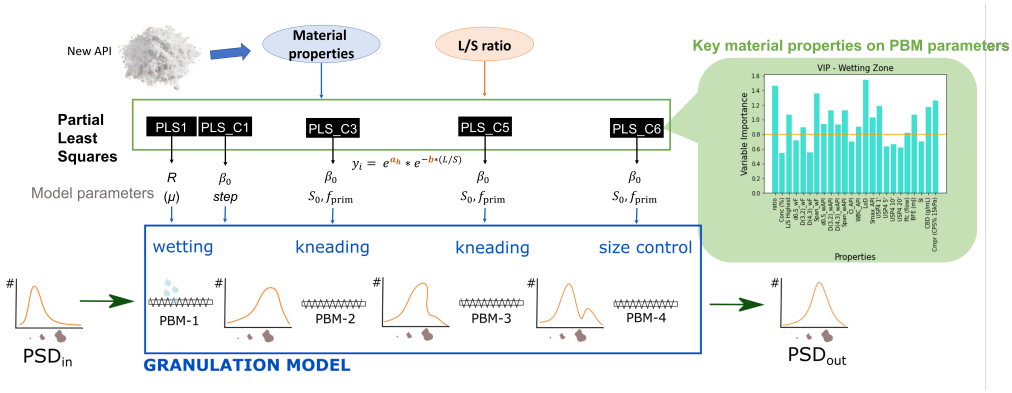
Graphical Abstract

Partial least squares regression to calculate population balance model parameters from material properties in continuous twin-screw wet granulation

Ana Alejandra Barrera Jiménez, Kensaku Matsunami, Daan Van Hauwermeiren, Michiel Peeters, Fanny Stauffer, Eduardo dos Santos Schultz, Ashish Kumar, Thomas De Beer, Ingmar Nopens



Highlights

Partial least squares regression to calculate population balance model parameters from material properties in continuous twin-screw wet granulation

Ana Alejandra Barrera Jiménez, Kensaku Matsunami, Daan Van Hauwermeiren, Michiel Peeters, Fanny Stauffer, Eduardo dos Santos Schultz, Ashish Kumar, Thomas De Beer, Ingmar Nopens

- Material properties were linked with population balance model parameters.
- Partial least square regression models were used to simulate model parameters.
- Experimental data from 10 formulations and different process settings were used.
- The models identified key material properties in twin-screw wet granulation.
- The models lead to a generic population balance model applicable to new drugs.

Partial least squares regression to calculate population balance model parameters from material properties in continuous twin-screw wet granulation

Ana Alejandra Barrera Jiménez[a,b,1,*], Kensaku Matsunami[a,b,1,*], Daan Van Hauwermeiren[a,b], Michiel Peeters[b], Fanny Stauffer[c], Eduardo dos Santos Schultz[d], Ashish Kumar[d], Thomas De Beer[b], Ingmar Nopens[a]

[a]*BIOMATH, Department of Data Analysis and Mathematical Modelling, Ghent University, Coupure links 653, Ghent, 9000, Oost-Vlaanderen, Belgium*
[b]*Laboratory of Pharmaceutical Process Analytical Technology, Department of Pharmaceutical Analysis, Ghent University, Ottergemsesteenweg 460, Ghent, 9000, Oost-Vlaanderen, Belgium*
[c]*Product Design & Performance, UCB, Ottergemsesteenweg 460, Braine lAlleud, 1420, Belgium*
[d]*Discovery, Product Development & Supply, Janssen R&D, Beerse, B-2340, Belgium*

Abstract

In the pharmaceutical industry, twin-screw wet granulation has become a realistic option for the continuous manufacturing of solid drug products. Toward the efficient design, population balance models (PBMs) have been recognized as a tool to compute granule size distribution and understand physical phenomena. However, the missing link between material properties and the model parameters limits the swift applicability and generalization of new active pharmaceutical ingredients (APIs). This paper proposes partial least squares (PLS) regression models to assess the impact of material properties on PBM parameters. The parameters of the compartmental one-dimensional PBMs were derived for ten formulations with varying liquid-to-solid ratios and connected with material properties and liquid-to-solid ratios by PLS models. As a result, key material properties were identified in order to calculate it with the necessary accuracy. Size- and moisture-related properties were influential in the wetting zone whereas density-related properties were

*Corresponding author
[1]Shared first authorship.

more dominant in the kneading zones.

Keywords: Continuous manufacturing, solid drug products, granule size distribution, hybrid model, mechanistic model, data-driven model.

1. Introduction

In pharmaceutical applications, twin-screw wet granulation (TSWG) has shown advantages to address the transition from batch to continuous manufacturing of solid dosage forms, due to its design flexibility, short residence time, stability, and controlled throughput range [1, 2] to name but a few. A good understanding of the mechanisms that govern the wet granulation process is very important to be able to incorporate control strategies, as well as the Quality by Design (QbD) paradigm, into a new drug development process. In that sense, various types of experimental research have been conducted to understand the process itself, the influence of process parameters (PPs), formulation, and equipment design on the properties of the final granules [1, 3, 4, 5, 6, 7, 8, 9, 10, 11]. Numerous mechanistic modeling approaches have aimed to reproduce the observed experimental behavior of the TSWG process, including the physical phenomena that have been identified as part of the process such as wetting, nucleation, consolidation, growth, breakage, and attrition [12, 13].

Among the applicable mechanistic modeling strategies for wet granulation, the Population Balance Model (PBM) framework has emerged as a powerful prediction tool for granule properties [14]. PBM can be combined with other modeling techniques, e.g., Discrete Element Modeling (DEM) and Computational Fluid Dynamics (CFD), that provide complementary information of the system to improve the solution of Population Balance Equations (PBEs) [2, 13, 15]. Recently, PBM has been coupled with data-driven models, e.g., partial least square (PLS), to link outputs of the PBM model to critical quality attributes (CQAs) of the granules, obtaining more information from the system in addition to the particle size distribution (PSD) that it is attained directly from the PBM simulation. For instance, Liu et al. [16] proposed a PLS regression to model the relationship between the kernel parameters of a PBM and the manipulated operating variables of the TSWG process. As an application to other unit operations, Metta et al. [17] used a PBM to track particle mass change in a comill unit. They then utilized

the resulting PSD and granule moisture content to develop a PLS model for predicting milled product CQAs such as bulk density, tapped density and friability.

Despite all these mechanistic modeling contributions [13, 18, 19, 20, 21, 22, 23, 24, 25] to simulate the TSWG process, some links among material properties, PPs, and screw configurations remain unestablished within those models for them to be more generic and applicable for new drug developments. PBMs need to be calibrated for each formulation, which requires extensive experimental efforts. Furthermore, it is challenging to understand phenomena without the identification of key material properties. Pure data-driven modeling approaches have been performed to assess the impacts of material properties and process parameters on product quality attributes [26, 27, 28, 29, 10]. However, data-driven models do not help to understand the detailed phenomena of the process, e.g., physical meanings of key parameters, as well as do extrapolation. In addition, data-driven models to predict full granule size distributions are limited to derived quantities, e.g., d10, d50, and d90 [30, 31] due to the difficulties in reproducing the distributions. The research question exists in the missing link between material properties and the model parameters, which limits the swift applicability and generalization to different formulations.

This work presents data-driven models to link material properties and process parameters with a PBM as a key step to develop a generic PBM, which predicts specific CQAs of granules produced by the TSWG process. Barrera Jiménez et al. [32] presented a one-dimensional compartment PBM, which can be performed with a less number of the model parameters. In this paper, PLS models were developed to generalize the PBM described in Barrera Jiménez et al. [32] and identify key material properties for each phenomenon and compartment. Using material properties and process parameters, the developed PLS models can compute PBM parameters, which can subsequently be used for the simulation of granule size distribution. Both experimental design and pre-processing of PLS development were crucial to compute PBM parameters for new formulations. In the next section, details about the experimental setup are given such as equipment type, process settings, and the formulation and active ingredients under study. Subsequently, the mathematical frameworks of PBMs and PLS models are described, which will be used to address the research question in this paper.

2. Materials and Methods

2.1. Design of the experiments

The workflow of the interaction between the formulation selection and both, the experimental and modeling work, is illustrated in Figure 1. A five-level central composite Design of Experiments (DoE) with two or three factors for the majority of formulations was conducted to investigate the relationship between formulation properties (% contain active pharmaceutical ingredient), PPs (liquid-to-solid (L/S) ratio, screw speed (SS), and mass feed rate (MFR)) and granule characteristics such as PSD, and granule friability. The factors (PPs) that were varied for each formulation are indicated in Table 1. Then, each DoE resulted in 8 and 14 experiments when using two and three factors respectively. Three center-point experiments were performed in each DoE. This work is a continuation of the work of Barrera Jiménez et al. [32, 33]. For more details on the process and characterization techniques, the reader is referred to the aforementioned work.

The formulations contained the same excipient base and differed solely in the Active Pharmaceutical Ingredient (API). The excipient base consists of lactose (30 or 75 % w/w depending on API content; Pharmatose 200M, DEF Pharma, Veghel, The Netherlands), microcrystalline cellulose (15 % w/w; Avicel PH101, FMC BioPolymer, Cork, Ireland), and hydroxypropyl methylcellulose (5 % w/w; Methocel E15, Dow Chemical Company, Midland, USA). The API differed in terms of its nature, e.g., both hydrophilic and hydrophobic, and concentration, both low and high concentrations. Then, in this work, ten formulations were included in the calibration process of the PBM. Formulations that were possible to process with the same screw configuration and maintain process stability were considered.

