POWDER MIXING.* in *Proceedings of the JANNAF 36th Propellant and Explosives Development and Characterization Joint Subcommittee Meeting.* 2010.

175. Schulze, D., *Flow properties of powders and bulk solids.* Braunschweig/Wolfenbu ttel, Germany: University of Applied Sciences, 2006.

176. Koynov, S., B. Glasser, and F. Muzzio, *Comparison of three rotational shear cell testers: powder flowability and bulk density.* Powder Technology, 2015. **283**: p. 103-112.

177. Navaneethan, C.V., S. Missaghi, and R. Fassihi, *Application of powder rheometer to determine powder flow properties and lubrication efficiency of pharmaceutical particulate systems.* Aaps Pharmscitech, 2005. **6**(3): p. E398-E404.

178. Mackin, L., et al., *The impact of low levels of amorphous material (< 5%) on the blending characteristics of a direct compression formulation.* International journal of pharmaceutics, 2002. **231**(2): p. 213-226.

179. Zhou, Q.T., et al., *Characterization of the surface properties of a model pharmaceutical fine powder modified with a pharmaceutical lubricant to improve flow via a mechanical dry coating approach.* Journal of pharmaceutical sciences, 2011. **100**(8): p. 3421-3430.

180. Jonat, S., et al., *Investigation of compacted hydrophilic and hydrophobic colloidal silicon dioxides as glidants for pharmaceutical excipients.* Powder technology, 2004. **141**(1): p. 31-43.

181. Ebba, F., et al., *Stress relaxation studies of granules as a function of different lubricants.* European journal of pharmaceutics and biopharmaceutics, 2001. **52**(2): p. 211-220.

182. Vromans, H. and C. Lerk, *Densification properties and compactibility of mixtures of pharmaceutical excipients with and without magnesium stearate.* International journal of pharmaceutics, 1988. **46**(3): p. 183-192.

183. Johansson, M.E., *Granular magnesium stearate as a lubricant in tablet formulations.* International journal of pharmaceutics, 1984. **21**(3): p. 307-315.

184. Freeman, T., et al., *Capsule filling performance of powdered formulations in relation to flow characteristics.* 2011.

185. Mehrotra, A., et al., *Influence of shear intensity and total shear on properties of blends and tablets of lactose and cellulose lubricated with magnesium stearate.* International journal of pharmaceutics, 2007. **336**(2): p. 284-291.

186. Llusa, M. and F. Muzzio, *The effect of shear mixing on the blending of cohesive lubricants and drugs.* 2005.

187. Fukui, E., N. Miyamura, and M. Kobayashi, *Effect of magnesium stearate or calcium stearate as additives on dissolution profiles of diltiazem hydrochloride from press-coated*

tablets with hydroxypropylmethylcellulose acetate succinate in the outer shell. International journal of pharmaceutics, 2001. **216**(1): p. 137-146.

188. Van Veen, B., et al., *Compaction mechanism and tablet strength of unlubricated and lubricated (silicified) microcrystalline cellulose.* European journal of Pharmaceutics and Biopharmaceutics, 2005. **59**(1): p. 133-138.

189. Zuurman, K., K. Van der Voort Maarschalk, and G. Bolhuis, *Effect of magnesium stearate on bonding and porosity expansion of tablets produced from materials with different consolidation properties.* International journal of pharmaceutics, 1999. **179**(1): p. 107-115.

190. Mendez, R., F.J. Muzzio, and C. Velazquez, *Powder hydrophobicity and flow properties: effect of feed frame design and operating parameters.* AIChE Journal, 2012. **58**(3): p. 697-706.

191. Ragnarsson, G., A. Hölzer, and J. Sjögren, *The influence of mixing time and colloidal silica on the lubricating properties of magnesium stearate.* International Journal of Pharmaceutics, 1979. **3**(2): p. 127-131.

192. Portillo, P.M., M.G. Ierapetritou, and F.J. Muzzio, *Characterization of continuous convective powder mixing processes.* Powder Technology, 2008. **182**(3): p. 368-378.

193. Portillo, P.M., M.G. Ierapetritou, and F.J. Muzzio, *Effects of rotation rate, mixing angle, and cohesion in two continuous powder mixers—A statistical approach.* Powder Technology, 2009. **194**(3): p. 217-227.

194. Osorio, J., K. Sowrirajan, and F. Muzzio, *Effect of resonant acoustic mixing on pharmaceutical powder blends and tablets.* 2014.

