

Development of a Virtual Sensor for Real-Time Prediction of Granule Flow Properties

Rexonni B. Lagare,^a Mariana Araujo da Conceicao,^a Ariana Camille Acevedo Rosario,^a Katherine Leigh Young,^a Yan-Shu Huang,^a M. Ziyen Sheriff,^a Clairmont Clementson,^b Paul Mort,^c Zoltan Nagy,^a Gintaras V. Reklaitis^{a b}

^a*Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN 47907, USA*

^b*Purdue Center for Particulate Products and Processes, West Lafayette, IN 47907, USA*

^c*School of Materials Engineering, Purdue University, West Lafayette, IN 47907, USA*

**rlagare@purdue.edu*

Abstract

We report progress of an ongoing work to develop a virtual sensor for flowability, which is a critical tool for enabling real time process monitoring in a granulation line. The sensor is based on camera imaging to measure the size and shape distribution of granules produced by wet granulation. Then, statistical methods were used to correlate them with flowability measurements such as ring shear tests, drained angle of repose, dynamic angle of repose, and tapped density. The virtual sensor addresses the issue with these flowability measurements, which are based on off-line characterization methods that can take hours to perform. With a virtual sensor based on real-time measurement methods, the prediction of granule flowability become faster, allowing for timely decisions regarding process control and the supply chain.

Keywords: virtual sensor, monitoring, flowability, machine learning, size and shape distribution

1. Introduction

The manufacturing of tablets often relies on a granulation step to improve the processability of a pharmaceutical powder blend. By converting them into granules, flowability, tabletability, compressibility, and compactibility can be improved. These properties are considered to be the critical quality attributes (CQA) of the granulation unit; and in a wet granulator, the monitoring of these properties is considered critical. Unfortunately, characterization tests for these CQAs are usually off-line methods that can take hours to measure and require sample reduction procedures that can lead to significant sampling errors. It is thus important to develop faster ways to estimate the CQAs and minimize sampling error.

In a wet granulation platform employing a fluidized bed, it is possible to measure the size and shape distribution of a finished batch of granules as it discharges from the product hopper. If these real-time measurements of size and shape can be used to automatically predict the CQA of the discharging granules, decisions regarding the batch and the process could be made much faster. This time advantage could save future batches from failure, provide valuable information about the raw material, and allow optimization of

the control process parameters of downstream unit operations to match the characteristics of each batch of granules.

Particle size and shape distributions are known to be strong indicators of granule properties, so they have a great potential to be reliable predictors of a granule's CQA's. However, their measurement result in a large number of data points that are difficult to manage and process. In practice, these distributions often get reduced to 1 to 3 D-values (i.e., D10, D50, D90) prior to analysis. There has been demonstrated success in this strategy, but this practice can lead to significant loss of information from the dataset, especially when the distributions have statistical central tendencies that do not fall close to those selected D-values. An ideal solution would be to employ all available information from the size and shape distribution measurements, and then to use an appropriate data reduction technique that maximizes the relevant information from the distribution measurements. By implementing this with an appropriate feature extraction technique, the reduced dataset should maximize correlatability with properties of interest such as flowability.

Aside from the predictor variables (i.e., size and shape distribution), the need for data reduction and feature extraction is also applicable for the predicted variables (i.e., the CQA's), especially for flowability. Since there is no singular measure for it, several methods exist to characterize flowability. Often, the goal in characterization is to select the method with test conditions that can closely match the conditions to which the granules are subjected to during processing. For some applications, a single method might suffice. But if granules will be subjected to tablet pressing, they will be subjected to quasi-static flow conditions in the hopper of the tablet press, as well as dynamic flow conditions inside the feed frame. Hence, several methods are required to ensure that the granules would result in quality tablets. Furthermore, each of these methods produce multiple test result parameters that are highly correlated. This can potentially result into a large dataset that needs to be appropriately reduced to make it more manageable and maximize its predictability with real-time measurements of granule size and shape distributions.