The granules collected and characterized in this study were produced in the high shear twin-screw wet granulator module of the ConsiGma-25 (GEA Pharma Systems, ColletteTM, Wommelgem, Belgium), continuous line. First, the blended materials flowed through the wetting zone of the granulator (denominated as C1 in Figure 2), where demineralized water was added by a twin-peristaltic pump. The wet materials were further processed to the two kneading zones (C3 and C5 in Figure 2), and size control elements (C6 in Figure 2). Process parameters such as screw speed, material throughput, and L/S ratio were chosen to operate the twin-screw granulator according to the processability of each formulation to obtain similar granules in all the formulations studied (see Table 1). From the experimental work,

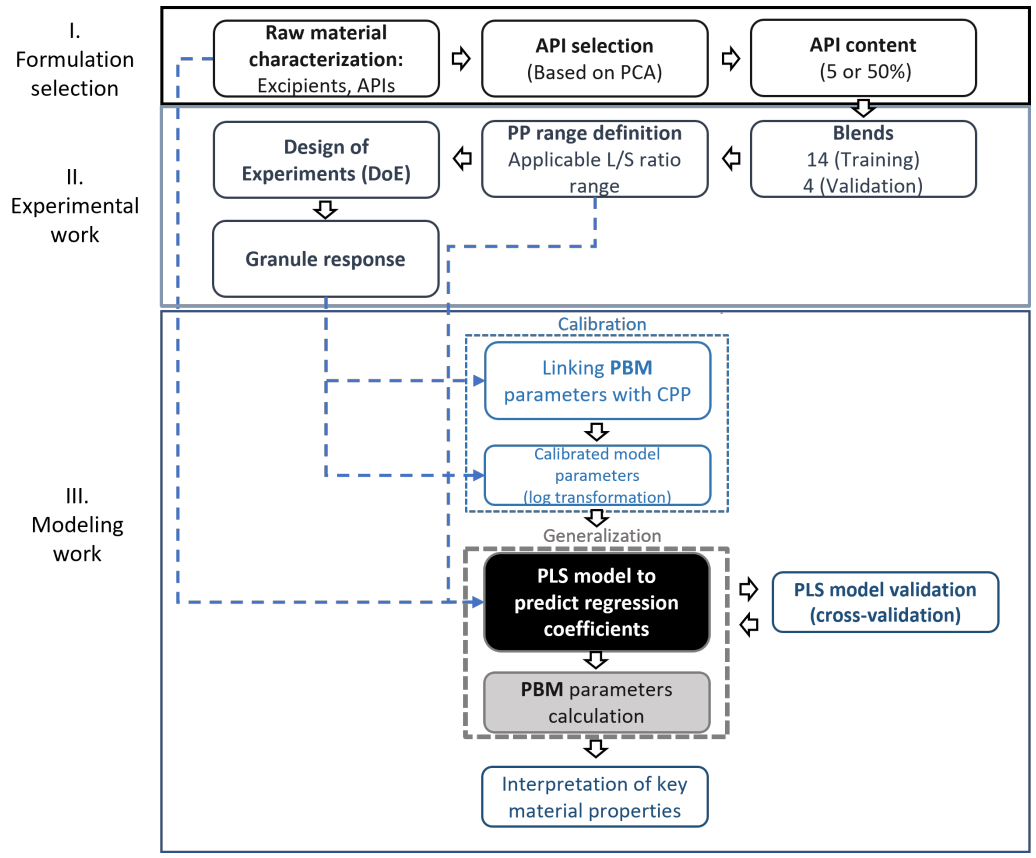


Figure 1: Workflow from material characterization to critical CQAs (particle size) using PLS models to predict PBM parameters.

it was concluded that the L/S ratio has the highest impact on the CQAs of the granules for any formulations in this study, which is in accordance with previously reported experimental research [7, 34]. Thus, in this work towards a generic PBM model as a tool to define the manufacturable ranges of a given new formulation, the L/S ratio was focused on and linked to the PBM parameters as a representative of PPs.

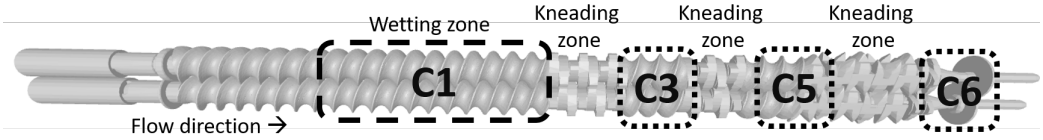


Figure 2: Compartments along the twin-screw.

Table 1 indicates the total number of experiments required according to the DoE plan. However, Figure 3 shows that after testing the first two APIs, it was possible to identify five critical experimental points that were sufficient to establish correlations. These points consist of the lowest, low, center point (in triplicate), high, and highest L/S ratio. Remarkably, a consistent trend was observed in the correlations between the calibrated model's parameters and the L/S ratio across all formulations. Therefore, these five points were deemed to be adequate to build the correlations. In order to reduce the experimental workload, fewer experiments were conducted for the rest of the APIs. From API 4 onwards, the five points were collected and characterized, that is: at the lowest, lowest (in triplicate), highest and highest central point, the L/S ratio granules. The number of experiments included in the calibration, and therefore, in the construction of the models are also reported in Table 1. API content, SS, and MFR were fixed as 50%, 675 rpm, and 20 kg/h, respectively, whereas L/S ratios, were varied from the lowest to the highest for each formulation. The wet granules were oven-dried, and their PSDs were measured using a QICPIC particle size analyzer (Sympatec, Etten-Leur, The Netherlands).

2.2. Population Balance Model description

The experimental data from section 2.1 was used to calibrate the one-dimensional compartmental PBM described in Barrera Jiménez et al. [32], which employs as internal coordinate the particle size. The model considers

Table 1: Process conditions of all formulations used in the experiments.

API	Formulation number	Number of experiments	Experiments included	Screw Speed (rpm)	Throughput (kg/h)	L/S (%)
API 1 (5 %)	1	11	11	675	15 - 25	8.9 - 20.2
API 1 (50 %)	2	11	11	675	15 - 25	13.6 - 23.7
API 2 (5 %)	3	17	16	450 - 675	15 - 25	8 - 18
API 2 (50 %)	4	17	17	450 - 675	15 - 25	5.2 - 16
API 3 (5 %)	5	17	5	750 - 900	15 - 25	15.2 - 18.5
API 3 (50 %)	6	17	5	450 - 675	15 - 25	5.2 - 13.4
API 4 (50 %)	7	17	6	450 - 675	15 - 25	18 - 28
API 5 (50 %)	8	17	7	450 - 900	15 - 25	21 - 26.4
API 6 (50 %)	9	17	7	450 - 900	15 - 25	21.6 - 28
API 7 (50 %)	10	7	7	675	20	15 - 20

four compartments, consisting of one wetting zone and three kneading zones (Figure 2). In the wetting zone, the binder liquid is added to the granulator, and pure aggregation is assumed to be the dominant physical phenomenon due to the low-shear environment [7]. In the three subsequent kneading zones, the model represents the change in PSD by a combination of aggregation and breakage kernels.

The population balance equation expressed as the temporal change of the particle numbers in a spatially homogeneous system is described as [35]:

$$\frac{\partial n}{\partial t} = \frac{1}{2}\int_0^x B(x-\varepsilon,\varepsilon,t)n(x-\varepsilon,t)n(\varepsilon,t)d\varepsilon - n(x,t)\int_0^\infty B(x,\varepsilon,t)n(\varepsilon,t)d\varepsilon + \int_0^\infty b(x,\varepsilon)S(\varepsilon)n(\varepsilon,t)d\varepsilon - S(x)n(x,t). \quad (1)$$

In Equation (1), $n(x,t)$ represents the density of particle numbers as a function of time (t) and its internal coordinate (i.e., granule size (x) expressed as particle volume). Time is interpreted as the residence time within the TSWG, considering that the process is modeled in steady state. Therefore, each simulation tracks the number of entities from t_o (initial) to t_e (final). The first term on the right-hand side of Equation (1) means the birth of particles because of aggregation whereas the second term refers to the death of particles caused by aggregation dynamics. The birth term can be interpreted as the formation of particles of size x, after the aggregation of a particle with size ε and a particle with size $x-\varepsilon$. As a consequence, the death of the initial particles that were aggregated results in their loss. In Equation (1), B means the aggregation kernel, which is used to describe the aggregation rate of two particles ε and $x-\varepsilon$. The aggregation kernel $B(t,x,\varepsilon)$ is expressed

as $\beta_0(t)\beta(x,\varepsilon)$, where $\beta_0(t)$ and $\beta(x,\varepsilon)$ mean the aggregation efficiency and the collision frequency. The last two terms of Equation (1) express the particle birth and death caused by breakage. The parameter b is the probability density function for the formation of a particle x from a particle ε, whereas S is the selection function that establishes the ratio of breakage of particle x.

The collision frequency is presented in Equation (2) [32]:

$$\beta(x,\varepsilon) = \begin{cases} \left(x^{2/5}\cdot\varepsilon^{2/5}\right), & x \wedge \varepsilon < R \\ \text{step } \cdot\left(x^{2/5}\cdot\varepsilon^{2/5}\right), & x \vee \varepsilon > R. \end{cases} \tag{2}$$

In Equation (2), R represents the critical particle size, and *step* is the scaling parameter related to the amount of non-granulated material in the wetting zone.

To model the kneading zones, a linear selection function was chosen: $S(m) = S_0 m^{1/3}$. S_0 is the breakage rate constant at a mother particle with the initial size of m (expressed as volume) before breakage. Moreover, the breakage function $b(m,d)$ in the expression (Equation (3)) represents the probability of obtaining granules of a daughter particle with the size d after the breakage of a granule of size m. The breakage function considers binary breakage as well as attrition [36, 32], as shown in Equation (3).

$$b(m,d) = f_{\text{prim}}\,\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{\left(d^{\frac{1}{3}}-\mu\right)^2}{2\sigma^2}}\frac{m}{\mu^3}\frac{1}{3d^{\frac{2}{3}}} + (1-f_{\text{prim}})\,\frac{2}{m}. \tag{3}$$

In Equation (3), f_{prim} represents the ratio of granules selected to break into smaller fragments, while σ and μ denote the standard deviation and mean of the Gaussian normal distribution. μ can also be interpreted as the size of primary particles generated by attrition. Therefore, the location of the smaller particles in the resultant PSD can determine μ parameter, that is, it can be estimated from the experimental data. m and d are the volumes of the mother and daughter particles, respectively.

