Preprint Cambridge Centre for Computational Chemical Engineering ISSN 1473 – 4273

Impact of powder dispersion on a wet granulation system

Catharine A. Kastner, Jethro Akroyd, Sebastian Mosbach, Markus Kraft [1]

released: 16 October 2013

[1] Department of Chemical Engineering
and Biotechnology
University of Cambridge
New Museums Site
Pembroke Street
Cambridge, CB2 3RA
United Kingdom
E-mail: mk306@cam.ac.uk

Preprint No. 135

Keywords: Particle granulation, modelling, experimental characterisation, population balance equation, process control

Edited by
Computational Modelling Group
Department of Chemical Engineering and Biotechnology
University of Cambridge
New Museums Site
Pembroke Street
Cambridge CB2 3RA
United Kingdom

Fax: + 44 (0)1223 334796
E-Mail: mk306@cam.ac.uk
World Wide Web: http://como.cheng.cam.ac.uk/

Abstract

In this paper we present an experimental investigation of the effects of the shape of the initial powder distribution on a granulation system. Experimental data was produced by granulating designed powder distributions of lactose monohydrate with deionised water acting as the binder. Three initial powder distributions were constructed from sieve cuts such that the mean volume particle diameters are similar while the shapes of the distributions differ. The initial distribution shapes consist of a narrow unimodal distribution, a wide unimodal distribution and a wide bimodal distribution. The resulting product distributions are characterised and compared using a variety of statistical metrics. The product distributions are found to show markedly different characteristics. Increasing the variance while maintaining a unimodal initial powder distribution produced a broader, more evenly distributed end-product with the mode of the resulting distribution being the same as the narrow distribution case. In comparison to the unimodal cases, the bimodal distribution produced an end-product with an even larger variance, an increase in oversized particles, a decrease in fines and a right-shift of the product distribution of one sieve class. These results were found to be independent of the methods of characterising the product distributions.

Contents

1 Introduction

Particle granulation encompasses a vast collection of agglomeration techniques wherein a fine powder and a binder are mixed to cause the particles to grow in size. Particles are agglomerated to improve flowability, dispersibility, bulk density and dusting behaviours while decreasing caking. The pharmaceuticals industry makes extensive use of granulation for these reasons, as well as to mix active pharmaceutical ingredients with excipients to produce particles that are easier to dispense. Granulation is also used in the manufacture of fertilisers, detergents and is used in food processing concerns.

Granulation is performed by mixing a powder with either a liquid binder or by melt granulation. The powder is placed inside a mixer, which may be any one of a rotating drum, a fluidised bed or a high shear mixer. The binder is added to the mixing chamber and the agglomeration process commences; where the properties of the end-product are dependent upon the equipment process conditions and the materials employed. The system of interest in this paper is a high-shear wet granulation process using lactose monohydrate and deionised water in a horizontal axis plough-share mixer. Lactose, due to its wide use as an excipient, has been used in granulation experiments to investigate many aspects of particle granulation; to study the influence of droplet size on particle granulation [3], the influence of granulation method on compactability [30], scaling-up of granulation in high shear mixers [2], the sensitivity of the process conditions on the end-product [5] and the importance of spray flux on the nucleation stage of granulation [20], among others.

While on a cursory level this is a simple process, on a kinetic level it is very complex and poorly understood. Hence, the use of models and computer simulations has been incorporated into the study of granulation to further understanding of the systems while avoiding extensive experimentation [6, 8–11, 22]. In [17] a detailed population balance model [7, 12] was found to exhibit sensitivity to the initial powder size distribution to such an extent that the uncertainty attached to the values characterising the initial powder were directly related to the ability of the model to reproduce the experimental results. The work in this paper is designed to experimentally investigate the behaviour indicated by the model.

The effect of the initial powder size distribution has been studied under a variety of conditions and initial material characterisations. The majority of the existent work is based on effects caused by changing the mean particle size, or similar values, e. g., the median particle size [15, 21, 25, 28]. Other studies focus on the effect of the spread of the initial particle distribution [14, 26].

It was found that the changes in initial particle size distribution affected the resulting product size distribution [1, 4, 21], porosity [4, 5, 15, 28], strength [19], compactability [4] and intergranular forces [21]. Moreover, changes to the initial particle size have been found to not only affect the growth mechanism directly [4, 5, 19, 21, 25], but also have been found to interact with other process conditions such as the binder viscosity [15, 28], the droplet size [1] and liquid requirements [18].

Additionally, work has been reported in which two powder substances have been blended, which inherently will be combinations of different size distributions. Altering the mixture ratio of the two materials was found to affect the end-product particle distribution [13],

the particle strength [23] and granule shape [27]. However, as more than one substance was involved, it is not possible to make a clear statement about the impact of the initial size distribution.

With a few notable exceptions [21, 25], the descriptions of the initial powders used are typically reported in terms of one or two single-value characterisations, such as a mean particle size and the span of the distribution. The shape of the initial powder distributions has not been the focus of any of these studies and the commonly reported characteristic values do not contain enough information to make inferences about the shape of the distribution.

The question of how to adequately describe the initial powder is of particular concern to pharmaceutical process development, where ICH Q8 [29] requires the identification of a multidimensional *design space* which is the combination of material attributes and process parameters that have been demonstrated to provide assurance of quality. If the initial powder is inadequately described, it may not be possible to correctly identify the design space, nor to determine the reasons for any variability in the process output.

The **purpose of this paper** is to experimentally investigate the impact of the shape of the initial powder distribution on a granulation process. For this purpose, three initial powder distributions have been constructed with a narrow unimodal distribution, a wide unimodal distribution and a wide bimodal distribution, all with a single fixed volume mean diameter. The resulting product distributions are characterised to show that increasing the variance alters the end-product. Further, introducing a bimodality to increase the variance has a different effect on the product size distribution than increasing the variance while maintaining a unimodal distribution.

The structure of this paper is as follows: Section 2 contains a description of the experimental system and details on how the initial powder distributions are constructed. Section 3 contains details of how the end-product is described in quantitative terms. In section 4 we show the results of the granulation experiments and compare the results for the three end-product distributions. In section 5 we draw conclusions and discuss recommendations for future work.

2 Experiment

2.1 Powder distribution construction

Manufacture step. The powder used to construct the initial powder distributions was first separated into size fractions by sieving. A single tier of six sieves and a bottom pan were used with sieve apertures of 150, 106, 90, 75, 53, and 45 μm. Quantities of approximately 150 g of lactose monohydrate (Granulac 230, Granulac 200 and Granulac 140, Meggle, Germany) were subjected to 30 minutes at 2.0 mm amplitude on a sieve shaker (model EVL1, Endecotts, UK). Material retained in each sieve class was separately collected until sufficient quantities of each type had been collected to construct new distributions. The final accumulated powder samples of interest are referred to as M_{45}, M_{53}, M_{75}, M_{90}, and M_{106} which represents powder collected in the 45μm, 53μm, 75μm, 90μm and 106μm

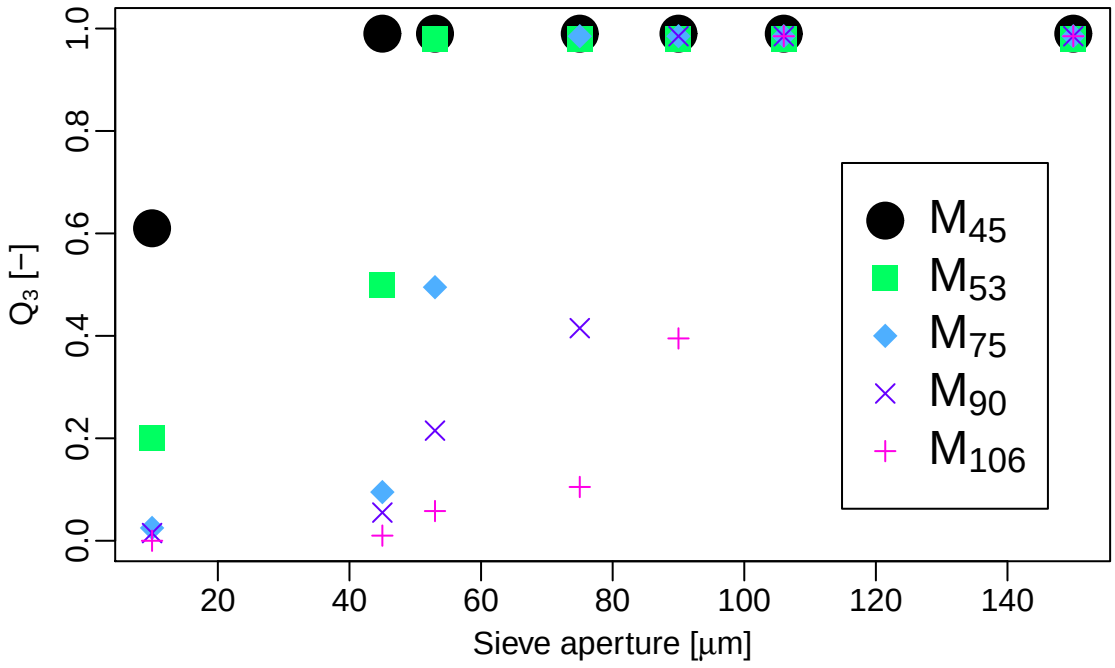

Figure 1: *$M_{[.]}$ initial size distributions.*

aperture sieves, respectively.