2. Methods

2.1. Data Reduction: Principal Component Analysis (PCA)

PCA is a method that reduces the dimensionality of large datasets while retaining most of its information. This is achieved by taking an orthogonal decomposition of the covariance matrix of process variables along the directions that explain the maximum variation of the data. (Wold *et al.*, 1987) While this method gives the same number of principal components as the original variables in the dataset, it also puts maximum possible variance in the first few principal components, making it possible to drop the rest of the principal components without losing much information. With the appropriate selection of principal components, data analysis and exploration can be performed on lower number of dimensions.

2.2. Latent Variable Regression

2.2.1. Linear Regression: Partial Least Squares (PLS)

With both the predictor and the predicted variables requiring data reduction through PCA, linear regression on their projections to latent spaces can be performed (i.e., projection to their principal components). This process is known as Partial Least Squares, and it is a

widely used technique in areas such as chemometrics bioinformatics, neurosciences, and sensor development, to name a few.(Liu and Chen, 2014)

3. Materials and Equipment

3.1. Granules

The granules used in this study are made with varying compositions of lactose and microcrystalline cellulose as the excipient, acetaminophen (APAP) as the active pharmaceutical ingredient (API), and either hydroxypropyl cellulose (HPC) or polyvinylpyrrolidone as the liquid binder solution. The excipient, API, and binders are prepared in varying compositions and wet granulation process conditions to produce granules with different flowability characteristics. This work studied four types of granules labelled as: HHIU1, HHIU2, HHIU3, and HHIU4.

3.2. Granulation Equipment

The granules are produced by wet granulation using the Xelum platform manufactured by Syntegon. Xelum employs a fluidized bed, where the pharmaceutical powders are automatically dosed and pneumatically charged with the liquid binders that facilitate the formation of granules. Moreover, granulation and drying takes place in the same process chamber, which eliminates the need to transfer wet granulate and improves the system's reliability.

3.3. Size and Shape Distribution Measurement

The size and shape distribution of the granules are measured using Eyecon₂, which is a direct imaging particle analyzer developed by Innopharma Technology. By using a camera to take images of the particles at-line or inline, this tool uses image analysis algorithms to detect particle boundaries and fit an ellipse around them. The ellipse gives a major and a minor diameter, which when averaged gives a third dimension to estimate a 3D volume of the particle using the equation:

$$Volume = \frac{\pi}{6} \times D_{min} \times D_{max} \times D_{ave} \quad \text{Equation 1}$$

Using this volume, an equivalent spherical diameter is computed, and this diameter is the basis for the size distribution reported by Eyecon₂. Size distributions are reported as D-values, which are based on the cumulative size distribution. Reporting distributions in this manner fixes the number of variables for every possible form of size distributions.

The major and minor diameters of each particle are also reflective of its shape, which may be quantified as eccentricity, as shown in the following equation.

$$Eccentricity = \sqrt{1 - \left(\frac{D_{min}}{D_{max}}\right)^2} \quad \text{Equation 2}$$

Eyecon₂ inherently acquires a distribution of eccentricity/shape but reports, by default, the distribution as a mean and relative standard deviation.

3.4. Flowability Measurements

The set of flowability measurements employed in this study covers both quasi-static flow and dynamic flow. Quasi-static flow is characterized by the ring shear tester and partly by tapped density analysis, while dynamic flow is characterized by drained and dynamic angle of repose.

3.4.1. Ring Shear Tester (RST)

The Schulze ring shear tester is an essential tool for hopper design since it is mainly concerned with quasi-static flow. In this technique, powder is loaded normally to a specific bulk density and then sheared until the material begins to flow. Data is collected as yield strength as a function of normal stress. From these measurements, the flow function coefficient can be computed, which may also be referred to as flowability. Additionally, other parameters such as internal friction, wall friction, and bulk density can be determined from the Schulze RST.