The cell average technique (CAT) is used to discretize and solve the integrals in Equation (1) numerically. This method assumed that each cell (or bin) in the defined domain has its representative value where all the particles of this cell are assumed to be concentrated. First, in this method, it is necessary to calculate averages of the properties of the newborn particles in a cell. Subsequently, the newborn particles are assigned according to

four possible situations (there are two possible options for the particles to be reallocated within the same cell, and there are two options for the new particles to be reallocated in the neighboring cells) to preserve the selected properties of the distribution [37, 38].

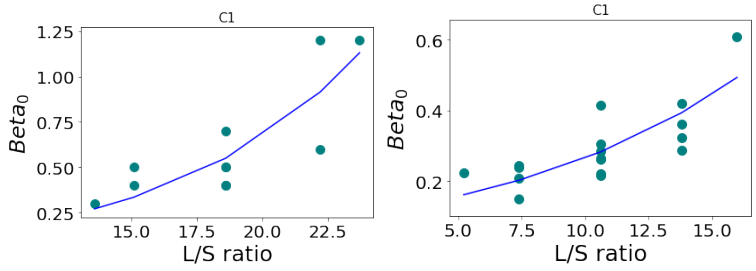
As a global stochastic optimization method, the Particle Swarm Optimization (PSO) method [39] was applied to the model calibration. Calibration was performed by compartments, experiments, and formulations in Table 1 individually. The best combination of each parameter set was chosen as the calibration results based on the minimum value of the objective function Equation (4) [40, 41].

$$D(u, v) = \left(2\mathbb{E}|X - Y| - \mathbb{E}\,|X - X'| - \mathbb{E}\,|Y - Y'|\right)^{1/2}. \tag{4}$$

In Equation (4), D represents the distance between probability distributions of the experimental data u and the computed data v. The parameters X and Y are independent random variables from the distributions of u and v, respectively. This equation is named as the energy distance in the `Python` package `Scipy` [42] and can transcribe the maximum mean discrepancy [43]. The presented PBM was tested with identifiability analysis [32], hence model calibration was performed for multiple PBM parameters simultaneously.

Two rounds of calibration were performed based on the experimental campaigns, where the number of experiments was adjusted based on the observation in the first round. Initially, calibration for APIs 1 to 4 was performed where it was calculated the parameters β_0 and *step* for the wetting zone and β_0, S_0 and f_{prim} for each of the kneading zones. The remaining unknown parameters were fixed as suggested in the authors' previous work [32]. Figure 3 and Figure 4 display the correlations of the calibrated model parameters in the wetting zone with the L/S ratio for those 50% API formulations. No clear trend emerged for the other process parameters that were varied in each DoE, therefore the L/S ratio was confirmed and remained the main process condition on which further developments are based. All the formulations exhibited the same trend. When the L/S ratio increases, the model parameter β_0 also increases. On the contrary, the *step* parameter showed a negative relationship with the L/S ratio. Increasing the L/S ratio reduces the amount of fine material remaining, therefore, a lower *step* is obtained. In other words, more successful collisions (represented by β_0) between particles occur due to the increase in interacting wet material after the formation of the initial nucleus, immediately, new larger granules can be formed and, as

a consequence, fewer non-granulated material is obtained, which is captured by the *step* parameter.



(a) API1 50%

(b) API2 50%

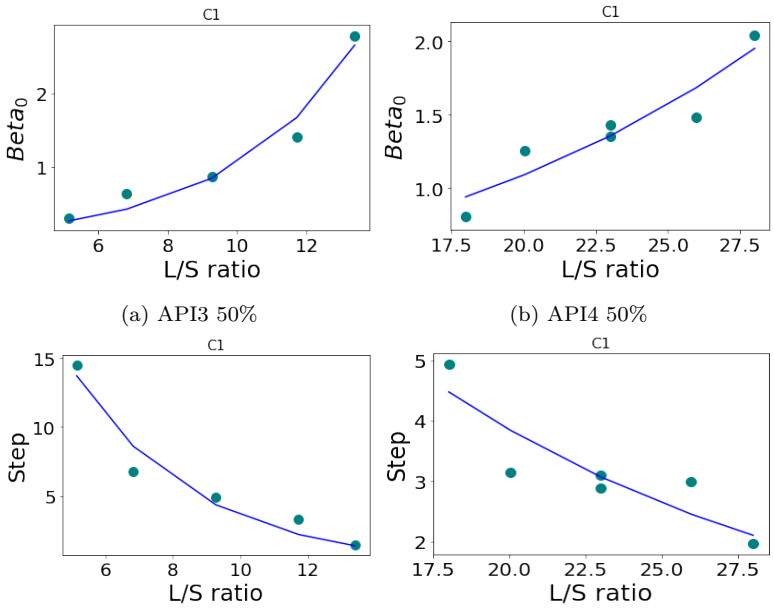
(c) API1 50%

(d) API2 50%

Figure 3: Resulting correlations for β_0 and *step* parameters in the wetting zone (C1) for 50% API formulations. The dots indicate each calibrated experiment. The blue line represents the exponential regression fit.

Figure 5 shows a comparison among the compartments in the kneading zones of the obtained correlations for each formulation for each calibrated model parameter. What is interesting about the results in this figure is that each parameter presents the same trend for all the formulations and in a similar way in each compartment. However, it can be observed that different slopes prevail from one compartment to another for each formulation. That could be attributed to the different mechanisms of formation of the nuclei and therefore the liquid distribution within the granule due to the nature of the formulation that in turn, could induce further aggregation or not in the kneading zones [7]. For the same reason, each formulation presents different scales on the y-axis of the plots (Figure 5). The results reflect the individual findings of each experiment during the calibration process for each formulation at each compartment.

The second round of the calibration stage was performed with the data



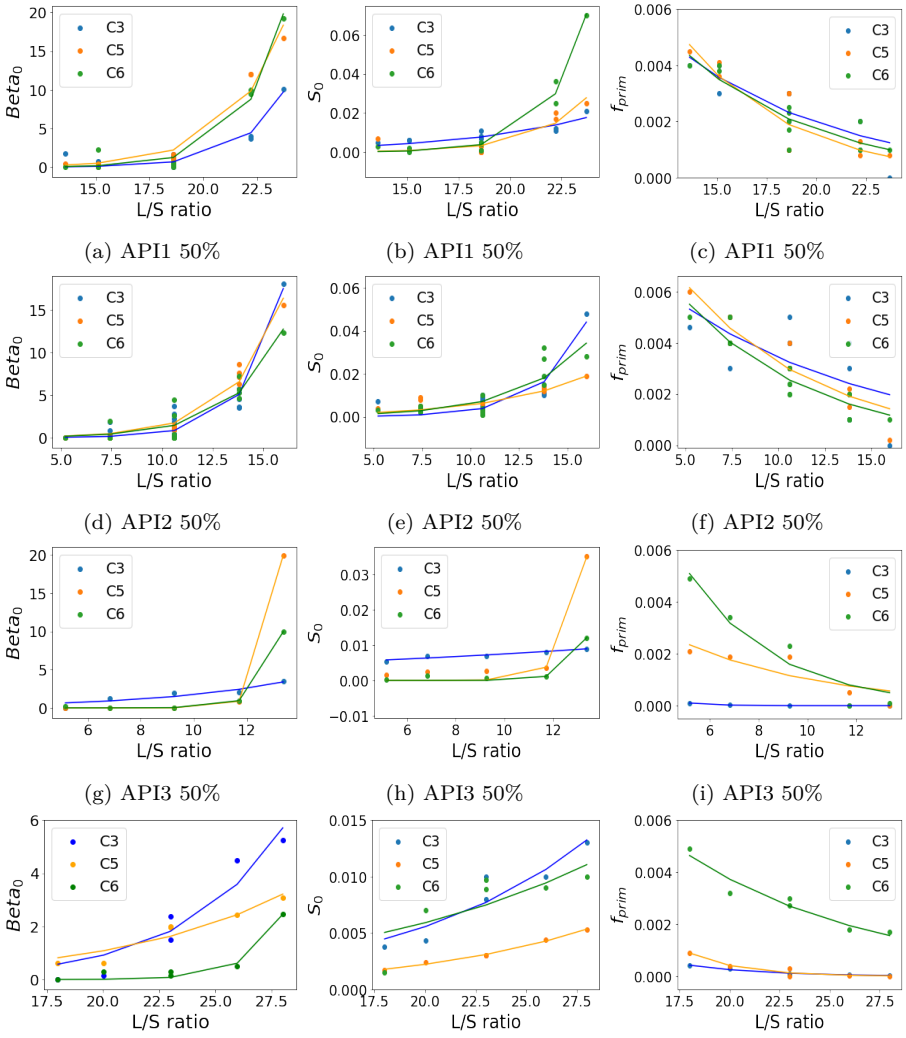
(a) API3 50%

(b) API4 50%

(c) API3 50%

(d) API4 50%

Figure 4: Resulting correlations for β_0 and *step* parameters in the wetting zone (C1) for 50% API formulations. The dots indicate each calibrated experiment. The blue line represents the exponential regression fit.



(a) API1 50% (b) API1 50% (c) API1 50%

(d) API2 50% (e) API2 50% (f) API2 50%

(g) API3 50% (h) API3 50% (i) API3 50%

(j) API4 50% (k) API4 50% (l) API4 50%

Figure 5: Comparison of the correlations for model parameters in the kneading zones (C3, C5, C6) for all 50% API formulations.

from API5, API6, and API7. Similar trends were obtained for these formulations (See Figure **??**).