195. Osorio, J.G. and F.J. Muzzio, *Evaluation of resonant acoustic mixing performance.* Powder Technology, 2015. **278**: p. 46-56.

196. Osorio Caicedo, J.G., *Macro and micro characterization of powder mixing processes.* 2014, Rutgers University-Graduate School-New Brunswick.

197. D7481-09, A., *Standard Test Methods for Determining Loose and Tapped Bulk Densities of Powders using a Graduated Cylinder.* 2009.

198. Abdullah, E. and D. Geldart, *The use of bulk density measurements as flowability indicators.* Powder technology, 1999. **102**(2): p. 151-165.

199. Pingali, K., et al., *Mixing order of glidant and lubricant–influence on powder and tablet properties.* International journal of pharmaceutics, 2011. **409**(1): p. 269-277.

200. Zhou, Q., et al., *Improving powder flow properties of a cohesive lactose monohydrate powder by intensive mechanical dry coating.* Journal of pharmaceutical sciences, 2010. **99**(2): p. 969-981.

201. Wornyoh, E.Y., V.K. Jasti, and C.F. Higgs, *A review of dry particulate lubrication: powder and granular materials.* Journal of Tribology, 2007. **129**(2): p. 438-449.

202. Godet, M., et al., *Wear Modeling: How Far Can We Get with First Principles?*, in *Tribological Modeling for Mechanical Designers.* 1991, ASTM International.

203. Wang, Y., et al., *Statistical comparison of dissolution profiles.* Drug Development and Industrial Pharmacy, 2015: p. 1-12.

204. Brone, D., A. Alexander, and F. Muzzio, *Quantitative characterization of mixing of dry powders in V-blenders.* American Institute of Chemical Engineers. AIChE Journal, 1998. **44**(2): p. 271.

205. Rasenack, N. and B.W. Müller, *Dissolution rate enhancement by in situ micronization of poorly water-soluble drugs.* Pharmaceutical research, 2002. **19**(12): p. 1894-1900.

206. Rasenack, N., H. Hartenhauer, and B.W. Müller, *Microcrystals for dissolution rate enhancement of poorly water-soluble drugs.* International journal of pharmaceutics, 2003. **254**(2): p. 137-145.

207. Yuksel, N., A.E. Kanik, and T. Baykara, *Comparison of in vitro dissolution profiles by ANOVA-based, model-dependent and -independent methods.* International Journal of Pharmaceutics, 2000. **209**(1-2): p. 57-67.

208. Van Vooren, L., et al., *A Novel Bending Point Criterion for Dissolution Profile Interpretation.* Drug Development and Industrial Pharmacy, 2001. **27**(8): p. 885-892.

209. J. W. Moore, H.H.F., *Mathematical comparison of curves with an emphasis on in-vitro dissolution profiles.* Pharmaceutical Technology, 1996. **20**(6): p. 64-74.

210. FDA, *Guidance for Industry: Modified Release Solid Oral Dosage Forms: Scale-up and Post Approval Changes (SUPAC-MR): Chemistry, Manufacturing and Controls; InVitro Dissolution Testing and In Vivo Bioequivalence Documentation.* 1997.

211. FDA, *Guidance for Industry: Extended Release Oral Dosage Forms- Development, Evaluation, and Application of In Vitro/In Vivo Correlations.* 1997.

212. FDA, *Guidance for Industry: Dissolution Testing of Immediate Release Solid Oral Dosage Forms.* 1997.

213. Peh, K.K. and C.F. Wong, *Application of Similarity Factor in Development of Controlled-Release Diltiazem Tablet.* Drug Development and Industrial Pharmacy, 2000. **26**(7): p. 723-730.

214. Williams, R.O., et al., *Investigation of Excipient Type and Level on Drug Release from Controlled Release Tablets Containing HPMC.* Pharmaceutical Development and Technology, 2002. **7**(2): p. 181-193.

215. Zhang, X., X. Tang, and R. Yang, *Development of a Tamsulosin Hydrochloride Controlled-Release Capsule Consisting of Two Different Coated Pellets.* Drug Development and Industrial Pharmacy, 2009. **35**(1): p. 26-33.

216. Kim H. Esbensen, a.P.P.-M., *Process Sampling: Theory of Sampling - the Missing Link in Process Analytical Technologies (PAT)*, in *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries, 2nd Edition*, K.A. Bakeev, Editor. 2010.

217. Stevens, R.E., et al., *Scientific and Regulatory Standards for Assessing Product Performance Using the Similarity Factor, f2.* The AAPS Journal, 2015. **17**(2): p. 301-306.