Powder analysis step. An approximately 75 g sample of powder was taken from each $M_{[.]}$ sample and was sieved under the same conditions as the manufacturing step. As can be seen in **Figure 1**, the cumulative mass measurements, or Q_3, found for each powder type are easily distinguishable. While each sample includes a notable proportion of smaller sized particles, a significant amount of the mass is in the appropriate size class and each distribution has a distinct median value.

P_n; Narrow unimodal distribution creation step. With the measurements found in the sieving process, we used the midpoints of each sieve class to calculate the arithmetic mean particle size and variance for the M_{90} powder to be 84.89 μm and 386.93 μm^2. The P_n initial powder consists of only powder taken from M_{90}.

P_b; Bimodal distribution creation step. The measured distribution values for the M_{53} and M_{106} powders were used to construct a second distribution, P_b which would have the same arithmetic mean particle size as P_n, but a larger variance induced by a bimodality. Specifically:

$$\sum_{i=1}^{N} q_{3i}(\mathrm{M}_{90})x_i = \sum_{i=1}^{N} c_1 q_{3i}(\mathrm{M}_{53})x_i + c_2 q_{3i}(\mathrm{M}_{106})x_i \,, \qquad (1)$$
$$c_1 + c_2 = 1 \,,$$

where $q_{3i}(\mathrm{M}_{[.]})$ is the i^{th} element of the mass frequency distribution, q_3, for the powder sample $\mathrm{M}_{[.]}$, x_i is the midpoint of the aperture of the i^{th} sieve class and N is the number of sieving measurements.

Using this method of creating a distribution, we calculated a blended distribution by solving for values of c_1 and c_2 such that:

$$q_{3i}(\mathrm{P_b}) = 0.431 q_{3i}(\mathrm{M}_{53}) + (1 - 0.431) q_{3i}(\mathrm{M}_{106}) \,, \qquad (2)$$

which has an arithmetic mean particle size and variance of 84.89 μm and 1387.56 μm^2, respectively. We then blended 431 g of powder from the M_{53} bag and 569 g from the M_{106} bag to create 1000 g of $\mathrm{P_b}$ powder.

$\mathrm{P_u}$; Uniform wide variance distribution creation step.

Following a similar procedure we constructed a distribution $\mathrm{P_u}$ where

$$\sum_{i=1}^{N} q_{3i}(\mathrm{M}_{90})x_i \approx \sum_{i=1}^{N} c_1 q_{3i}(\mathrm{M}_{45})x_i + c_2 q_{3i}(\mathrm{M}_{53})x_i + c_3 q_{3i}(\mathrm{M}_{75})x_i +$$
$$c_4 q_{3i}(\mathrm{M}_{90})x_i + c_5 q_{3i}(\mathrm{M}_{106})x_i \,, \qquad (3)$$
$$\sum_{j}^{5} c_j = 1 \,,$$

with the further restriction that the resulting distribution be unimodal with a maximal variance. This was accomplished by imposing ranges upon the c_i's which would force a unimodal distribution and then optimising over these ranges to minimise the difference from the mean particle size while maximising the variance. Thus we constructed a distribution such that:

$$q_{3i}(\mathrm{P_u}) = 0.11 q_{3i}(\mathrm{M}_{45}) + 0.06 q_{3i}(\mathrm{M}_{53}) + 0.24 q_{3i}(\mathrm{M}_{75}) 0.29 q_{3i}(\mathrm{M}_{90}) + 0.3 q_{3i}(\mathrm{M}_{106}) \,, \quad (4)$$

which gives a unimodal distribution with an arithmetic mean particle size of 81.39 μm with a variance of 1014.10 μm^2. The powders from the individual $\mathrm{M}_{[.]}$ sieve cuts were then blended with 110 g, 60 g, 239 g, 290 g and 300 g from samples M_{45}, M_{53}, M_{75}, M_{90}, and M_{106}, respectively.

Thus we have three 1000 g bags of powder for granulation with approximately the same arithmetic mean particle size where P_n is unimodal with a small variance, P_u unimodal

Table 1: *Constructed distribution defining (arithmetic) values.*

Distribution	Mean	Variance	Standard Deviation
P_n	84.89 μm	386.93 μm^2	19.67 μm
P_b	84.89 μm	1387.56 μm^2	32.24 μm
P_u	81.39 μm	1014.10 μm^2	31.84 μm

with a large variance and P_b is bimodal with a large variance. The final calculated mass fractions and cumulative mass fractions of the initial powder distributions are shown in **Figure 2** and **Figure 3**, respectively. The arithmetic values used to define the distributions are summarised in **Table 1**.

2.2 Granulation of constructed powder distributions

A wet granulation process using lactose monohydrate in a bench-scale mixer shall be the system of interest in this study. The equipment and procedures used in this system are identical to those used in [17], from which the following description is taken.

Experimental setup The equipment setup that was used is shown in **Figure 4**. The experiments were performed using a horizontal axis 5 litre ploughshare mixer (Kemeutec) which is described in detail in Jones and Bridgwater [16]. The mixer shaft is driven by a variable speed DC motor with a torque meter (DRBK-20-n, ETH Messtechnik, Germany) mounted between the shaft and the motor.

The binder is drawn from a reservoir by a magnetic drive gear pump (model DG.19, Tuthill Corporation, USA) to a single fluid nozzle (model 121, orifice $\oslash$ 0.5 mm, 60 ° spray angle, Düsen-Schlick, Germany) which is suspended at a fixed height above the powder bed. The pump speed is controlled by an inverter (model Altivar 31, Telemecanique, France). The inverter frequency is determined by pulses from a flowmeter (OG1, Nixon Flowmeters, UK) that is mounted between the pump and the nozzle. The equipment is controlled and monitored by a Labview application that allows for the specification of flow rate, mixer speed, and all timing elements of the granulation process. Controlling, as well as monitoring and recording the operating status of the equipment, is facilitated by two data recording cards (6009 and 6601, National Instruments, USA).

Materials Each granulation run was performed using 1000 g of lactose monohydrate material from P_n, P_b or P_u using 150 ml deionised water as the binder.

Procedure Upon loading the mixer with the powder, the nozzle is suspended at a fixed height above the mixing chamber with a drip cup immediately below it collecting the