2.3. PLS development

PLS models were developed to link the L/S ratio and material properties of APIs with PBM parameters. PLS regressions find latent variables (LVs) which maximize the covariance between the projections of input and output parameters. A PLS model was created for each compartment. The initial model configuration has the L/S ratio and 34 material properties as input parameters. The list of material properties is summarized in Table 2; eight parameters are material properties of a formulation while the other 26 are API material properties. The characterization techniques and equipment used for measuring material properties are listed in Table 3. Material properties can be divided into four categories, i.e., size, density, flowability, and moisture. For the parameter R, the L/S ratio was excluded from input parameters because the values of R are fixed for the same formulation regardless of the L/S ratio. By definition, R is the critical size, from which the aggregation of particles larger than this size will be favored over those smaller than that size. Thus, this parameter is directly affected and related to the initial particle size of the formulation and its nature. The PLS model of R is developed separately with other PBM parameters in the wetting zone to exclude L/S ratio from the input parameters. The outputs of each PLS model are summarized in Table 4.

The calibrated PBM parameters were transformed using the natural logarithm because some parameters have values spanning multiple orders of magnitude. Original PLS models without logarithm transformation are linear regressions, which could potentially lead to predicting negative aggregation or breakage depending on input values. In addition, the exponential relationship between L/S ratio and most of PBM parameters was confirmed in the calibration steps. This relationship is the equivalent to the log-transformation of PBM parameters. Each PLS model can be presented as shown in Equation (5):

$$\ln y_j = \sum_{i=1}^{n_{\text{mat}}} c_i x_i + c_{\text{LS}} x_{\text{LS}}, \tag{5}$$

where the parameters y_j, x_i, and x_{LS} represent a PBM parameter j, e.g., β_0, material property i, e.g., $d0,5_{\text{wF}}$, and L/S ratio, respectively. The parameters n_{mat}, c_i, and c_{LS} are the amount of material properties considered in

Table 2: List of material properties used as the initial input parameters of PLS models.

Category	Input parameters	Abbreviations
Size	50 and 90% cumulative undersize fraction of volumetric PSD of a blend (wet dispersion)	$d0,5_{\mathrm{wF}}$, $d0,9_{\mathrm{wF}}$
	Volume and surface-weighted mean particle size of a blend (wet dispersion)	$D(3,2)_{\mathrm{wF}}$, $D(4,3)_{\mathrm{wF}}$
	Span of volumetric PSD of a blend	$Span_{\mathrm{wF}}$
	50 and 90% cumulative undersize fraction of volumetric PSD of an API (wet dispersion)	$d0,5_{\mathrm{wAPI}}$, $d0,9_{\mathrm{wAPI}}$
	Volume and surface-weighted mean particle size of an API (wet dispersion)	$D(3,2)_{\mathrm{wAPI}}$, $D(4,3)_{\mathrm{wAPI}}$
	Span of volumetric PSD of an API	$Span_{\mathrm{wAPI}}$
Density	Bulk and tapped density of an API	ρ_{bulk}, ρ_{tap}
	Hausner ratio of an API	HR
	Compressibility index of an API	CI
	Conditioned bulk density of an API	CBD
Flowability	Flow function coefficient of an API	ffc
	Basic flow energy of an API	BFE
	Stability index of an API	SI
	Flow rate index of an API	FRI
	Specific energy of an API	SE
	Compressibility of an API at 15 kPa	$Cmpr$
Moisture	Water binding capacity of an API	WBC
	Residual moisture content of an API via loss on drying	LoD
	Maximum solubility of an API in water	S_{max}
	Fraction API powder dissolved after 1, 3, 5, 10, 20, 30, and 60 mins in a dissolution test	DR_1, DR_3, DR_5, DR_{10}, DR_{20}, DR_{30}, DR_{60}
	The lowest and the highest applicable L/S ratio for a formulation	L/S_{Low}, L/S_{High}
Other	API content of a formulation	C_{API}

Table 3: Characterization techniques used for measuring material properties.

Technique	Equipment	Material properties
Laser diffraction	Mastersizer® S (Malvern Instruments, Worcestershire, UK)	$d0,5_{\text{wF}}$, $d0,9_{\text{wF}}$, $D(3,2)_{\text{wF}}$, $D(4,3)_{\text{wF}}$, $Span_{\text{wF}}$, $d0,5_{\text{wAPI}}$, $d0,9_{\text{wAPI}}$, $D(3,2)_{\text{wAPI}}$, $D(4,3)_{\text{wAPI}}$, $Span_{\text{wAPI}}$
Density	Tapping device (J. Engelsman, Ludwigshafen, Germany)	ρ_{bulk}, ρ_{tap}, HR, CI
Powder rheology	FT4 powder rheometer (Freeman Technology, Tewkesbury, UK)	CBD, BFE, SI, FRI,SE, $Cmpr$
Ring shear test	Ring shear tester, RST-XS (Dietmar Schulze Schüttgutmesstechnik, Wolfenbüttel, Germany)	ffc
Water binding capacity	Heraeus Multifuge 3 S-R (Thermo Scientific, USA)	WBC
Loss on drying	HC103 Halogen Moisture Analyzer (Mettler-Toledo, Zaventem, Belgium).	LoD
Solubility	Cellulose-based filter (Grade 2, Whatman, USA) UV-1650 PC (Shimadzu, Suzhou New District, China)	S_{max}
Dissolution rate	USP4 Flow-Through Dissolution Systems (Sotax, Allschwil, Switzerland) UV-1650 PC (Shimadzu, Suzhou New District, China)	DR_1, DR_3, DR_5, DR_{10}, DR_{20}, DR_{30}, DR_{60}
L/S ratio	ConsiGma-25 (GEA Pharma Systems, ColletteTM, Wommelgem, Belgium)	L/S_{Low}, L/S_{High}

Table 4: List of output parameters per each PLS model.

Model number	Compartment	Parameters	Dependency of L/S ratio
1	Wetting	R	Independent
2	Wetting	$\beta_{0,1}$, $step$	Dependent
3	C3	$\beta_{0,3}$, $S_{0,3}$, $f_{\text{prim, 3}}$	Dependent
4	C5	$\beta_{0,5}$, $S_{0,5}$, $f_{\text{prim, 5}}$	Dependent
5	C6	$\beta_{0,6}$, $S_{0,6}$, $f_{\text{prim, 6}}$	Dependent

the model, PLS regression coefficients of material property i and L/S ratio, respectively. Since material properties are fixed for the same formulation, Equation (6) can be described for each formulation:

$$\ln y_j = a_h + c_{\mathrm{LS}} x_{\mathrm{LS}}, \tag{6}$$

where a_h represents the intercept of a PLS regression for formulation h and is equal to $\sum c_i x_i$ in Equation (5). These equations can reflect on the exponential effect of L/S ratio as shown in Figures 3 and 5. Prior to PLS model development, all input and output data were normalized by mean-centering and scaling to unit standard deviation.

During the PLS development step, the number of input parameters was reduced to avoid overfitting and increase the applicability of the models in the pharmaceutical industry. The industry focuses only on identifying relevant characteristics, so less critical properties for an envisioned prediction accuracy should be excluded from the models based on analysis. This could lead to accurate predictions with reduced training data. By changing input parameter combinations, we aimed to find the best combination that maximizes R^2 values in cross-validation (CV) for each model. The number of possible combinations for each model is more than 17 billion (2^{34}), which is computationally expensive. To address this, we reduced input parameters using the procedure in Figure 6. Among all material properties used as inputs for the model (initially 34), the least relevant property was excluded for the next round to avoid excluding critical properties. This step was repeated until all input properties were judged as relevant. Leave-One-Out Cross-Validation (LOOCV) was used as the CV method for R. For other PBM parameters, all data using the same formulation were excluded as test data instead of one dataset, whereas the framework of CV was the same as LOOCV. This enabled R^2 values to show predictability for new formulations. The number of LVs varied from one to five and was chosen based on R^2 CV.

The two steps of this pre-processing approach, i.e., log-transformation and reduction of the number of input parameters, were critical to enable the models to predict PBM parameters appropriately. For clarity, the results of PLS models without pre-processing as well as with different pre-processing approaches are presented in the supplementary material. In addition to CV, validation of R was made by using the model for four new APIs. The primary focus of this work is to identify key material properties on PBM parameters by developing PLS models. A comprehensive account of the validation phase

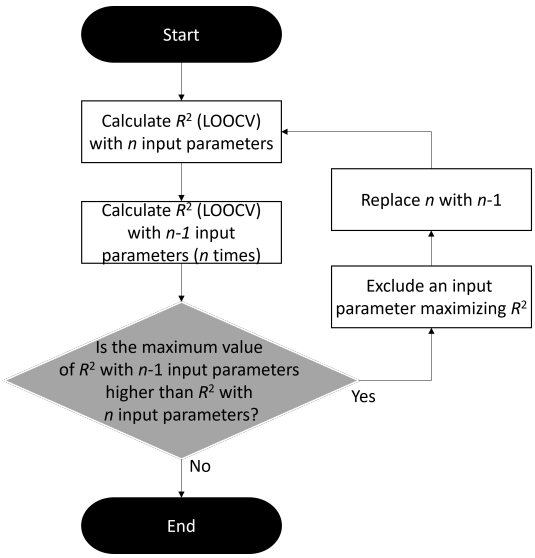


Figure 6: Flowchart of extracting material properties used for the PLS models.

for the hybrid model, consisting of the PBM component and the proposed PLS models, can be found in Barrera Jiménez et al. 2023 [33]. All works of PLS development were performed using the `Python` package `scikit-learn`.

To quantify the uncertainties of the developed PLS models, confidence intervals ($CInt$) were also calculated based on Hotelling T-squared values (T^2), as shown in Equation (7) [44]:

$$CInt = t_{n_{inp}-1,\alpha} \cdot RMSE\sqrt{1+T^2}, \tag{7}$$

where $t_{n_{inp}-1,\alpha}$ represents the percent point function of Student's T distributions with $n_{inp}-1$ degrees of freedom at significance level α. The parameters n_{inp} and $RMSE$ represent the number of data used for training of PLS models and root mean square errors of PLS models in training data, respectively. In addition to R^2 values, model validation was performed by checking if experimental data were within the ranges of confidence intervals.