218. Maggio, R.M., M.A. Rivero, and T.S. Kaufman, *Simultaneous acquisition of the dissolution curves of two active ingredients in a binary pharmaceutical association, employing an on-line circulation system and chemometrics-assistance.* Journal of pharmaceutical and biomedical analysis, 2013. **72**: p. 51-58.

219. Montgomery, D.C., *Design and Analysis of Experiments, 8th Edition*. 2012.

220. Liu, J.-p., M.-C. Ma, and S.-C. Chow, *Statistical Evaluation of Similarity Factor f2 as a Criterion for Assessment of Similarity between Dissolution Profiles.* Drug Information Journal, 1997. **31**(4): p. 1255-1271.

221. Hsu, J.Y., et al., *On the characteristics of the FDA's similarity factor for comparison of drug dissolution.* Journal of Food and Drug Analysis, 1998. **6**(3): p. 553-558.

222. Shah, V.P., et al., *In vitro dissolution profile comparison - Statistics and analysis of the similarity factor, f(2).* Pharmaceutical Research, 1998. **15**(6): p. 889-896.

223. Ma, M.C., R.P. Lin, and J.P. Liu, *Statistical evaluations of dissolution similarity.* Statistica Sinica, 1999. **9**(4): p. 1011-1027.

224. Ocana, J., G. Frutos, and P. Sanchez, *Using the similarity factor f(2) in practice: A critical revision and suggestions for its standard error estimation.* Chemometrics and Intelligent Laboratory Systems, 2009. **99**(1): p. 49-56.

225. Gohel, M.C. and M.K. Panchal, *Novel Use of Similarity Factors f2 and Sd for the Development of Diltiazem HCl Modified-Release Tablets Using a 32 Factorial Design.* Drug Development and Industrial Pharmacy, 2002. **28**(1): p. 77-87.

226. Mashru, R.C., et al., *Development and Evaluation of Fast-Dissolving Film of Salbutamol Sulphate.* Drug Development and Industrial Pharmacy, 2005. **31**(1): p. 25-34.

227. Seo, P.R., V.P. Shah, and J.E. Polli, *Novel Metrics to Compare Dissolution Profiles.* Pharmaceutical Development and Technology, 2002. **7**(2): p. 257-265.

228. Saranadasa, H. and K. Krishnamoorthy, *A multivariate test for similarity of two dissolution profiles.* Journal of Biopharmaceutical Statistics, 2005. **15**(2): p. 265-278.

229. Maggio, R.M., P.M. Castellano, and T.S. Kaufman, *A new principal component analysis-based approach for testing "similarity" of drug dissolution profiles.* European Journal of Pharmaceutical Sciences, 2008. **34**(1): p. 66-77.

230. Maggio, R.M., P.M. Castellano, and T.S. Kaufman, *PCA-CR analysis of dissolution profiles. A chemometric approach to probe the polymorphic form of the active pharmaceutical ingredient in a drug product.* International journal of pharmaceutics, 2009. **378**(1): p. 187-193.

231. Tsong, Y., T. Hammerstrom, and J.J. Chen, *Multipoint dissolution specification and acceptance sampling rule based on profile modeling and principal component analysis.* Journal of Biopharmaceutical Statistics, 1997. **7**(3): p. 423-439.

232. Adams, E., et al., *Evaluation of dissolution profiles using principal component analysis.* International journal of pharmaceutics, 2001. **212**(1): p. 41-53.

233. Adams, E., et al., *Principal component analysis of dissolution data with missing elements.* International journal of pharmaceutics, 2002. **234**(1): p. 169-178.

234. Gómez-Mantilla, J.-D., et al., *Permutation Test (PT) and Tolerated Difference Test (TDT): Two new, robust and powerful nonparametric tests for statistical comparison of dissolution profiles.* International journal of pharmaceutics, 2013. **441**(1): p. 458-467.

235. Enachescu, D., *Exploratory Data Analysis Methods for Comparison of Drug Dissolution Profiles.* Biocybernetics and Biomedical Engineering, 2010. **30**(4): p. 17-27.

236. Chow, S.C. and J. Shao, *On the assessment of similarity for dissolution profiles of two drug products.* J Biopharm Stat, 2002. **12**(3): p. 311-21.

237. Lansky, P. and M. Weiss, *Classification of dissolution profiles in terms of fractional dissolution rate and a novel measure of heterogeneity.* J Pharm Sci, 2003. **92**(8): p. 1632-47.

