Formulation-independent pharmaceutical dry granulation model via gray box approach

a Formulation Technology Research Laboratories, Daiichi Sankyo Co., Ltd., 1-12-1 Shinomiya, Hiratsuka, Kanagawa 254-0014, Japan

bDepartment of Systems Science, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan

sato.kanta.m6@daiichisankyo.co.jp

Abstract

In the roller compaction process, powder blends are compacted into ribbons by opposite-direction rotating rolls and subsequently milled to obtain granules. Since ribbon density is an important material attribute affecting product quality, controlling it is essential. We developed a novel mechanistic model describing the impact of process parameters and material attributes on ribbon density. By incorporating the pressure of a preconsolidation powder blend into the model, the compression coefficient of the powder blend in the small-scale uniaxial compression tests can be applied to the large-scale roller compactor. This allows us to understand the relationship between process parameters and ribbon density from small-scale experiments with few materials. A gray box model was developed to predict ribbon density, combining a first principle model based on the Johanson model and a statistical model. The accuracy of the gray box model was validated through manufacturing data, achieving precise predictions with a Root Mean Square Error of Cross Validation of 0.032 in relative ribbon density. This approach is expected to significantly reduce material consumption and enhance comprehension of the roller compaction process. The validation results demonstrate its capacity to predict ribbon density using only a minimal amount of experimental material, provided that a compressibility coefficient is within the range of the training data.

**Keywords**: Roller compaction, Gray box modeling, Pharmaceutical manufacturing, Process modeling, Scale up

* 1. Introduction

To achieve the desired product quality, it is essential to accurately understand the relationship between process parameters and product quality but understanding the relationship requires significant resources and costs. Predictive modeling is crucial in understanding these relationships. First principle modeling requires deep insight into the process. In some cases, parameters may be required that are not easily accessible during normal operation. The serial gray box model supplements certain parameters of a first principle model with statistical models (Ahmad et al., 2020). According to Von Stosch et al. (2014), statistical models can be applied with limited process knowledge, but they require large amounts of data compared to a first principle model, and the quality of the model can only be trusted in the vicinity of the data from which the model was derived. Gray box models balance the advantages and disadvantages of first principle models and statistical models.

The granulation process is a crucial step that improves the flowability of the powder and increases the manufacturing efficiency of subsequent processes (Kleinebudde, 2004). The roller compaction (RC) process is a continuous process in which powder blends are compacted into ribbons by rolls and subsequently milled into granules. In the RC process, ribbon density control is critical to ensure the quality of the final products. Johanson (1965) proposed a first principle model for RC, the rolling theory of granular solids that correlates materials attributes and process parameters with the ribbon density in a simplified two-dimensional space as shown in Figure 1(a).

To obtain the model parameters, RC manufacturing must be performed over a wide range of ribbon densities on the target manufacturing machine as material attributes (Raynolds et al., 2010). When data sets based on specific materials are limited, there may be a discrepancy between predicted and experimental ribbon densities. Amini and Akseli (2020) succeeded in predicting ribbon density utilizing the Johanson model by substituting the RC manufacturing with uniaxial compression tests, as shown in Figure 1(b). This was achieved under limited conditions such as the roll speeds ranging from 2 to 6 min−1 where the preconsolidation density would remain constant regardless of the process parameters. Considering the actual process development scenario requiring a comprehensive evaluation of the process parameters to identify the design space, a robust first principle modeling approach that can take the variation of preconsolidation density into account is required. To accommodate varying production speeds, manufacturing at higher roll speeds may be required.

This study aims to build a model predicting ribbon density under a wide range of process parameters and to set process parameters that achieve the desired ribbon density from a small amount of material. By regarding preconsolidation powder conditions as variable parameters, we resolved the discrepancy between RC and uniaxial compression tests. A gray box model was built to predict ribbon density. A statistical model was developed to estimate the preconsolidation parameters, which were difficult to measure, and the ribbon density was calculated using a first principle model with the estimated parameters.