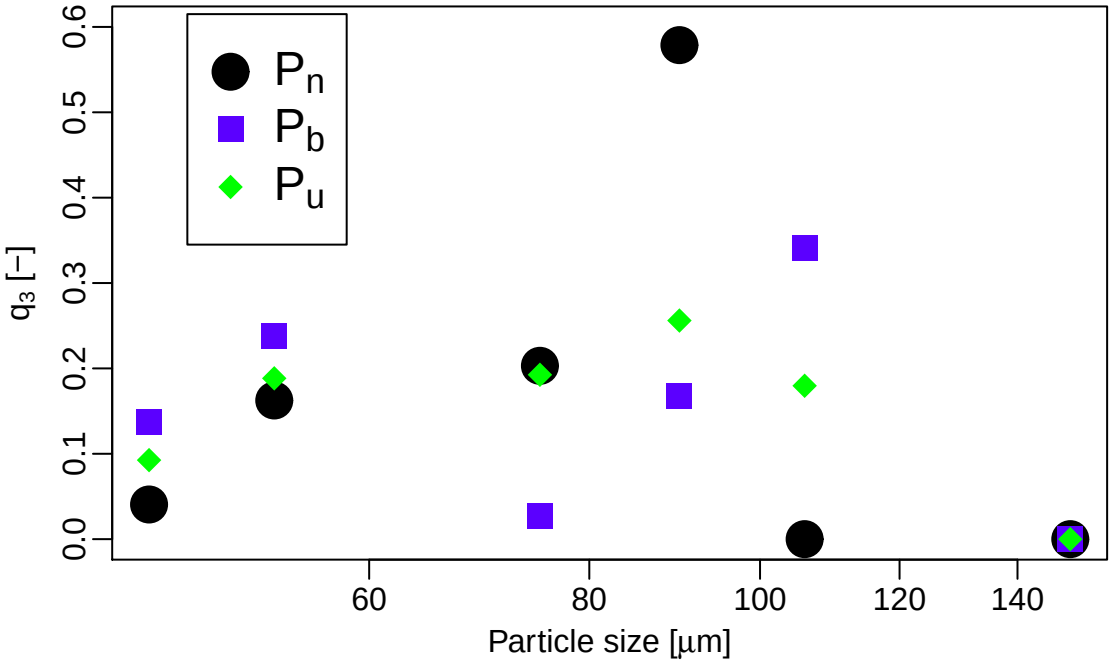

Figure 2: *Initial constructed powder distribution mass fractions.*

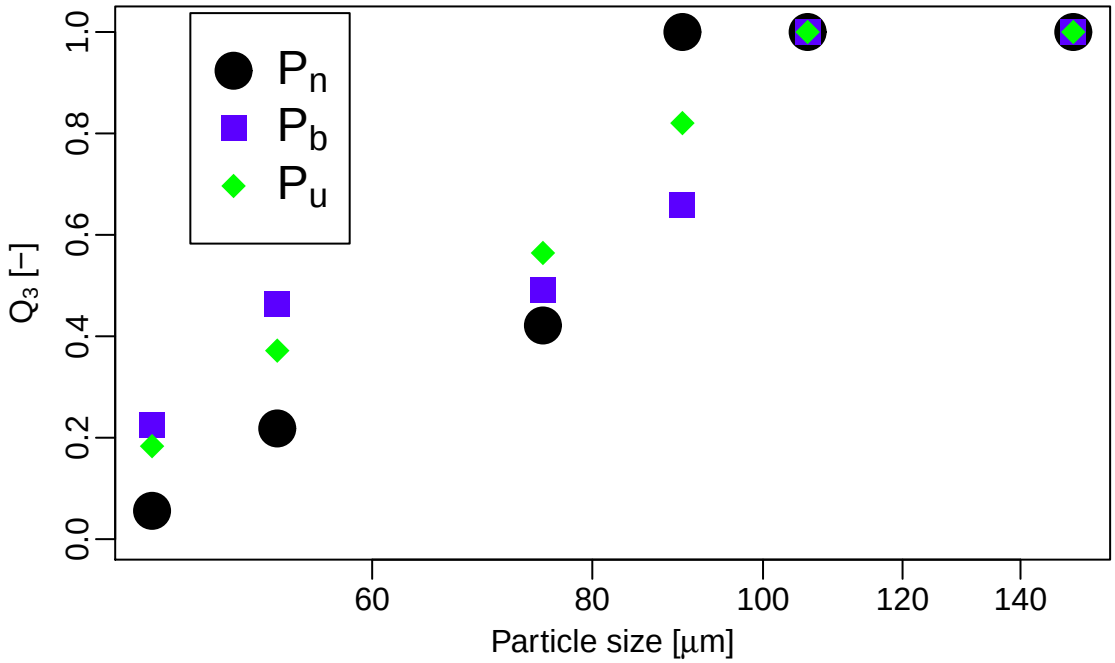

Figure 3: *Initial powder cumulative mass fractions distributions.*

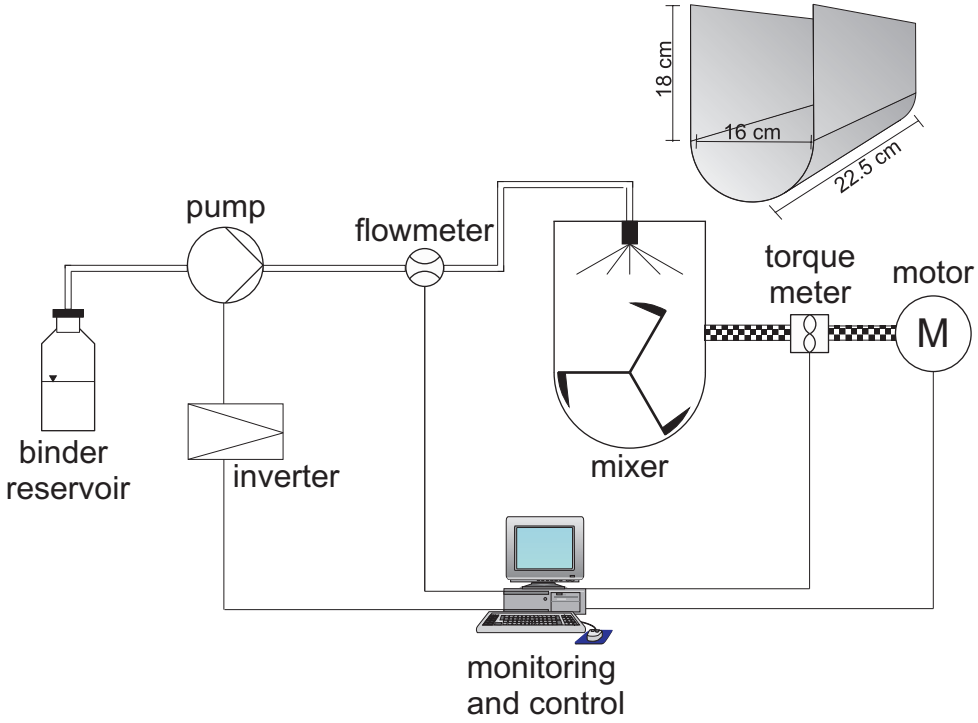

Figure 4: *Experimental setup.*

binder flow. After two minutes of dry mixing to aerate the powder, the drip cup is removed. When the allotted amount of time for the binder addition has passed, the binder stream is automatically shut off by the controller program and the drip cup is placed under the nozzle again to prevent any additional droplets from reaching the powder bed. The mixer continues to run at the specified speed until the desired amount of time has elapsed.

The resulting product is removed from the mixer and distributed onto metal trays. The trays are placed in an atmospheric pressure drying cabinet (INC 95SF, Genlab, UK) at $50\,^{\circ}$C with $55\,^{\circ}$C designated as the overheat temperature, until no significant change in mass is recorded. The dried product is recombined and samples for analysis are chosen by putting the entire dried product through a sample splitter until a sample between 60–100 g is obtained.

Particle size analysis is performed on each end-product sample by sieving. Three tiers, each consisting of six sieves and a bottom pan, were constructed using a $\sqrt{2}$ progression from 53–16000 μm. Each tier was subjected to 25 minutes at approximately 1.5 mm amplitude on a sieve shaker (model EVL1, Endecotts, UK). Material in each of the sieves, as well as the bottom pan of the third tier, was weighed and recorded giving 19 mass measurements per sample. The data set of measured mass values is included as a supplementary .csv file to this paper.

Measurement precision. A preliminary granulation run was performed in order to detect any sources of error or sub-optimal procedures. Prior to each granulation run, the binder flow rate was tested for 150 ml of binder over the specified addition time. Each

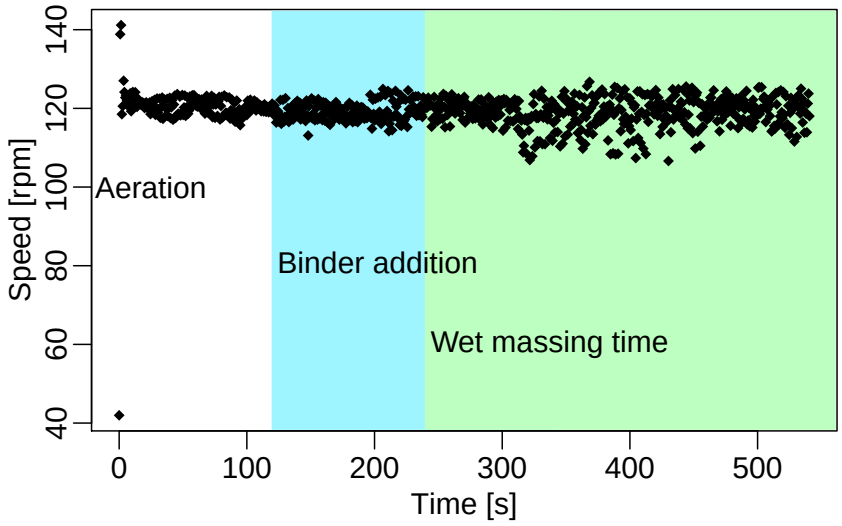

Figure 5: *Measured speed of mixer during P_b material granulation run. Mean speed is 119.28 rpm with a variance of 18.55 rpm^2.*

granulation run was performed using control conditions that produced 150 ml $\pm$ 2 ml of binder under the relevant process conditions immediately prior to each run. The sieving analysis of the preliminary run allowed the selection of an appropriate range of sieves. Through this preliminary work it was determined that three samples taken from a single granulation run would yield acceptable results for the product size distribution. To minimize sampling error in this case, five samples were taken for each of the P_n, P_b and P_u granulation runs. The balance employed for the material and sieving measurements (XB 3200C, Precisa) is precise to 0.01 g. The motor controller, in conjunction with the torque meter, monitors and adjusts the impeller speed at least once per second. Analysis of the recorded measurements during the granulation runs indicate that the average deviation from the specified speed for all the granulation runs is less than $\pm$ 2 rpm. **Figure 5**, **Figure 6** and **Figure 7** show the measured speed values for individual experimental conditions. As can be seen, while there are fluctuations, the speed does stay centred on the designated value, in this case 120 rpm.

Previous experiments [17] were used to select process conditions which would have the most apparent impact on the end-product. Aside from the initial powder distribution, all other process conditions, including the binder to powder ratio of 150 ml:1000 g, are held constant:

1. Binder addition flow rate of 75 ml/min (flow rate);
2. Impeller speed of the mixer of 120 rpm (impeller speed) and
3. Allowing the mixer to continue after all the binder had been added for 5 minutes (wet massing time).