3.4.2. Drained Angle of Repose and Jamming Onset

The drained angle of repose is measured using the Flodex™ tool, which essentially measures the ability of a powder to fall freely under gravity through an orifice. Initially, the powder is contained in a hopper with a flow disk at the bottom. The disk has an orifice that can be opened via a discharge valve to start the powder flow. After opening the discharge valve, not all the powder in the hopper would be able to flow out and this residual powder would remain between the edge of the orifice and the hopper walls. The angle between the surface of this residual powder and the orifice disk is called the drained angle of repose and is correlated with the flowability of the powder.

3.4.3. Dynamic Angle of Repose

The dynamic angle of repose is measured using a rotary drum developed by GranuTools called the GranuDrum™. The powder is loaded into a drum that can be rotated at a set rotating speed. As the drum is rotated from rest, the angle of the powder surface increases from horizontal until an avalanche occurs. The angle at which this happens may be referred to as the yield point and is correlated with flowability. Thereafter, the powder surface is maintained at an angle from horizontal, and this is recorded automatically using back-lit cameras as the dynamic angle of repose. As the rotation speed of the drum changes, the dynamic angle of repose also changes, revealing interesting rheological behaviors of powder during flow.

3.4.4. Tapped Density Analyzer

Tapped density analysis is performed by another tool developed by GranuTools called GranuPack™. This tool minimizes operator error during filling and volume measurements using automation and sensor technologies. Powder is loaded onto a cylinder container and its density is monitored as the container is tapped continuously. As the powder is tapped, the density increases until it asymptotically approaches a maximum. The density may be expressed as the Hausner ratio, which is basically the ratio between the tapped density and the poured density. The dynamics of the compaction during tapping is also automatically captured via the parameters characteristic number and tau. The characteristic number is the number of taps at which the density is between between the poured density and the asymptotic density (i.e., density at infinite number of taps), while Tau is another characteristic number extrapolated from an exponential model (Philippe and Bideau 2003) fitted onto the compaction curve.

4. Results and Discussion

4.1. Principal Component Analysis on Size and Shape Distributions

Size and shape distributions measurements can result in at least 24 variables as shown in the x-axis of Figure 1. Applying principal components analysis (PCA) on the dataset reduced the number of variables into just 3 principal components (PC), which can explain up to 97% of the variance in the original dataset.

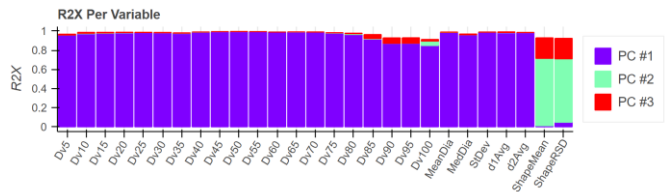


Figure 1. Explained variance per size and shape distribution variable

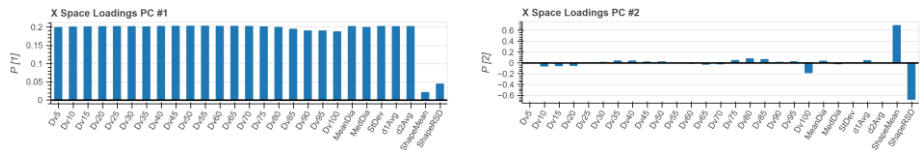


Figure 2. Loadings of size and shape distribution variables on principal component 1 (left figure) and principal component 2 (right figure)

This drastic reduction of variables suggests that many of them are highly correlated, as shown in the loadings plot in Figure 2. This is the case for the size distribution variables, as they dominate influence on the first principal component, which explains 89% of the variation in the original data. On the other hand, shape-related variables (i.e., the shape mean and relative standard deviation) have the strongest influence on the second principal component, supporting the importance of measuring shape distributions, and not just size.