* 1. Material and Methods
		1. Materials

This study used four formulations A to D. Formulation A consisted of 72 % PEARLITOL 50 C as mannitol (Roquette, France), 18 % CEOLUS PH-101 as microcrystalline cellulose (Asahi Kasei, Japan), and 8 % NISSO HPC-SL as hydroxypropyl cellulose (HPC) (Nippon Soda, Japan), to provide the ideal relative ribbon density range (Zinchuk et al., 2004). Formulations B to D consisted of different ratios of plastic material such as CEOLUS KG-1000, microcrystalline cellulose (MCCKG) (Asahi Kasei, Japan), and brittle material such as dibasic calcium phosphate anhydrous (DCPA) (Fuji Chemical Industries, Japan) to provide different compaction behavior. Compositions of MCCKG and DCPA are 78 % and 20 % for formulation B, 49 % each for formulation C, and 20 % and 78 % for formulation D. 2 % HyQual 5712, magnesium stearate (SpecGx, USA) was included as a lubricant in all formulations.

Figure 1 Schematic diagram of (a) roller compaction process, (b) uniaxial compression test

(a)

(b)

RC manufacturing was conducted using a roller compactor FP90×30 (Freund Turbo, Japan) over a wide range of process parameters, i.e., roll pressure of 8–27 MPa, roll speed of 6–20 min−1, and roll gap of 0.75–1.04 mm.

The compressibility coefficient (*K*) for each formulation was determined through uniaxial compression tests. 300 mg of the powder blends were placed in a 10 mm diameter die and compacted with a flat-faced punch at 10 mm/min, and then the density was measured. The density of the compacts obtained in the uniaxial compression test was measured in the range of 2.5 MPa to 100 MPa to cover the ribbon density obtained in RC manufacturing.

* + 1. First principle model

The powder blend fed by the screw experiences the screw feed pressure (*P0*) in the slip region, resulting in the preconsolidation relative density (*γ0*). The relative density is the ratio of the true density of the material to the density of the ribbon. When the roll angle (*θ*) becomes less than the nip angle (*α*), compaction of the powder begins. This area is called the nip region, and the compaction of the powder increases as *θ* decreases. At *θ=0*, the distance between the rolls is the shortest, and the powder experiences the maximum pressure (*Pmax*) from the rolls. For a detailed derivation of *Pmax*, please refer to Johanson's original paper (Johanson, 1965). The powder has *K*, and the relative density of the compacted ribbon (*γR*) is expressed by Eq. (1).

|  |  |
| --- | --- |
|  | (1) |

Eq. (2) is given by rearranging the logarithm of Eq. (1).

|  |  |
| --- | --- |
|  | (2) |

Raynolds et al. (2010) set *P0*=1, regardless of the material attributes and process parameters, and set the slope of ln*γR* and ln*Pmax* in Eq. (2) as 1/*K* with an intercept of ln*γ0*. They found differences in *K* and *γ0* between the uniaxial compression test and RC. In RC, a velocity gradient occurs in the nip region, and the powder blend near the roll surface moves faster than the powder blend farther away (Muliadi et al., 2012), which deviates from Johanson's assumption that the powder blend moves at a uniform velocity. We assumed that *K* remains constant regardless of the machines. Under the assumption, the discrepancy in *K* between RC and uniaxial compression test calibration was due to the different assumed values of *P0* and *γ0*. If there is a velocity gradient in the preconsolidation powder blend around the rolls in RC due to process parameters, both *P0* and *γ0* expected to vary accordingly. It is usually difficult to measure *P0* and *γ0* during normal operation because the preconsolidated powder blend in RC exists in the closed space. Since the slip region remains constant within the same machines, its volume can also be considered constant, given a sufficient feed of powder blend. Once *P0* is determined, *γ0* can be deduced consequently. We introduced an unmeasurable parameter *ζ* to represent the relationship between *P0* and *γ0* in Eq. (3).

|  |  |
| --- | --- |
|  | (3) |

*γR* is expressed by Eq. (4) based on Eq. (1) and Eq. (3).

|  |  |
| --- | --- |
|  | (4) |

Using a statistical model, *ζ* was calculated from measured manufacturing results and predicted from the powder blend attributes and process parameters. It was used to calculate *γR* in Eq. (4). Eq. (2) can be transformed into Eq. (5).