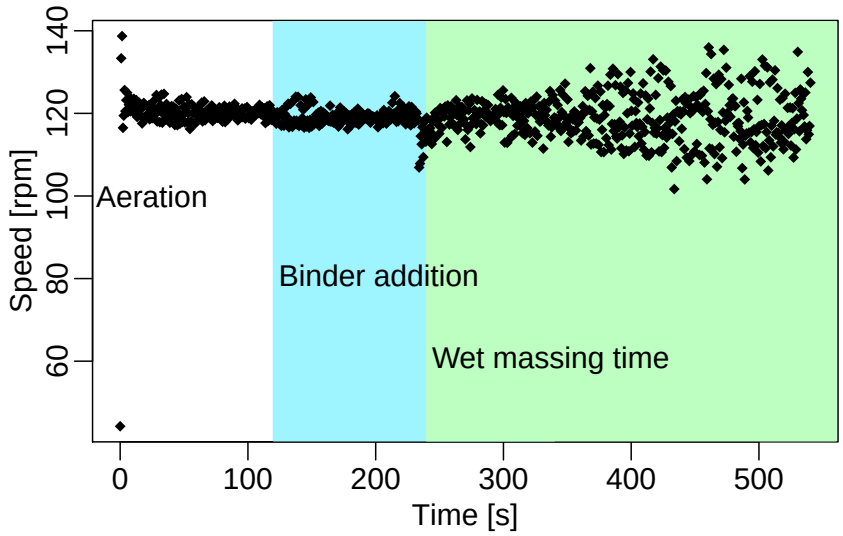

Figure 6: *Measured speed of mixer during P_n material granulation run. Mean speed is 119.17 rpm with a variance of 26.28 rpm^2.*

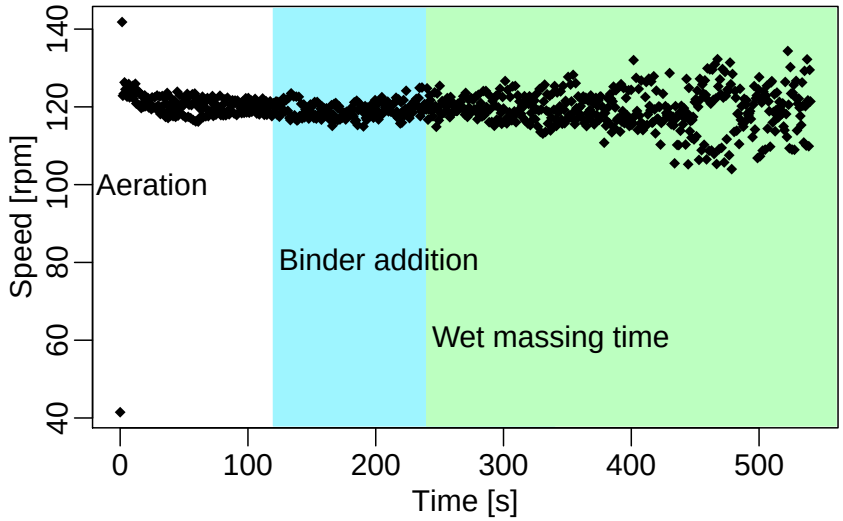

Figure 7: *Measured speed of mixer during P_u material granulation run. Mean speed is 119.52 rpm with a variance of 24.10 rpm^2.*

3 Methodology

In this section, the methods and calculations used to quantify the measurements obtained from the experimental system are described in detail. It should be noted that unless explicitly stated otherwise, the sieve diameter values, x, used in this section are the midpoints of the experimental sieve cuts, e. g., particles collected in the 75μm aperture sieve are treated as having a diameter of $(75 + 106)/2 = 90.50\mu$m. Since any given sieve measurement is associated with a diameter range, using the midpoint for calculation purposes represents a compromise between the known upper and lower boundaries and is intended to minimise over- or under-estimation.

3.1 Physical measurements

The complete end-product for each of the three experimental cases was run through the sample splitter until five samples of 60–100 g were obtained for each case. Each of these samples was subjected to sieving analysis giving a mass associated with one of $N = 19$ sieve cuts. Each sieve cut mass is then expressed as a mass fraction by dividing the sieve cut mass by the total sample mass. This is expressed as a vector w_k, where the i^{th} element is defined as:

$$w_{ki} = \frac{m_{ki}}{\sum_{j=1}^{N} m_{kj}}\,, \tag{5}$$

where m_{ki} is the mass measured of the k^{th} sample in the i^{th} sieve class. In this case $k = 1, \ldots, 5$ and $i = 1, \ldots, 19$. Each m_{ki} directly corresponds to the measured mass in a sieve cut for a given sample.

As each granulation run has five samples associated with it, we calculate a vector of averaged mass fractions, or probability density function (pdf), d, for a given granulation run across the k samples for each sieve class where each element d_i is defined as:

$$d_i = \frac{1}{k}\sum_{j=1}^{k} w_{ji}\,. \tag{6}$$

From this point, we convert the values into an empirical cumulative distribution function (ECDF) by summing all fractions of mass that are smaller than a given sieve class. This is equivalent to sieving measurements in terms of the proportion of material that has passed through any given sieve aperture. The ECDF, D, is calculated directly from the averaged mass fractions as:

$$D_i = \sum_{j=1}^{i} d_j\,, \tag{7}$$

where $i = 1, \ldots, 19$.

The inverse process, to derive mass fractions from a given ECDF is performed by measuring the jumps between the sieve cuts values or:

$$d_i = \begin{cases} D_i & i = 1\,, \\ D_i - D_{i-1} & i > 1\,. \end{cases} \tag{8}$$

Further, we construct error bars around the averaged mass fractions and ECDF using 95% confidence intervals with the Student's t distribution based on the k samples for each measurement.

Using the mass fractions, d and the ECDF, D, we are able to calculate a wide variety of characterisations to describe the results.

3.2 Mean particle diameters

Describing a particle size distribution in terms of a mean diameter is a common practice. There are a variety of methods to calculate this value, but overall it is intended to portray an average particle size. Here, we briefly describe the methods used in this paper.

3.2.1 Empirical arithmetic mean particle size

The empirical arithmetic mean particle diameter, M_{a}, is the simplest single value expression used to describe particle size distributions. The empirical arithmetic mean particle diameter is calculated by first multiplying the mass fraction, d, by the associated sieve diameter, x, and then summing the results:

$$M_{\mathrm{a}} = \sum_{i=1}^{N} d_i x_i\,, \tag{9}$$

where d_i is the i^{th} element of the density vector d, and x_i is the midpoint of the ith sieve class.

3.2.2 Empirical geometric mean particle size

The empirical geometric mean particle diameter, M_{g}, is an alternative to the arithmetic mean. In this context, we transform the sieve aperture values, x, by taking the natural logarithm, $x' = \ln(x)$. This transformation is used so that the larger particle size classes are not disproportionately represented. In the sieve series that we use, the range of the sieve classes is 53–16000 μm which, by taking the natural log, transforms into the range 3.97–9.68. The empirical geometric mean particle size is calculated by first multiplying the density by the transformed sieve class and then taking the exponential of the summed

the results:

$$M_{\mathrm{g}} = \exp\left(\sum_{i=1}^{N} d_i x_i'\right), \tag{10}$$

where each d_i is the i^{th} element of the density vector, d, and x_i' is the natural logarithm of the midpoint of the i^{th} sieve class.

3.3 Dispersion and shape of distribution

3.3.1 Variance and standard deviation

The variance, $\sigma^2_{[.]}$ is a single value expression of the spread of the resultant particle size distribution. For this value, we shall use the untransformed midpoints of the sieve class values. The arithmetic variance, σ^2_{a}, is calculated as:

$$\sigma^2_{\mathrm{a}} = \sum_{i=1}^{N} d_i x_i^2 - \left(\sum_{i=1}^{N} d_i x_i\right)^2, \tag{11}$$

where each d_i is the i^{th} element of the density vector d and x_i is the midpoint of the i^{th} sieve class.

The standard deviation, σ_{a}, is the square root of the variance, or:

$$\sigma_{\mathrm{a}} = \sqrt{\sigma^2_{\mathrm{a}}}\,. \tag{12}$$

Alternately, we can calculate an empirical geometric standard deviation, or σ_{g} by using the $x' = \ln(x)$ transformation, as before. In this paper, we calculate the empirical geometric standard deviation as:

$$\sigma_{\mathrm{g}} = \exp\left(\sqrt{\sum_{i=1}^{N} d_i (x_i' - \ln M_{\mathrm{g}})^2}\right). \tag{13}$$

3.3.2 Skewness

The skewness, or γ_1, of a distribution quantifies the deviation from symmetric, or how lopsided the distribution is, where larger values indicate more lop-sidedness. Using only the arithmetic-based values skewness is calculated as:

$$\gamma_1 = \frac{\left(\sum_{i=1}^{N} d_i x_i^3\right) - 3M_{\mathrm{a}}\sigma^2_{\mathrm{a}} - M^3_{\mathrm{a}}}{\sigma^3_{\mathrm{a}}}\,. \tag{14}$$

3.3.3 Kurtosis

The kurtosis, or β_2, of a distribution quantifies the peakedness of the distribution. An arithmetic based calculation for kurtosis is defined as:

$$\beta_2 = \frac{\sum_{i=1}^{N} d_i x_i^4 - 4M_\mathrm{a} \sum_{i=1}^{N} d_i x_i^3 + 6M_\mathrm{a}^2 \sum_{i=1}^{N} d_i x_i^2 - 4M_\mathrm{a}^3 \sum_{i=1}^{N} d_i x_i + M_\mathrm{a}^4}{\sigma_\mathrm{a}^4} \,. \quad (15)$$

3.4 Percentile and span by interpolation

Another common metric is the percentiles of the distribution, specifically the 10th, 50th or median and 90th percentiles, or $\tilde{d}_{10}$, $\tilde{d}_{50}$ and $\tilde{d}_{90}$, respectively. In this context, the pth percentile is a diameter value where p percent of the mass belongs to particles with a diameter less than the percentile diameter. Here, we estimate the percentile diameter by linear interpolation between the upper and lower bounds of the relevant sieve class.