3. Results and Discussion

3.1. PLS model construction

The input parameters as well as the number of LV of each PLS model were determined after CV, as summarized in Table 5. The indices R^2X and

R^2Y represent the sum of the percentages of variation of the X (input) and Y (output) explained by each LV, respectively. The number of material properties used for the models was reduced from 34 to between six and ten based on the flowchart shown in Figure 6. Overall, 25 material properties were necessary to develop the five PLS models. Since some material properties are measured by the same characterization method, nine out of eleven material characterization methods were needed to predict all model parameters. Model 1 (R) had the smallest R^2X value due to the smaller number of LVs used and the complex underlying relationship among the material properties incorporated into the model. The values of R^2Y are related to the impacts of PBM parameters on the granule size distribution. The model parameters in the wetting zone (Model 2) have a high impact on the whole simulation, due to the influence of the bimodality in that compartment. On the other hand, the parameters in C3 (Model 3) change granule size distribution towards intermediate sizes slightly.

Table 5: Summary of the developed PLS models.

Model	Number of LV	Selected material properties	R^2X	R^2Y
1	2	$d0,5_{\mathrm{wF}}$, $D(4,3)_{\mathrm{wF}}$, $Span_{\mathrm{wAPI}}$, ρ_{bulk}, SI, $Cmpr$, DR_{10}, L/S_{High}	0.555	0.901
2	4	$Span_{\mathrm{wF}}$, $Span_{\mathrm{wAPI}}$, ρ_{tap}, HR, ffc, S_{max}, DR_1, DR_5	0.916	0.887
3	3	$D(4,3)_{\mathrm{wF}}$, $d0,5_{\mathrm{wAPI}}$, $d0,9_{\mathrm{wAPI}}$, $D(3,2)_{\mathrm{wAPI}}$, $D(4,3)_{\mathrm{wAPI}}$, ρ_{tap}, SI, FRI, L/S_{High}	0.879	0.273
4	5	$d0,9_{\mathrm{wF}}$, ρ_{tap}, HR, CI, SE, L/S_{High}	0.955	0.633
5	3	$D(3,2)_{\mathrm{wAPI}}$, ρ_{bulk}, HR, CI, CBD, BFE, SI, FRI, SE, L/S_{High}	0.626	0.431

3.2. PLS model validation

Based on the selected LV numbers and material properties, the PLS models were validated. The developed PLS models were used to compute the PBM parameters of the experimental data. The accuracy of all of the developed PLS models is summarized in Table 6, where R^2 and root mean square error ($RMSE$) were used. The results of fitting and prediction of R with 90% confidence intervals are presented in Figure 7; where predicted values

were calculated based on LOOCV. Figure 7 shows that the prediction as well as the fitting could capture the general trends of R. Observed values of R were within the ranges of 90% confidence intervals for all formulations except for Formulation 6. Particle sizes, which should be important for R, of Formulation 6 were much smaller than the other formulations. While smaller particles of Formulation 6 resulted in the lowest R, the PLS model over-predicted because of limited data of formulations consisting of smaller particles.

The prediction results of the other PBM parameters with 90% confidence intervals are presented in Figure 8. For the comparison, the fitting results of the PBM parameters are presented in Figure S3 in the supplementary material. For the wetting zone, R^2-values were quite high for fitting and remained high for CV as well, which fulfilled the model requirement sufficiently. This fact is crucial because PBM parameters in the wetting zone have the highest impacts on the final granule size distributions according to the results presented in [32].

While the computation accuracy in the kneading zones is lower, all PLS models capture the general trends of the PBM parameters. The CV results in Table 6 were not sufficient for accurate estimations of PBM parameters for new formulations. The values of R^2 in the CV were negative for most of the PBM parameters in the kneading zone whereas they were the highest among possible combinations of material properties as inputs. Negative R^2 values indicate the insufficiency of the models for computing the values of PBM parameters for new formulations. This is due to either missing critical material properties in the measurement or the necessity of nonlinear modeling. On the other hand, fitting results as well as the fact that the models captured the general trends suggest that the models are sufficient for interpreting data. Two experimental data points in C6 had high prediction errors, as they originated from API 5 (Formulation 8) where seven out of ten input material properties had either the highest or the lowest values among the ten formulations. The predicted f_{prim} for these data points was significantly higher. Since CV always excluded data to be validated from training data, the simulation of PBM parameters in C6 for Formulation 8 required extrapolation of the trained PLS. When all input data was used, the fitting results of Formulation 8 were much better, as shown in the supplementary material. A similar tendency can be observed for C3 and S_0 in C5. Except for them, no data points were showing inverse trends between predicted and observed values. The potential causes of the lower prediction are calibration

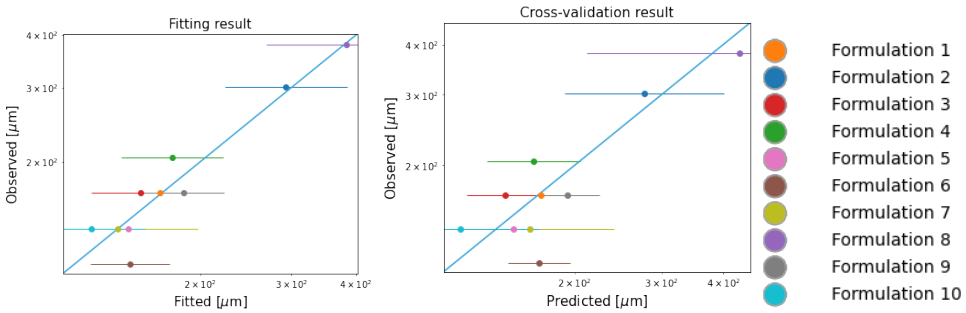
of PBM parameters in the kneading zone, which in turn can be affected by experimental uncertainty and linear regression of the models.

These causes also explain wide ranges of confidence intervals especially in the kneading zones and for specific formulations (e.g., Formulation 8). The calibrated values of the PBM parameters in the wetting zone were within the ranges of confidence intervals, which confirmed the accuracy and appropriateness of the developed PLS model. On the other hand, in the CV results of the PBM parameters in the kneading zones, some calibrated values were outside of confidence intervals, and other confidence intervals were too broad. The impact of experimental uncertainty could be large for $f_{\text{prim, 6}}$ since it is not sensitive to granule size distribution compared to other PBM parameters. The validation of PLS models was also tried without pre-processing methods (i.e., log transformation and the reduction of material properties), as shown in the supplementary material. The prediction accuracy was found to be much lower, and sometimes the models even predicted physically unrealistic negative values for PBM parameters. These results proved the validity of applying both pre-processing methods.

Table 6: The overview of the PLS model performances.

		Fitting		CV	
Model	PBM parameter	R^2	$RMSE$	R^2	$RMSE$
1	R	0.906	0.109	0.740	0.181
2	β_0	0.876	0.339	0.724	0.505
	$step$	0.899	0.408	0.707	0.695
3	β_0	0.376	0.881	-2.265	2.017
	S_0	0.254	0.481	-4.565	1.314
	f_{prim}	0.188	1.478	-2.651	3.135
4	β_0	0.726	0.667	0.434	0.959
	S_0	0.463	0.750	-0.146	1.096
	f_{prim}	0.709	0.653	-0.004	1.213
5	β_0	0.422	1.274	-5.888	4.399
	S_0	0.590	0.831	-3.744	2.827
	f_{prim}	0.281	1.423	-12.07	6.070

The obtained PLS model was used to predict the R parameter of new formulations, which were not included in the training data (Table 7). Figure 9 shows a comparison between the predicted R parameter and experimental



(a) Fitting (b) CV

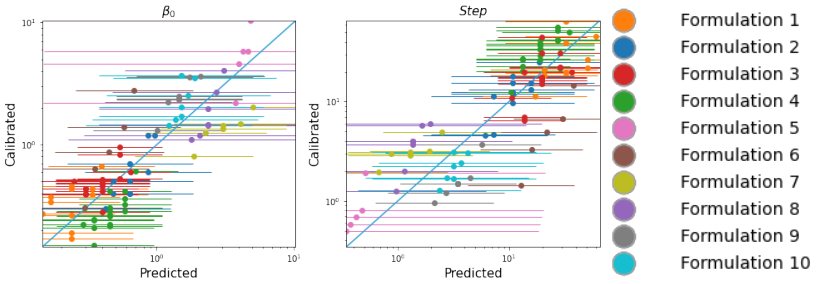
Figure 7: Fitted, predicted, and observed values of R with 90% confidence interval.

observations. The vertical lines in the figure represent the μ parameter defined in Equation 3. In this work, we followed a strategy presented in [32] to reduce the number of parameters that need calibration. Specifically, we set μ to be the same as R defined in the wetting zone. By identifying μ in the kneading zones, we can obtain the R value for the wetting zone. Therefore, it is not necessary to collect data from the wetting zone (C1) to validate the model. This approach significantly reduces the experimental work and maintains a focus on the model's industrial applicability. The predicted R values were close to the experimentally observed μ values, and the prediction errors in the proposed PLS model were lower than those in the PLS models without pre-processing (see the supplementary material). In the conventional PBM, the identification of R values requires multiple experiments with different process settings. The proposed PLS enabled designers to calculate R values without experiments.

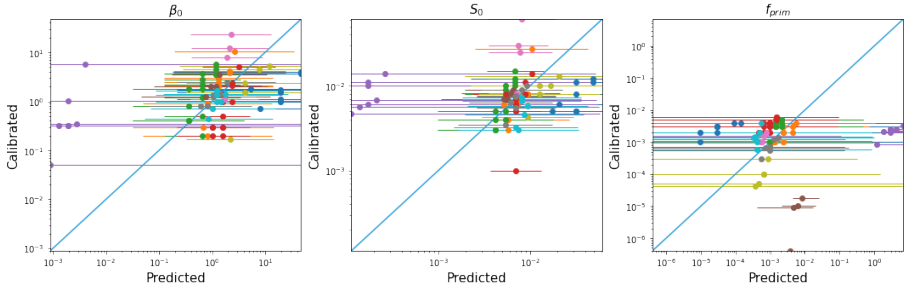
Table 7: Process conditions of validation formulations.