|  |  |
| --- | --- |
|  | (5) |

*K* was obtained from the uniaxial compression test and used for RC. Since the only pressure applied to the preconsolidated powder blend in the uniaxial compression test is atmospheric, the preconsolidation relative density can be considered constant for the same material. Consequently, *K* can be calculated from the slope of ln*Pmax* and (ln*γR−*ln*γ0*)*.*

* + 1. Statistic model

Gaussian Process Regression (GPR) was used to build the statistical model with K-Fold Cross Validation (KFCV; K=10). In KFCV, the dataset is divided into K equal subsets, with one subset used as test data and the remaining K−1 subsets used as training data. This process is repeated K times and the results are averaged to evaluate the performance of the model. Prediction performance was evaluated using the coefficient of determination R² and the Root Mean Square Error Cross Validation (RMSECV) or Root Mean Square Error Prediction (RMSEP). The input parameters used for the GPR model were the same as those for the first principle model, i.e., true density, bulk density, wall friction angle, effective angle of internal friction, roll force, roll gap, roll speed, and *K*. The importance of the input parameters was evaluated using Permutation Importance (PI). *ζ* was calculated from Eq. (4) using the dataset for each run.

* + 1. Case Study

All formulations were used in the model building and cross validation confirmed the accuracy of the ribbon density predictions. Then, to predict RC manufacturing results using only small-scale experiments, a model was built using RC manufacturing data for three of the four formulations to predict the relative ribbon density of the untrained formulation, and the relative ribbon density of the remaining formulation was predicted.

* 1. Results

Figure 2 shows the relationship between *K*, roll force, roll gap, roll speed, and *ζ*, where *ζ* is derived from manufacturing data. *K* increased with higher DCPA content in formulations B to D. Formulation A showed a significantly higher value than the other three formulations. One factor that can be inferred is the existence of the binder HPC in Formulation A, which is not included in Formulations B to D. *ζ* was found to vary not only with formulation but also with manufacturing conditions. Therefore, it is suggested that *ζ* should not be treated as a fixed parameter based solely on material attributes, but rather as a variable parameter that also depends on process parameters.

Figure 2 Scatter plots matrix of process parameters and *ζ*.

Table 1 shows the prediction accuracy of *ζ* using the statistical model. All formulations achieved excellent prediction accuracy for *ζ*, with RMSECV of 0.010 and R2 of 0.993. The prediction of *ζ* was significantly affected by *K*, as indicated by the PI. Figure 3(a) shows the scatter plots of observed versus predicted values of *γR* using *ζ* obtained in cross validation. An accurate prediction of the RMSECV of 0.032 and R2 of 0.962 indicates that *γR* can be accurately predicted using the gray box model approach.

|  |
| --- |
| Table 1 Prediction accuracy of *ζ* in case studies and PI for modeling |
| Target formulation | All data | A | B | C | D |
| *ζ* prediction | Cross validation | RMSECV | 0.010 | 0.009 | 0.010 | 0.011 | 0.011 |
| R2 | 0.993 | 0.860 | 0.993 | 0.992 | 0.993 |
| Prediction of the untrained formulation | RMSEP | – | 0.244 | 0.013 | 0.006 | 0.009 |
| R2 | – | −40.266 | 0.838 | 0.910 | 0.801 |
| Permutation Importance in cross validation | True density (g / mL) | 0.263 | 0.095 | 0.087 | 0.295 | 0.320 |
| Bulk density (g / mL) | 0.081 | 0.045 | 0.078 | 0.044 | 0.040 |
| Wall friction angle (°) | 0.001 | 0.210 | 0.001 | 0.001 | 0.000 |
| Effective angle of internal friction (°) | 0.023 | 0.309 | 0.027 | 0.030 | 0.012 |
| Roll force (MPa) | 0.037 | 0.710 | 0.036 | 0.035 | 0.042 |
| Roll gap (mm) | 0.084 | 1.331 | 0.076 | 0.076 | 0.091 |
| Roll speed (min−1) | 0.003 | 0.057 | 0.003 | 0.003 | 0.002 |
| *K* (-) | 0.348 | 0.039 | 0.376 | 0.393 | 0.290 |