Further, the spread of a particle size distribution is often expressed in terms of the span of the distribution, which is defined as:

$$\mathrm{Span} = \frac{\tilde{d}_{90} - \tilde{d}_{10}}{\tilde{d}_{50}} \,. \quad (16)$$

3.5 Lognormal curve fitting

As the mass fractions and ECDF are of the form of probability distributions, it is common practice to fit a probability curve to the results and describe the results in terms of the characteristics of the fitted curve. For particle size distributions, a lognormal distribution is frequently used. This curve fitting can be accomplished using either the ECDF or the mass fractions. The density of a lognormal distribution has two parameters, μ_fit and σ_fit, and is defined as:

$$\ln \mathcal{N}(x, \mu_\mathrm{fit}, \sigma_\mathrm{fit}) = \frac{1}{x\sqrt{2\pi\sigma_\mathrm{fit}^2}} \exp - \frac{(\ln x - \mu_\mathrm{fit})^2}{2\sigma_\mathrm{fit}^2} \,, \quad (17)$$

The curves are fit by minimising a defined objective function, here the Euclidian distance between two N-element vectors F and G:

$$\mathrm{Dist} = \sqrt{\sum_{i=1}^{N} (F_i - G_i)^2} \,. \quad (18)$$

A lognormal curve can be fit to the mass fractions by substituting the density vector d for F and $\ln \mathcal{N}(x, \mu_\mathrm{fit}, \sigma_\mathrm{fit})$ for G in Equation 18. Similarly, a lognormal curve can be fit to the ECDF by substituting the empirical ECDF D for F and a lognormal cumulative

distribution with parameters μ_{fit} and σ_{fit} for G. In either case, the objective function Dist can then be minimized with respect to values for μ_{fit} and σ_{fit} using an optimisation technique, here the `optim` function in R [24]. While the ECDF is most commonly used for curve fitting in this context, in this paper curves are fit to both forms of the distribution for the purpose of comparison. When making this comparison, it is important to note the density d is defined in terms of the midpoints of the sieve cuts, whilst the empirical ECDF D is defined in terms of the upper bounds of the sieve classes.

Further, the fitted curves have metrics, such as mean and standard deviation which can be used to describe the distribution, in lieu of the empirically computed values. The lognormal distribution, with parameters μ_{fit} and σ_{fit} in particular, has defined values of:

$$\tilde{d}_{50} = \exp\left(\mu_{\text{fit}}\right) \tag{19}$$

$$M_{\text{a}} = \exp\left(\mu_{\text{fit}} + \frac{\sigma_{\text{fit}}^2}{2}\right) \tag{20}$$

$$M_{\text{g}} = \exp\left(\mu_{\text{fit}}\right) \tag{21}$$

$$\sigma_{\text{a}} = \sqrt{\left(\exp(\sigma_{\text{fit}}^2) - 1\right) \exp\left(2\mu_{\text{fit}} + \sigma_{\text{fit}}^2\right)}\,. \tag{22}$$

4 Results

In this section we present the experimental results as measured and as quantified by the methods described in Section 3.

4.1 Particle size distribution measurements

The measured mass fractions are averaged for each sieve class over the five samples taken from each granulation run using Eqn. 6. The resulting averaged size distributions are shown in **Figure 8**. From this figure it is apparent that the resultant distributions are different. As all process conditions were constant, except for the construction of the initial particle size distribution, this indicates that the initial powder distribution has a significant effect on the size distribution of the end-product. Some cursory observations which can be made from this data set are that the P_b bimodal distribution is right-shifted in comparison to the others. The P_b and the unimodal P_n distributions have the largest peaks, while the unimodal P_u distribution has a lower primary peak and more widely spread end-product distribution when compared to P_n. All three distributions have a small amount of fines and additional modes in the larger particle sizes. Further, it is interesting to note that, while the largest sieve aperture used for the initial powder was material collected in the 106 μm sieve, the product 106 μm sieve class is the first product sieve class to have a substantial amount of material, with the first significant jump in mass in the 150 μm sieve.

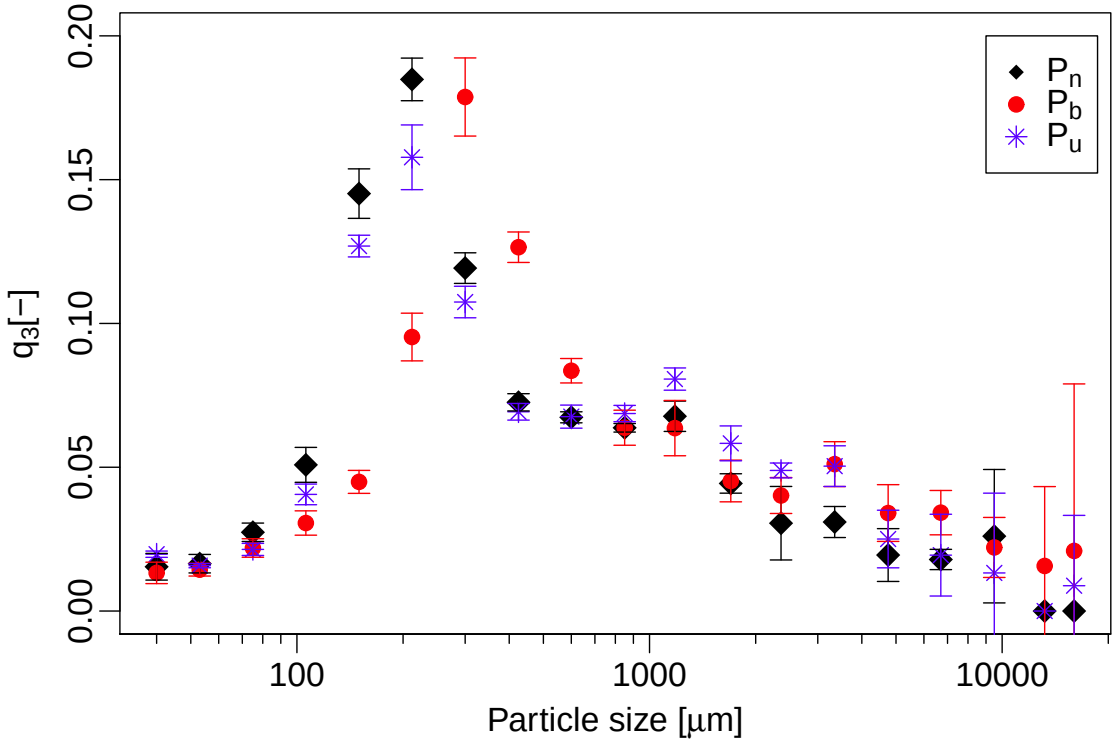

Figure 8: *Experimental results of sieving particle size analysis. Density distribution with error bars representing 95% confidence intervals.*

This suggests that most, if not all, of the material in the powder bed has been involved in the granulation process, particularly when one considers the attrition inherent in sieving analysis.

Based on these observations, it is of interest to quantify what effects altering the initial powder distribution has on the end-product.

4.2 Curve fitting to distribution results

It is common practice to fit particle size distributions to lognormal curves for analysis. Two lognormal curves are fit to the experimental data, one by fitting to the empirical cumulative distribution, the other by fitting to the mass fractions. As a side note, a number of alternate distributions were considered, and all such alternatives were found to deviate from the data points more than the lognormal distributions.

4.2.1 Lognormal fit to ECDF

The parameters of the curves fitted to the measured ECDF using Eqn. 18 can be seen in **Table 2**. **Figure 9** shows the empirical cumulative size distribution of the three end-product distributions and the lognormal curves fit to the data points. From these figures, we can observe that, to varying degrees, the data points are initially overestimated, then underestimated around the primary mode by the curves. The good agreement occurring at the upper end of the distributions can be largely disregarded as a positive feature because, by definition, an ECDF will always asymptote to 1. This attribute, as well as the cumulative form, may conceal flaws in the fitting process. The systematic deviation of the curves fitted to the ECDF implies that they may not be suitable for this data.

Another feature to notice is the variance attached to the experimental data points when expressed as an ECDF. The P_b bimodal case appears to have noticeably larger sample variance from the five measurements taken to obtain the averaged ECDF, than either of the unimodal cases.