API	Number of experiments	Screw Speed (rpm)	Throughput (kg/h)	L/S (%)
API 8 (50 %)	5	675	20	20.127.3
API 9 (50 %)	5	675	20	14.827.2
API 10 (50 %)	5	675	20	17.023.0
API 11 (50 %)	3	675	20	5.97.0

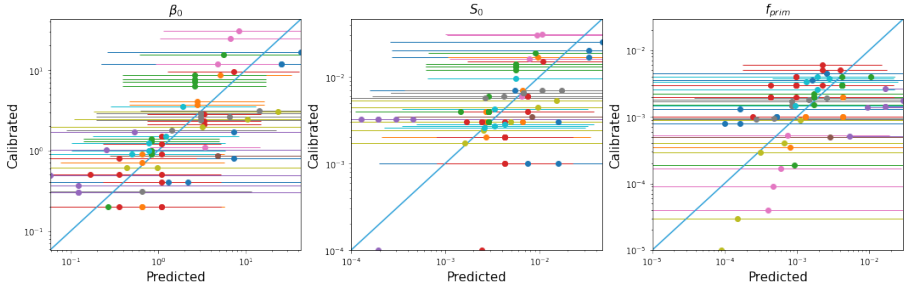
Figure 10 presents a comparison of the experimental and predicted R/μ-values for each validation formulation. The simulated PSDs were obtained



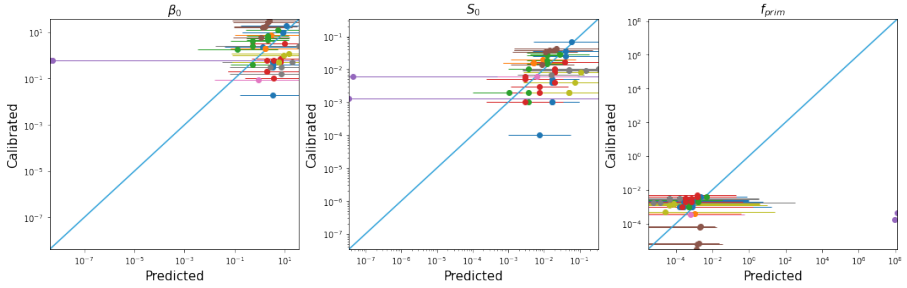
(a) Wetting zone



(b) C3

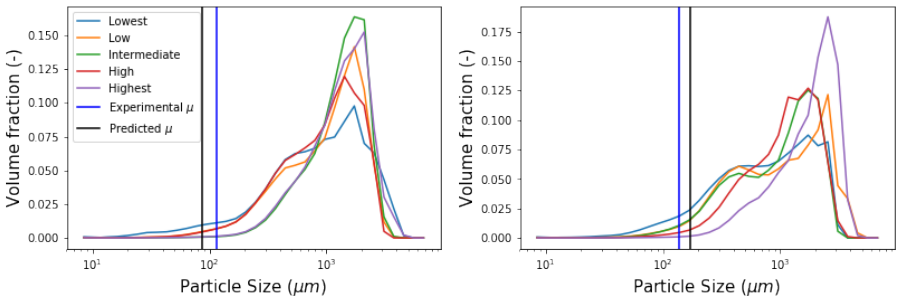


(c) C5



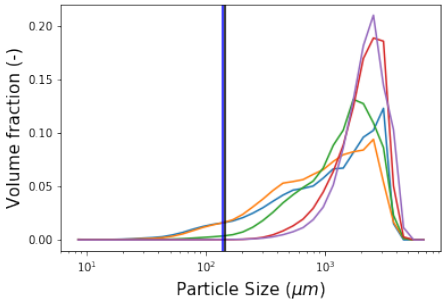
(d) C6

Figure 8: Predicted PBM parameters from material properties vs calibrated PBM parameters in the wetting (C1) and the kneading (C3, C5, and C6) zones with 90% confidence interval.

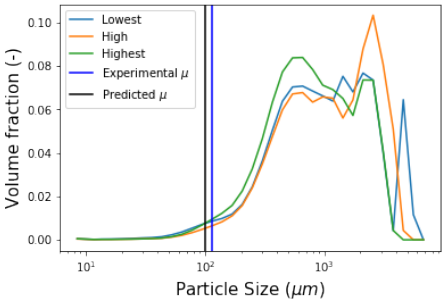


(a) API 8

(b) API 9



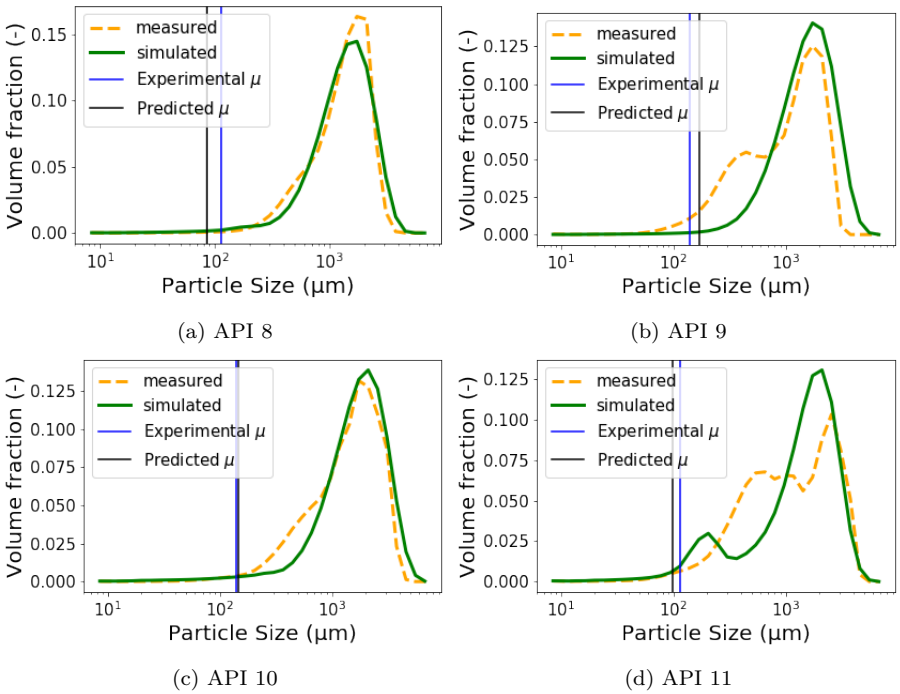
(c) API 10



(d) API 11

Figure 9: Predicted R for four new APIs and its comparison with the experimental R-value. Here, μ represents R-values. For each formulation, multiple size distributions are presented from the lowest to the highest L/S ratio.

using the values predicted by the PLS models to calculate the PBM parameters. In our previous publication [33], it was also validated that the computed PBM parameters can be used for the simulation of granule size distributions for new formulations.



(a) API 8

(b) API 9

(c) API 10

(d) API 11

Figure 10: Predicted and experimental R/μ for the four validation formulations against the simulated and experimental Particle size distribution at the intermediate L/S ratio for each formulation.

As described above, some improvement opportunities in the model development were found from the validation results. Model calibration and linearity assumption resulted in lower prediction and wide confidence intervals of PBM parameters in the kneading zones. Advanced analyses of uncertainty and relationships among model parameters, e.g, application of the Monte Carlo method for uncertainty analysis [45], could improve the model performance and uncertainty quantification.

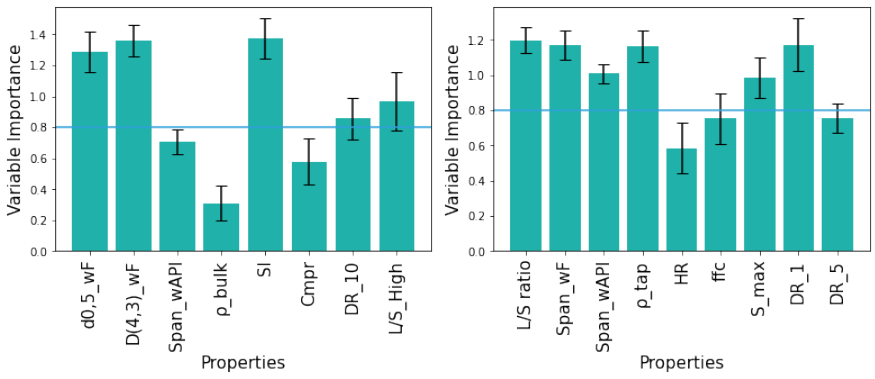
3.3. PLS model interpretation

Key material properties were further interpreted by Variable Importance for the Projection (VIP) plotting. VIP represents the impact of each in-

put parameter on the PLS models, where critical input parameters have large values of VIP. To analyze the accuracy and reliability of the VIP results, the standard deviations of the VIP scores were calculated based on the VIP results in the CV procedure. The VIP plots of the PLS models are shown in Figure 11, where error bars show the standard deviations of VIP scores. The standard deviations were not so large for most of the parameters that the results of key material properties and process settings were stable. Several material properties have similar variable importance with the L/S ratio in the wetting zone while the L/S ratio is the highest influential factor in the kneading zone. By comparing the VIP plots of different PLS models, the differences in key material properties can be interpreted. For example, size-related properties (e.g., $d0,5_{\mathrm{wF}}$, $D(3,2)_{\mathrm{wAPI}}$, and $Span_{\mathrm{wAPI}}$) affected R and the model parameters in the upstream zones more than the model parameters in the downstream zones. In addition, moisture-related parameters (e.g., $S_{\max}$) were important in the wetting zone, whereas density-related parameters (CI and HR) proved to be more important in C5 and C6. Density-related parameters (CI and HR) were critical in C5, which is in agreement with experimental observations. According to the experimental observations, a shift in the distribution after the second kneading zone is related to the compaction of large granules [46, 47]. Therefore, it is reasonable that density-related parameters prevailed as key parameters in the later kneading zones.