238. Berry, M.R. and M.D. Likar, *Statistical assessment of dissolution and drug release profile similarity using a model-dependent approach.* Journal of Pharmaceutical and Biomedical Analysis, 2007. **45**(2): p. 194-200.

239. Mauger, J.W., D. Chilko, and S. Howard, *On the Analysis of Dissolution Data.* Drug Development and Industrial Pharmacy, 1986. **12**(7): p. 969-992.

240. J., C., *Statistical power analysis for the behavior ciences*. 1988: Lawrenced Erlbaum Associates, Inc. 283-286.

241. Hsu, J.C. and M. Peruggia, *Graphical Representations of Tukey's Multiple Comparison Method.* Journal of Computational and Graphical Statistics, 1994. **3**(2): p. 143-161.

242. Mascia, S., et al., *End-to-End Continuous Manufacturing of Pharmaceuticals: Integrated Synthesis, Purification, and Final Dosage Formation.* Angewandte Chemie International Edition, 2013. **52**(47): p. 12359-12363.

243. Spanos, C.J., et al., *Real-time statistical process control using tool data [semiconductor manufacturing].* Semiconductor Manufacturing, IEEE Transactions on, 1992. **5**(4): p. 308-318.

244. Montgomery, D.C., *Introduction to statistical quality control*. 2007: John Wiley & Sons.

245. Hinz, D.C., *Process analytical technologies in the pharmaceutical industry: the FDA's PAT initiative.* Analytical and bioanalytical chemistry, 2006. **384**(5): p. 1036-1042.

246. Dressman, J.B., et al., *Dissolution testing as a prognostic tool for oral drug absorption: immediate release dosage forms.* Pharmaceutical research, 1998. **15**(1): p. 11-22.

247. Uppoor, V.R.S., *Regulatory perspectives on in vitro (dissolution)/in vivo (bioavailability) correlations.* Journal of controlled release, 2001. **72**(1): p. 127-132.

248. Roggo, Y., et al., *Characterizing process effects on pharmaceutical solid forms using near-infrared spectroscopy and infrared imaging.* European Journal of Pharmaceutics and Biopharmaceutics, 2005. **61**(1): p. 100-110.

249. Otsuka, M., J.I. Gao, and Y. Matsuda, *Effects of mixer and mixing time on the pharmaceutical properties of theophylline tablets containing various kinds of lactose as diluents.* Drug development and industrial pharmacy, 1993. **19**(3): p. 333-348.

250. Abe, H. and M. Otsuka, *Effects of lubricant-mixing time on prolongation of dissolution time and its prediction by measuring near infrared spectra from tablets.* Drug development and industrial pharmacy, 2012. **38**(4): p. 412-419.

251. Spencer, J.A., et al., *Delayed release tablet dissolution related to coating thickness by terahertz pulsed image mapping.* Journal of pharmaceutical sciences, 2008. **97**(4): p. 1543-1550.

252. Gao, Z., et al., *Gauge repeatability and reproducibility for accessing variability during dissolution testing: a Technical Note.* AAPS PharmSciTech, 2007. **8**(4): p. 11-15.

253. Zannikos, P.N., et al., *Spectrophotometric prediction of the dissolution rate of carbamazepine tablets.* Pharmaceutical research, 1991. **8**(8): p. 974-978.

254. Blanco, M., et al., *Determination of dissolution profiles in intact pharmaceutical tablets by NIR spectroscopy.* J. Process Anal. Technol, 2006. **3**(5): p. 25-28.

255. Tabasi, S.H., et al., *Sustained release dosage forms dissolution behavior prediction: a study of matrix tablets using NIR spectroscopy.* International journal of pharmaceutics, 2009. **382**(1): p. 1-6.

256. Basalious, E.B., W. El-Sebaie, and O. El-Gazayerly, *Application of pharmaceutical QbD for enhancement of the solubility and dissolution of a class II BCS drug using polymeric surfactants and crystallization inhibitors: development of controlled-release tablets.* Aaps Pharmscitech, 2011. **12**(3): p. 799-810.

257. Beebe, K.R., R.J. Pell, and M.B. Seasholtz, *Chemometrics: a practical guide.* Vol. 4. 1998: Wiley-Interscience.

258. Savitzky, A. and M.J. Golay, *Smoothing and differentiation of data by simplified least squares procedures.* Analytical chemistry, 1964. **36**(8): p. 1627-1639.

259. Pingali, K., et al., *Evaluation of strain-induced hydrophobicity of pharmaceutical blends and its effect on drug release rate under multiple compression conditions.* Drug development and industrial pharmacy, 2011. **37**(4): p. 428-435.