Table 2: *ECDF lognormal curve fit parameters.*

	P_n	P_b	P_u
Lognormal curve μ_{fit} parameter	6.18	6.73	6.39
Lognormal curve σ_{fit} parameter	1.25	1.33	1.38

4.2.2 Lognormal fit to mass fractions

The parameters of the lognormal curves when fit to the measured mass fractions using Eqn. 18 can be seen in **Table 3**. The change in values from **Table 2** to **Table 3** is indicative of an aspect of using the ECDFs as a form of analysis. Whilst the changes in μ_{fit} can be ascribed to the use of the mid-points, as opposed to the upper bounds of the sieve classes in the definition of the mass fraction density d, the P_b and P_u distributions show a significant increase in σ_{fit}, whereas the P_n distribution shows a decrease in σ_{fit}. This indicates that the mass fractions are being fitted to a wider distribution than the ECDF, while P_n is being fitted to a narrower distribution.

Table 3: *Mass fractions lognormal curve fit parameters.*

	P_n	P_b	P_u
Lognormal curve μ_{fit} parameter	5.85	6.60	6.20
Lognormal curve σ_{fit} parameter	1.19	1.55	1.42

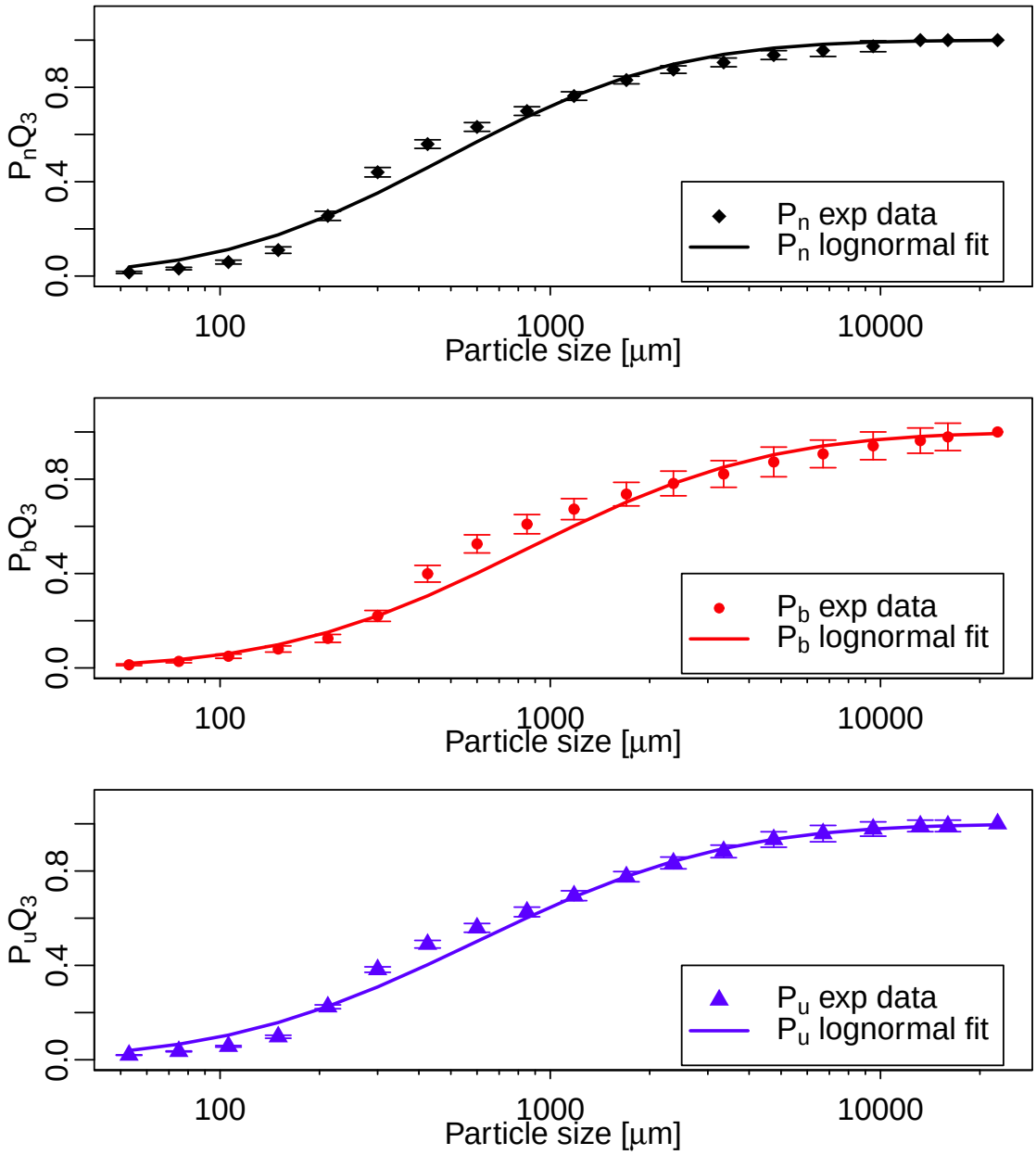

Figure 9: *Lognormal curves fit to Empirical CDFs. Error bars representing 95% confidence intervals.*

We can observe the outcome of these changes in **Figure 10**. Of particular interest is the increase in the tail-weight of the distributions such that P_b has more probability allocated to the upper tail than the other distributions. Regarding the initial upward spike in the fitted curves in the lower tail, this is an artefact of the discrete sieve class structures imposed upon the continuous lognormal function. The fitted distributions are universally right-shifted with respect to the highest point of the experimental distributions and the probability is not allocated to the primary modes in correct magnitudes. Overall, the quality of the density fitted curves is debatable.

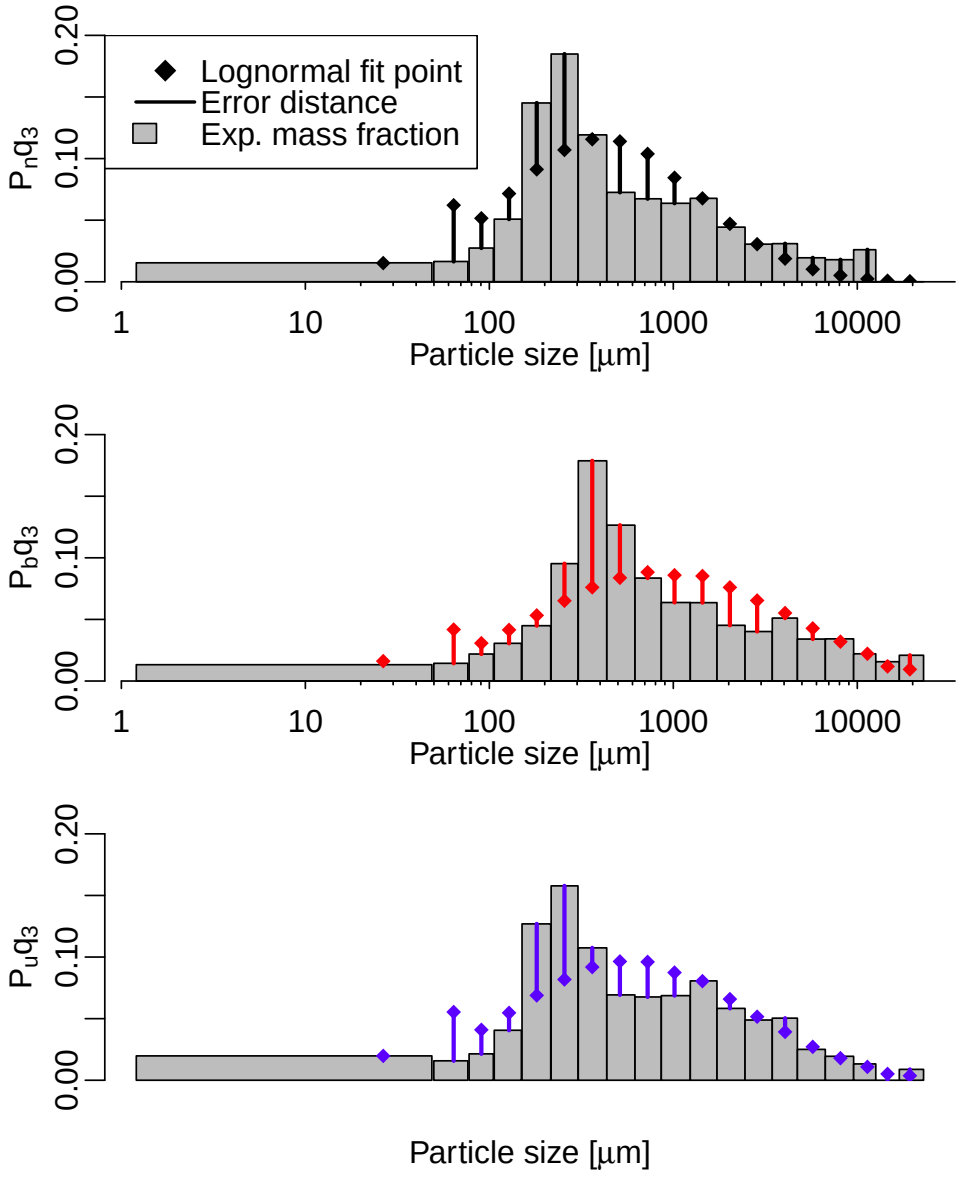

Figure 10: *Lognormal curves fit to mass fractions.*

4.3 Single value characterisations

A number of widely used single value characterisations are presented in this section using both the original experimental data and the fitted curves.