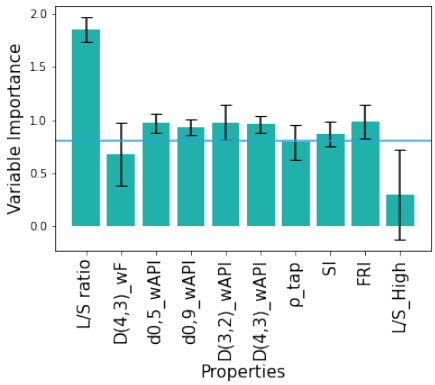
Figure S5 in the supplementary material shows the PLS regression coefficients of all PBM parameters. The data show that β_0 and S_0 have positive correlations, while β_0 and *step* or f_{prim} have negative correlations. By comparing the same PBM parameters in the different zones, several interesting findings were observed. For β_0, moisture-related properties had a similar tendency of the impacts, i.e., lower L/S_{High} increased β_0 regardless of the zones. This corroborates the differences obtained for each formulation in terms of the L/S ratio range, to obtain similar granules and allow a fair comparison between the formulations studied [34, 48]. In addition, this shows that the PBM model parameters can be linked to the nature of the material.

The proposed approach enabled the computation of PBM parameters for new drugs based on identified key material properties. The information is useful to reduce experiments in development and understand the phenomena of wet granulation deeply. Moreover, the proposed PLS models visualized different impacts of material properties and L/S ratio on aggregation and breakage in different compartments. The results help to understand the role

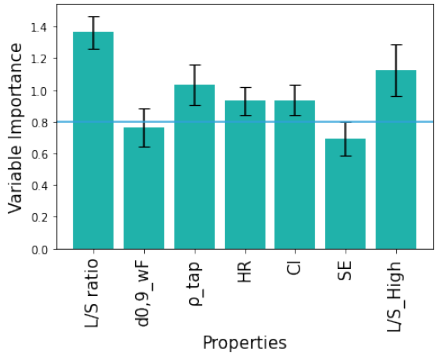


(a) R

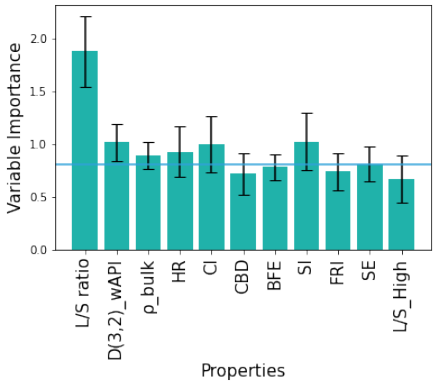
(b) Wetting zone



(c) C3



(d) C5



(e) C6

Figure 11: VIP plots of the developed PLS models with standard deviations obtained in CV.

of each compartment deeply and improve the design of screws and equipment. The proposed approach can be extended to other fields, where the links between available mechanistic models and material properties are missing. On the other hand, the prediction accuracy can be improved especially for the PBM parameters in the kneading zones. Possible causes of lower prediction accuracy are lack of non-linearity due to the PLS method and low accuracy of calibrated PBM parameters. One solution is using other data-driven models which can reflect non-linear relationships between input and output parameters. Improvement of the 1D-PBM model as well as the expansion of data sets could also lead to high prediction accuracy.

4. Conclusions

The presented PLS models could assess the impact of material properties on PBM parameters in the context of continuous TSWG. Five PLS models were developed (one for each compartment) after the log-transformation of PBM parameters and the reduction of the number of input parameters. The developed models showed sufficient fitting accuracy for interpretation and had predictive power for the PBM parameters in the wetting zone. Successful PLS development was not possible without comprehensive experimental design, the PBM having fewer model parameters through identifiability analysis, and pre-processing of model parameters. This study is the first attempt to determine the values of the PBM parameters based on material properties and process parameters.

Through VIP plots and PLS regression coefficients, the key material properties were observed by compartments and PBM parameters. Size- and moisture-related properties were influential on the upstream zones, whereas density-related properties showed a significant impact on the downstream zones. These insights can reduce experiments and material characterization significantly for the process design of new drugs.

In the future, the prediction accuracy of the proposed approach can be further improved by using other data-driven models such as non-linear regression models. A hybrid model of data-driven models and the PBM [32] can be used for the simulation of granule size distributions for new drugs. The model has a generic nature under development with the same excipient base and produced with the same screw configuration [33]. The model applicability can be extended by adding other factors, e.g., different excipient bases and screw configurations, as inputs of PLS models. A deeper

analysis of the impacts of material properties is valuable for the industry, e.g., the application of sensitivity analysis. Furthermore, an advanced uncertainty analysis through stochastic approaches is expected to narrow down confidence intervals and clarify the causes of model uncertainty.

Acknowledgement

The authors would like to acknowledge, in no particular order, Janssen Pharmaceutica and UCB for their financial support and fruitful collaboration in this project.

References

[1] A. Kumar, J. Vercruysse, S. T. Mortier, C. Vervaet, J. P. Remon, K. V. Gernaey, T. De Beer, I. Nopens, Model-based analysis of a twin-screw wet granulation system for continuous solid dosage manufacturing, Computers and Chemical Engineering 89 (2016). doi:10.1016/j.compchemeng.2016.03.007.

[2] D. Barrasso, R. Ramachandran, Qualitative Assessment of a Multi-Scale, Compartmental PBM-DEM Model of a Continuous Twin-Screw Wet Granulation Process, Journal of Pharmaceutical Innovation 11 (3) (2016) 231–249. doi:10.1007/s12247-015-9240-7.

[3] S. M. Iveson, J. D. Litster, B. J. Ennis, Fundamental studies of granule consolidation part 1: Effects of binder content and binder viscosity, Powder Technology (1996). doi:10.1016/0032-5910(96)03096-3.

[4] S. M. Iveson, J. D. Litster, Fundamental studies of granule consolidation. Part 2: Quantifying the effects of particle and binder properties, Powder Technology (1998). doi:10.1016/S0032-5910(98)00116-8.

[5] R. M. Dhenge, R. S. Fyles, J. J. Cartwright, D. G. Doughty, M. J. Hounslow, A. D. Salman, Twin screw wet granulation: Granule properties, Chemical Engineering Journal (2010). doi:10.1016/j.cej.2010.05.023.

[6] A. S. El Hagrasy, J. R. Hennenkamp, M. D. Burke, J. J. Cartwright, J. D. Litster, Twin screw wet granulation: Influence of formulation parameters on granule properties and growth behavior, Powder Technology (2013). doi:10.1016/j.powtec.2012.04.035.

[7] M. Verstraeten, D. Van Hauwermeiren, K. Lee, N. Turnbull, D. Wilsdon, M. am Ende, P. Doshi, C. Vervaet, D. Brouckaert, S. T. Mortier,

I. Nopens, T. D. Beer, In-depth experimental analysis of pharmaceutical twin-screw wet granulation in view of detailed process understanding, International Journal of Pharmaceutics 529 (1-2) (2017) 678–693. doi:10.1016/j.ijpharm.2017.07.045.

[8] S. V. Lute, R. M. Dhenge, A. D. Salman, Twin screw granulation: An investigation of the effect of barrel fill level, Pharmaceutics 10 (2) (2018). doi:10.3390/pharmaceutics10020067.

[9] B. Khorsheed, I. Gabbott, G. K. Reynolds, S. C. Taylor, R. J. Roberts, A. D. Salman, Twin-screw granulation: Understanding the mechanical properties from powder to tablets, Powder Technology 341 (2019). doi:10.1016/j.powtec.2018.05.013.

[10] K. M. Hwang, C. H. Cho, S. D. Yoo, K. I. Cha, E. S. Park, Continuous twin screw granulation: Impact of the starting material properties and various process parameters, Powder Technology 356 (2019) 847–857. doi:10.1016/j.powtec.2019.08.062.

[11] C. Portier, T. Vigh, G. Di Pretoro, T. De Beer, C. Vervaet, V. Vanhoorne, Continuous twin screw granulation: Impact of binder addition method and surfactants on granulation of a high-dosed, poorly soluble API, International Journal of Pharmaceutics 577 (mar 2020). doi:10.1016/j.ijpharm.2020.119068.

[12] S. M. Iveson, J. D. Litster, K. Hapgood, B. J. Ennis, Nucleation, growth and breakage phenomena in agitated wet granulation processes: a review, Powder Technology 117 (1-2) (2001) 3–39. doi:10.1016/S0032-5910(01)00313-8.

[13] K. F. Lee, S. Mosbach, M. Kraft, W. Wagner, A multi-compartment population balance model for high shear granulation, Computers and Chemical Engineering 75 (2015) 1–13. doi:10.1016/j.compchemeng.2015.01.009.

[14] S. Shirazian, H. Y. Ismail, M. Singh, R. Shaikh, D. M. Croker, G. M. Walker, Multi-dimensional population balance modelling of pharmaceutical formulations for continuous twin-screw wet granulation: Determination of liquid distribution, International Journal of Pharmaceutics 566 (June) (2019) 352–360. doi:10.1016/j.ijpharm.2019.06.001.

[15] A. Chaudhury, A. Kapadia, A. V. Prakash, D. Barrasso, R. Ramachandran, An extended cell-average technique for a multi-dimensional population balance of granulation describing aggregation and breakage, Advanced Powder Technology 24 (6) (2013) 962–971. doi:10.1016/j.apt.2013.01.006.

[16] H. Liu, S. C. Galbraith, S. Y. Park, B. Cha, Z. Huang, R. F. Meyer, M. H. Flamm, T. O'Connor, S. Lee, S. Yoon, Assessment of spatial heterogeneity in continuous twin screw wet granulation process using three-compartmental population balance model (2018). doi:10.1080/10837450.2018.1427106.