4.2. Predicting Flowability Data

Measurements from the Schulze RST can lead to 9 different parameters (or variables) that are related to flowability. Similar to the Eyecon2 data, most of the variation in these parameters (up to 98%) can be explained by only three principal components. Hence, by applying partial least squares using three principal components onto the Eyecon2 data (as the predictor variables) and the RST data (as the predicted variables), the parity plot of the flow function coefficient (FFC) shown in Figure 3 show good prediction performance. Although not shown, similar performance was also observed for the rest of the variables.

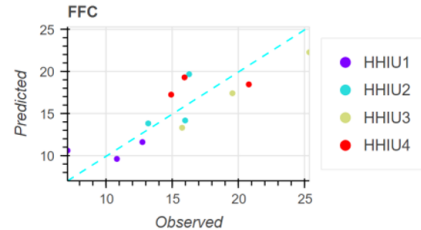


Figure 3. Predicted vs observed flow function coefficients measured from Schulze RST.

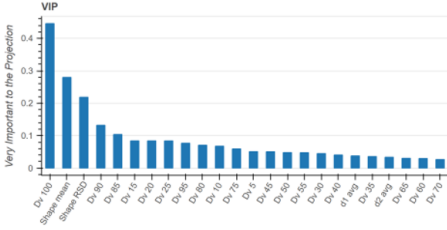


Figure 4. Ranking of variable importance to the PLS projections.

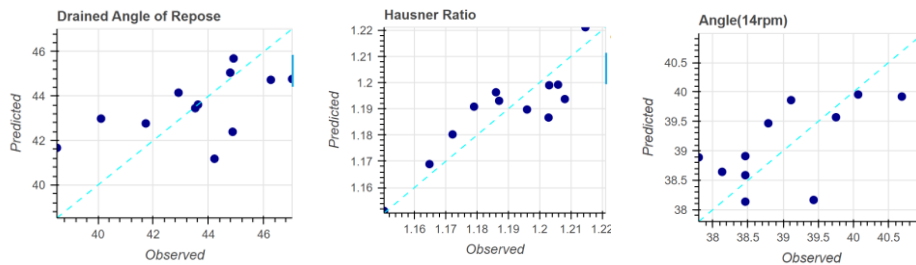


Figure 5. Parity plots for selected dynamic flow test parameters: drained angle of repose (left), Hausner ratio (middle), and dynamic angle of repose (right).

The performance of the PLS model can be attributed to the effectiveness of using all available information instead of selecting some and then ignoring the rest. Figure 4 shows how the larger D-values (e.g., D85, D90, and D100) and the shape parameters contribute the most to the PLS projections and hence its performance. This not only corroborates the importance of measuring shape distributions, but also the folly of selecting certain D-values such as D50, D10, and D90. As shown in Figure 4, those variables are not the most important. Using the aforementioned techniques, similar results were achieved from the dynamic flow tests, as shown in Figure 5 for selected parameters from the Flodex (left figure), GranuDrum (middle figure), and GranuPack (right figure) measurements.

5. Conclusions

Using PLS regression, sensor models were developed to predict flowability measurements based on size and shape distribution of granules, and parity plots show good predictability for all flowability measurements. The importance of shape measurements as well as using the complete size distribution, instead of selecting a few D-values, in the predictive performance was highlighted.

6. Acknowledgements

PCA and PLS computations were performed using the Phi toolbox (github.com/salvadorgarciamunoz/pyphi) for multivariate analysis by Dr. Sal Garcia.

This work was supported by United States Food and Drug Administration under grant 1U01FD006487-01.

7. References

- S. Wold, K. Esbensen, and P. Geladi. 1987. "Principal Component Analysis", *Chemometrics and Intelligent Laboratory Systems*, 2.1-3,37-52
- J. Liu, D.S. Chen. 2014. "Developing a Soft Sensor for an Air Separation Process Based on Variable Selection in Dynamic Partial Least Squares", *Computer Aided Chemical Engineering*, 33, 685-690
- P. Philippe, and D. Bideau. 2003. "Granular Medium under Vertical Tapping: Change of Compaction and Convection Dynamics around the Liftoff Threshold." *Physical Review Letters*, 91 (10): 104302.