4.3.1 Mean particle diameter

As discussed in the methodology section, the manner of calculating a mean particle diameter can vary. In **Figure 11** we can observe how the various methods of calculating a mean diameter are realised. As expected from the methods of calculation, the arithmetic based means, even when based on the lognormal distribution fit parameters, are significantly larger than the values which have been transformed to reduce the influence of the larger sieve aperture classes. However, the qualitative behaviour is the same for all calculations, insofar that the smallest to largest mean diameters for the granulation runs are in the order of P_n, then P_u with the largest mean diameter belonging to P_b, independent of method of calculation. The different calculation methods all give noticeably different values and affect the magnitude of the differences; however the overall trend is not affected.

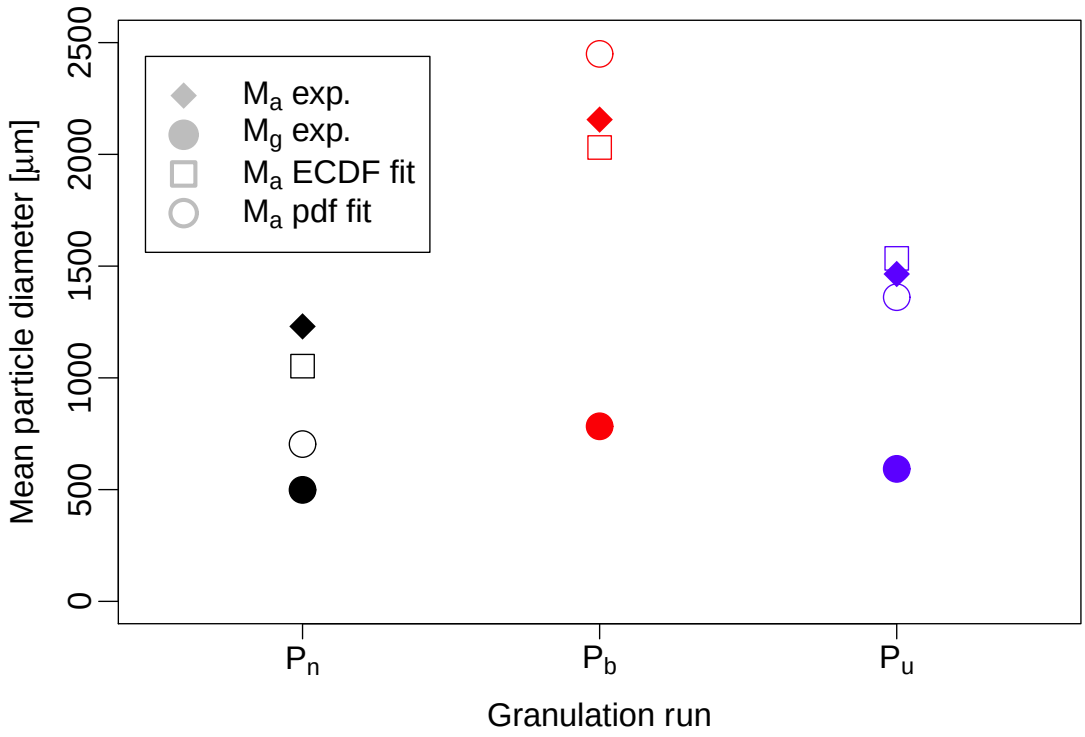

Figure 11: *Calculated forms of mean diameter particle size.*

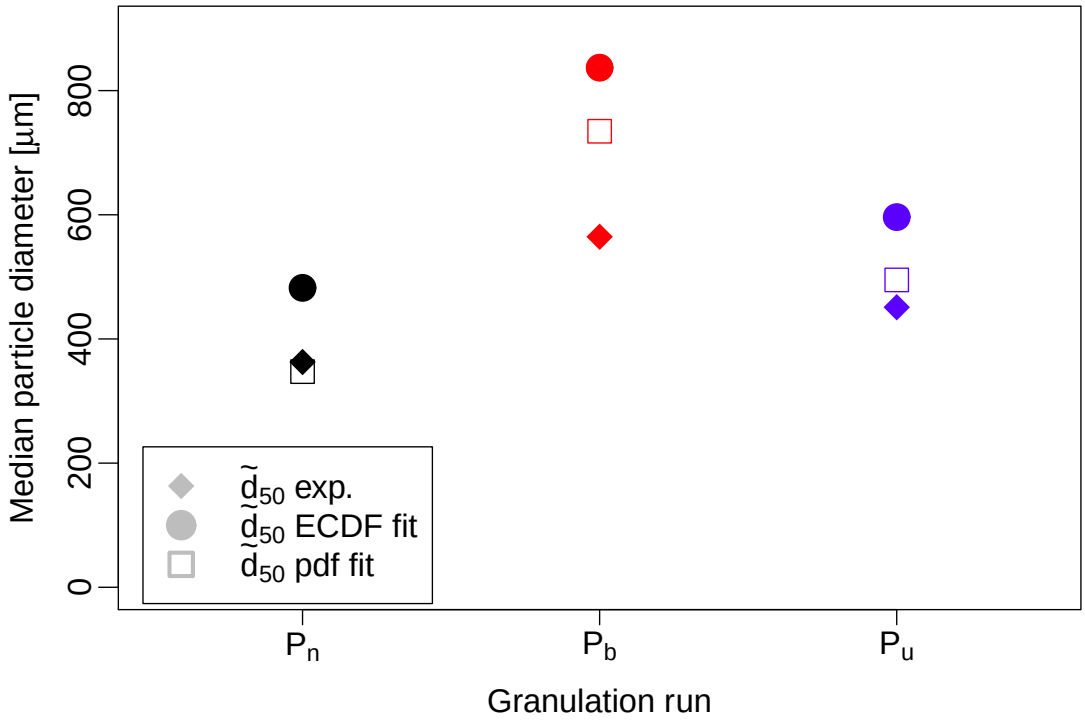

Figure 12: *Median diameter particle size.*

4.3.2 Median particle diameter

In **Figure 12** we can observe the median, or $\tilde{d}_{50}$ of the characteristic distributions follow the same pattern as the means of the distributions. The product from P_n is the smallest followed by P_u, with the P_b having the largest median particle diameter. We can further observe the effects of the fitting process, as the fit values are significantly larger than the experimentally found values for the EDCF fit, while the mass fraction based fit is only significantly off for the P_b case.

4.3.3 Particle size distribution spread and shape

The spread, or how narrow or wide the distribution is can be characterised in multiple ways. As the variance and standard deviation are expressions of the same feature of a distribution, we shall neglect reporting the variance in favour of the more widely used empirical standard deviation. In **Figure 13** the standard deviation as calculated using the arithmetic values, and the values derived from the lognormal fit curves are displayed. Here we observe the same qualitative agreement in the methods as seen with the mean particle sizes for the experimental and the mass fraction fit curve. From smallest to largest, the particle size distributions follow the order of P_n, P_u, with P_b having the largest standard deviation where the magnitude of the differences is strongly dependent upon the calcula-

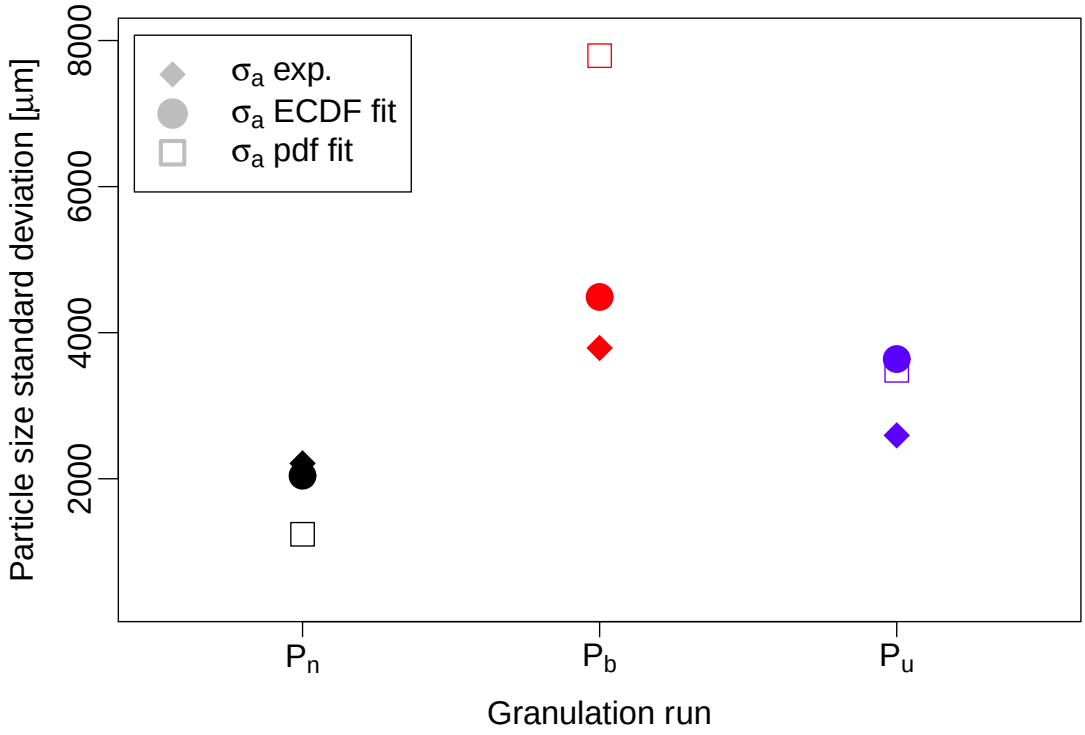

Figure 13: *Calculated forms of particle size distribution standard deviation.*

tion technique. The mass fraction fit curve seems to particularly exaggerate the standard deviation.