Table 1 displays predicted results for the untrained formulation. The RMSEP for *ζ* from formulations A to D were 0.244, 0.013, 0.006, and 0.009, respectively. Figure 3(b) shows the results for *γR* calculated using the predicted *ζ*, with corresponding RMSEP values of 0.518, 0.044, 0.023, and 0.039. These results indicate practical prediction accuracy, except for formulation A. In formulation A, a significant difference in the distribution of the largest PI in the statistical model compared to the other formulations was observed. This difference can be attributed to the different distribution of the manufacturing data in Figure 2. Specifically, *K* values, which is the largest PI in the statistical model, were significantly different for formulation A from the other three formulations, resulting in an extrapolation in the GPR model built with the other three formulations.

Figure 3 (a) Cross validation of *γR* for all formulations: experimental vs predicted. Data fitted with RMSECV=0.032, R2=0.962. (b) Predictive validation of *γR* on untrained formulation. Data fitted with (A) RMSECV=0.518, R2=-61.411, (B) RMSECV=0.044, R2=0.776, (C) RMSECV=0.023, R2=0.897, (D) RMSECV=0.039, R2=0.720.

(a)

(b)

While the dataset used in this study was limited to specific formulations, future research should use a wider range of datasets to improve the generalizability and robustness of the model. This study showed the potential variability in manufacturing parameters for *ζ* and the effect of *K* on *ζ*, demonstrating the complexity of the interaction between formulation and manufacturing processes.

The gray box modeling used in this study captured the differences between formulations with different material attributes while incorporating the variability in manufacturing parameters to achieve high predictive accuracy. By combining a statistical approach based on experimental data with mechanistic modeling based on phenomena, this method offers a new way to predict the behavior of formulations under different manufacturing conditions. This advancement is expected to enable innovative formulation design that significantly reduces development time and cost. In addition, by increasing the predictability of manufacturing processes, the gray box model is expected to contribute to improved robustness and quality in pharmaceutical formulation development.

* 1. Conclusion

This study presents a gray box modeling approach to predict ribbon density in pharmaceutical manufacturing, which is essential for product quality in the roller compaction process. This gray box model achieved accurate predictions with RMSECV of 0.010 and 0.032 for *ζ* and relative ribbon density. We have established a predictive relationship between preconsolidation pressure and relative density, which is critical for optimizing the manufacturing process. Model validation results show that the ribbon density can be predicted with only a small amount of experimental material as long as *K* is within the range of the training data. Our research contributes a validated, adaptable model to the pharmaceutical industry that promises improvements in product quality and streamlined the manufacturing process. The model's implications include resource reduction and an improved understanding of the manufacturing process.

References

I. Ahmad, A. Ayub, M. Kano, I.I. Cheema, 2020, Gray-box soft sensors in process industry: current practice, and future prospects in era of big data, Processes, 8, 243.

H. Amini, I. Akseli, 2020, A first principle model for simulating the ribbon solid fraction during pharmaceutical roller compaction process, Powder Technology, 368, 32–44.

J. Johanson, 1965, A rolling theory for granular solids, Journal of Applied Mechanics, Transaction of ASME, 32, 842–848.

P. Kleinebudde, 2004, Roll compaction/dry granulation: pharmaceutical applications, European Journal of Pharmaceutics and Biopharmaceutics, 58, 317–326.

A. R. Muliadi, J. D. Litster, C. R. Wassgren, 2012, Modeling the powder roll compactionn process: Comparison of 2-D finite element method and the rolling theory for granular solids (Johanson’s model), Powder Technology, 221, 90–100.

G. Reynolds, R. Ingale, R. Roberts, S. Kothari, B. Gururajan, 2010, Practical application of roller compaction process modeling, Computers and Chemical Engineering, 34, 1049–1057.

M. V Stosch, R. Oliveira, J. Peres, S. F.de Azevedo, 2014, Hybrid semi-parametric modeling in process systems engineering: Past, present and future, Computer and Chemical Engineering, 60, 86-101.

A. V. Zinchuk, M. P. Mullarney, B. C. Hancock, 2004, Simulation of roller compaction using a laboratory scale compaction simulator, International Journal of Pharmaceutics, 269, 403–415.