[17] N. Metta, M. Verstraeten, M. Ghijs, A. Kumar, E. Schafer, R. Singh, T. D. Beer, I. Nopens, P. Cappuyns, I. V. Assche, Model development and prediction of particle size distribution , density and friability of a comilling operation in a continuous pharmaceutical manufacturing process, International Journal of Pharmaceutics 549 (1-2) (2018) 271–282. doi:10.1016/j.ijpharm.2018.07.056.

[18] D. Barrasso, S. Walia, R. Ramachandran, Multi-component population balance modeling of continuous granulation processes : A parametric study and comparison with experimental trends, Powder Technology 241 (2013) 85–97. doi:10.1016/j.powtec.2013.03.001.

[19] D. Barrasso, T. Eppinger, F. E. Pereira, R. Aglave, K. Debus, S. K. Bermingham, R. Ramachandran, A multi-scale, mechanistic model of a wet granulation process using a novel bi-directional PBM-DEM coupling algorithm, Chemical Engineering Science 123 (2015) 500–513. doi:10.1016/j.ces.2014.11.011.

[20] A. D. Mcguire, S. Mosbach, K. F. Lee, G. Reynolds, M. Kraft, A high-dimensional, stochastic model for twin-screw granulation Part 1 : Model description, Chemical Engineering Science 188 (2018) 221–237.

[21] A. D. McGuire, S. Mosbach, K. F. Lee, G. Reynolds, M. Kraft, A high-dimensional, stochastic model for twin-screw granulation Part 2: Numerical methodology, Chemical Engineering Science 188 (2018) 18–33. doi:10.1016/j.ces.2018.04.077.

[22] S. Shirazian, S. Darwish, M. Kuhs, D. M. Croker, G. M. Walker, Regime-separated approach for population balance modelling of continuous wet granulation of pharmaceutical formulations, Powder Technology 325 (2018) 420–428. doi:10.1016/j.powtec.2017.11.047.

[23] H. Liu, S. C. Galbraith, S. Y. Park, B. Cha, Z. Huang, R. F. Meyer, M. H. Flamm, T. O'Connor, S. Lee, S. Yoon, Assessment of spatial heterogeneity in continuous twin screw wet granulation process using three-compartmental population balance model, Pharmaceutical Development and Technology 24 (1) (2019) 105–117. doi:10.1080/10837450.2018.1427106.

[24] H. Y. Ismail, M. Singh, A. B. Albadarin, G. M. Walker, Complete two dimensional population balance modelling of wet granulation in twin screw, International Journal of Pharmaceutics 591 (2020) 120018. doi:10.1016/j.ijpharm.2020.120018.

[25] L. G. Wang, J. P. Morrissey, D. Barrasso, D. Slade, S. Clifford, G. Reynolds, J. Y. Ooi, J. D. Litster, Model driven design for twin screw granulation using mechanistic-based population balance model, International Journal of Pharmaceutics 607 (2021) 120939. doi:10.1016/j.ijpharm.2021.120939.

[26] A. Ryckaert, D. Van Hauwermeiren, J. Dhondt, A. D. Man, A. Funke, D. Djuric, C. Vervaet, I. Nopens, T. De Beer, Tpls as predictive platform for twin-screw wet granulation process and formulation development, International Journal of Pharmaceutics 605 (2021) 120785. doi:10.1016/J.IJPHARM.2021.120785.

[27] B. V. Snick, W. Grymonpré, J. Dhondt, K. Pandelaere, G. D. Pretoro, J. P. Remon, T. D. Beer, C. Vervaet, V. Vanhoorne, Impact of blend properties on die filling during tableting, International Journal of Pharmaceutics 549 (2018) 476–488. doi:10.1016/j.ijpharm.2018.08.015.

[28] M. A. Järvinen, M. Paavola, S. Poutiainen, P. Itkonen, V. Pasanen, K. Uljas, K. Leiviskä, M. Juuti, J. Ketolainen, K. Järvinen, Comparison of a continuous ring layer wet granulation process with batch high shear and fluidized bed granulation processes, Powder Technology 275 (2015) 113–120. doi:10.1016/j.powtec.2015.01.071.

[29] K. Matsunami, T. Nagato, K. Hasegawa, H. Sugiyama, Determining key parameters of continuous wet granulation for tablet quality and productivity: A case in ethenzamide, International Journal of Pharmaceutics 579 (2020) 119160. doi:10.1016/j.ijpharm.2020.119160.

[30] P. Kazemi, M. H. Khalid, A. P. Gago, P. Kleinebudde, R. Jachowicz, J. Szlk, A. Mendyk, Effect of roll compaction on granule size distribution of microcrystalline cellulosemannitol mixtures: computational intelligence modeling and parametric analysis, Drug Design, Development and Therapy 11 (2017) 241. doi:10.2147/DDDT.S124670.

[31] D. Van Hauwermeiren, M. Stock, T. De Beer, I. Nopens, Predicting pharmaceutical particle size distributions using kernel mean embedding, Pharmaceutics 12 (2020) 271. doi:10.3390/pharmaceutics12030271.

[32] A. A. Barrera Jiménez, D. Van Hauwermeiren, M. Peeters, T. De Beer, I. Nopens, Improvement of a 1D Population Balance Model for Twin-Screw Wet Granulation by Using Identifiability Analysis, Pharmaceutics 13 (692) (2021) 1–22.

[33] A. A. Barrera Jiménez, K. Matsunami, M. Peeters, T. De Beer, I. Nopens, Linking material properties to 1D-PBM parameters towards a generic model for twin-screw wet granulation, Chemical Engineering Research and Design (2023). doi:https://doi.org/10.1016/j.cherd.2023.04.009.

[34] C. Portier, C. Vervaet, V. Vanhoorne, Continuous twin screw granulation: A review of recent progress and opportunities in formulation and equipment design, Pharmaceutics 13 (2021) 668. doi:10.3390/pharmaceutics13050668.

[35] D. Ramkrishna, Theory and Applications to Particulate systems in Engineering, Academic Press, Orlando Florida, 2000.

[36] D. Van Hauwermeiren, M. Verstraeten, P. Doshi, M. T. Ende, N. Turnbull, K. Lee, T. D. Beer, I. Nopens, On the modelling of granule size distributions in twin-screw wet granulation: calibration of a novel compartmental population balance model, Powder Technology (2018) 116–125doi:10.1016/j.powtec.2018.05.025.
URL `https://doi.org/10.1016/j.powtec.2018.05.025`

[37] J. Kumar, Numerical approximations of population balance equations in particulate systems, Ph.D. thesis, Otto-von-Guericke-University Magdeburg (2006).

[38] J. Kumar, G. Warnecke, Convergence analysis of sectional methods for solving breakage population balance equations-II: The cell average technique, Numerische Mathematik 110 (4) (2008) 539–559. doi:10.1007/s00211-008-0173-7.

[39] K. Parsopoulos, M. Vrahatis, Recent approaches to global optimization problems through Particle Swarm Optimization, Natural computing 1 (2-3) (2002) 235–306. doi:10.1023/A:1016568309421.

[40] G. J. Székely, M. L. Rizzo, Energy statistics: A class of statistics based on distances, Journal of Statistical Planning and Inference 143 (8) (2013) 1249–1272. doi:10.1016/j.jspi.2013.03.018.

[41] M. L. Rizzo, G. J. Székely, Energy distance, Wiley Interdisciplinary Reviews: Computational Statistics 8 (1) (2016) 27–38. doi:10.1002/wics.1375.

[42] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, . Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar, A. P. Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C. N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D. A. Nicholson, D. R. Hagen, D. V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G. A. Price, G. L. Ingold, G. E. Allen, G. R. Lee, H. Audren, I. Probst, J. P. Dietrich, J. Silterra, J. T. Webber, J. Slavi, J. Nothman, J. Buchner, J. Kulick, J. L. Schönberger, J. V. de Miranda Cardoso, J. Reimer, J. Harrington, J. L. C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N. J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P. A. Brodtkorb, P. Lee, R. T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna,

S. Peterson, S. More, T. Pudlik, T. Oshima, T. J. Pingel, T. P. Robitaille, T. Spura, T. R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y. O. Halchenko, Y. Vázquez-Baeza, Scipy 1.0: fundamental algorithms for scientific computing in python, Nature Methods 2020 17:3 17 (2020) 261–272. doi:10.1038/s41592-019-0686-2.

[43] D. Van Hauwermeiren, on the Simulation of Particle Size Distributions in Continuous Pharmaceutical Wet Granulation, Ph.D. thesis, Ghent University, Belgium (2019).
URL https://biblio.ugent.be/publication/8667346/file/8667347

[44] P. Nomikos, J. F. MacGregor, Multi-way partial least squares in monitoring batch processes, Chemometrics and Intelligent Laboratory Systems 30 (1995) 97–108. doi:10.1016/0169-7439(95)00043-7.

[45] Z. Nagy, R. Braatz, Distributional uncertainty analysis using power series and polynomial chaos expansions, Journal of Process Control 17 (3) (2007) 229–240, special Issue ADCHEM 2006 Symposium. doi:https://doi.org/10.1016/j.jprocont.2006.10.008.
URL https://www.sciencedirect.com/science/article/pii/S0959152406001119

[46] S. V. Lute, R. M. Dhenge, M. J. Hounslow, A. D. Salman, Twin screw granulation: Understanding the mechanism of granule formation along the barrel length, Chemical Engineering Research and Design 110 (2016) 43–53. doi:10.1016/j.cherd.2016.03.008.

[47] M. Verstraeten, Model-based process analytical technology in continuous manufacturing, Ph.D. thesis, Ghent University, Belgium (2018).

[48] L. Vandevivere, E. Van Wijmeersch, O. Häusler, T. De Beer, C. Vervaet, V. Vanhoorne, The effect of screw configuration and formulation variables on liquid requirements and granule quality in a continuous twin screw wet granulation process, Journal of Drug Delivery Science and Technology 68 (September 2021) (2022) 103042. doi:10.1016/j.jddst.2021.103042.