There are other metrics which can be used to further describe the found particle size distribution. In **Table 4** multiple measurements are presented based on the experimental measurements and the curve fits for consideration. The span, interestingly, changes the order of the granulation run with respect to size of the spread. Although P_b still is clearly the largest, P_n is attributed to having a larger span then P_u, contrary to all other metrics, although the values are fairly close numerically. The geometric standard deviations maintain the same pattern as the arithmetic standard deviations, except for the ECDF fit curve, where P_n is the smallest but P_u is larger than P_b. Bearing in mind the poor fit of the curves, and the changes when the mass fractions were used to fit the curves, this can be attributed to failure of the lognormal ECDF fitted curve to match the behaviour of the data points. The 10th and 90th percentiles also maintain the established pattern; however these metrics have the additional values of addressing oversized and undersized materials. The lower 10% of the material in the P_n run is within the range of the starting material, while the P_b is larger than any of the starting materials and the P_u just crosses the borderline of the upper size boundary of the initial materials. The P_n and P_u have a similar amount of oversized materials, however the P_b has a marked increase in the size for the 90th percentile. This indicates that that P_b has significantly more oversized material.

Table 4: *Single value empirically computed characterisations of the end-product distributions.*

	P_n	P_b	P_u
Mode	212	300	212
Span	8.34	10.83	8.13
Exp. geometric std	3.49	3.95	3.71
ECDF fit geometric std	3.48	3.78	3.95
Mass fraction fit geometric std	3.28	4.72	4.14
10th percentile ($\tilde{d}_{10}$)	141	177	151
90th percentile ($\tilde{d}_{90}$)	3167	6294	3819
Skewness	3.21	2.91	4.03
Kurtosis	13.54	11.64	23.63

The skewness and kurtosis are more descriptive of the shape of the product distributions. The experimentally based calculations for the skewness indicate that the P_b distribution is the least skewed, with the P_u being the most skewed. Similarly, the kurtosis indicates that the P_b distribution is the most peaked, while the P_u distribution is the least. This suggests that merely increasing the variance in the initial powder creates a wider and more skewed distribution, while increasing the variance by introducing bimodality causes a more central symmetric distribution at a larger central point.

5 Conclusions

We have presented an investigation of the effects of the shape of the initial powder distribution on a wet granulation system. Experimental data was produced using carefully constructed distributions of lactose monohydrate for the initial powder and deionised water as the binder. Initial powder distributions were constructed from sieve cuts to have the same arithmetic mean particle diameter while the shapes of the entire distributions are narrowly unimodal, widely unimodal and widely bimodal. The resulting end-product distributions are characterised and compared by a variety of metrics.

The product distributions were found to show markedly different characteristics. Increasing the variance while maintaining a unimodal initial powder distribution demonstrated a broader, more evenly distributed product with the peak of the resulting distribution being in the same size class as the peak of the product distribution resulting from the initial narrow unimodal distribution. The end-product from the bimodal distribution showed an even larger variance, an increase in oversized particles, a decrease in fines and a right-shift of the peak product distribution of at least one sieve class when compared to both unimodal cases. These results were found to be independent of the methods of characterising the product distributions.

These findings are of particular relevance to pharmaceutical process development. The failure to adequately describe the initial powder may prevent the correct identification of the *design space* for a granulation process, and may impede the ability to determine the reasons for any variability in the process output.

Future work will use this experimental data with a population balance model to investigate the physical mechanisms for this observed behaviour. Further, with the significance of the shape of the initial powder distribution being established, methods to experimentally establish and mathematically describe the initial material will be investigated, as the current standard methods do not encompass the distribution shape.

Acknowledgements

Support of the Singapore National Research Foundation under its Campus for Research Excellence And Technological Enterprise (CREATE) programme is gratefully acknowledged.

Nomenclature

Roman symbols

$\oslash$	Diameter	mm
c_i	i^{th} constant	[-]
d	Averaged density function	[-]
d_i	i^{th} element of density function	[-]
$\tilde{d}_{10}$	10th percentile	μm
$\tilde{d}_{50}$	50th percentile or median	μm
$\tilde{d}_{90}$	90th percentile	μm
D	Cumulative mass function	[-]
D_i	i^{th} element of cumulative mass function	[-]
Dist	Euclidian distance objective function	[-]
ECDF	Empirical cumulative density function	[-]
F	vector of values for objective function	[-]
G	vector of values for objective function	[-]
$\ln \mathcal{N}(x, \mu, \sigma)$	Lognormal distribution with parameters μ and σ	[-]
m_{ki}	measured mass of k^{th} sample in the i^{th} sieve class	[-]
$\mathrm{M}_{[.]}$	Powder accumulated in any single given aperture sieve.	[-]
M_{45}	Powder accumulated in 45 μm sieve	[-]
M_{53}	Powder accumulated in 53 μm sieve	[-]
M_{75}	Powder accumulated in 75 μm sieve	[-]
M_{90}	Powder accumulated in 90 μm sieve	[-]
M_{106}	Powder accumulated in 106 μm sieve	[-]
M_{a}	Arithmetic mean particle size	μm
M_{g}	Geometric mean particle size	μm
N	Number of sieving measurements for a single sample	[-]
p	percentile	[-]
P_{b}	Constructed powder distribution, bimodal large variance	[-]
P_{n}	Constructed powder distribution, unimodal small variance	[-]
P_{u}	Constructed powder distribution, unimodal large variance	[-]
q_3	Mass density distribution	[-]
$q_{3i}(M_{[.]})$	the i^{th} element of the mass density function for the powder sample $\mathrm{M}_{[.]}$	[-]
Q_3	Cumulative mass distribution	[-]
rpm	Revolutions per minute	[-]
w_k	vector of normalised mass fractions of k^{th} sample	[-]
w_{ki}	normalised mass fraction of k^{th} sample in the i^{th} sieve class	[-]
x_i	i^{th} sieve class diameter	μm
x_i'	Natural log of i^{th} sieve class	[-]

Greek symbols

β_2	Kurtosis	[-]
γ_1	Skewness	[-]
μ_{fit}	Lognormal distribution location parameter	[-]
σ_{fit}	Lognormal distribution shape parameter	[-]
σ_{a}^2	Arithmetic variance	$\mu\mathrm{m}^2$
σ_{a}	Arithmetic standard deviation	$\mu\mathrm{m}$
σ_{g}	Geometric standard deviation	[-]

References

[1] T. Abberger, A. Seo, and T. Schaefer. The effect of droplet size and powder particle size on the mechanisms of nucleation and growth in fluid bed melt agglomeration. *International Journal of Pharmaceutics*, 249(1–2):185–197, 2002. doi:10.1016/S0378-5173(02)00530-6.

[2] D. Ameye, E. Keleb, C. Vervaet, J. P. Remon, E. Adams, and D. Massart. Scaling-up of a lactose wet granulation process in Mi-Pro high shear mixers. *European Journal of Pharmaceutical Sciences*, 17(4–5):247–251, 2002. doi:10.1016/S0928-0987(02)00218-X.

[3] K. Ax, H. Feise, R. Sochon, M. Hounslow, and A. Salman. Influence of liquid binder dispersion on agglomeration in an intensive mixer. *Powder Technology*, 179 (3):190–194, 2008. doi:10.1016/j.powtec.2007.06.010.

[4] S. I. F. Badawy and M. A. Hussain. Effect of starting material particle size on its agglomeration behavior in high shear wet granulation. *AAPS PharmSciTech*, 5(3): 16–22, 2004. doi:10.1208/pt010433.

[5] S. I. F. Badawy, T. J. Lee, and M. M. Menning. Effect of drug substance particle size on the characteristics of granulation manufactured in a high-shear mixer. *AAPS PharmSciTech*, 1(4):55–61, 2000. doi:10.1208/pt010433.

[6] A. Braumann and M. Kraft. Incorporating experimental uncertainties into multivariate granulation modelling. *Chemical Engineering Science*, 65(3):1088–1100, 2010. doi:10.1016/j.ces.2009.09.063.

[7] A. Braumann, M. J. Goodson, M. Kraft, and P. R. Mort. Modelling and validation of granulation with heterogeneous binder dispersion and chemical reaction. *Chemical Engineering Science*, 62(17):4717–4728, 2007. doi:10.1016/j.ces.2007.05.028.

[8] A. Braumann, M. Kraft, and P. R. Mort. Parameter estimation in a multidimensional granulation model. *Powder Technology*, 197(3):196–210, 2010. doi:10.1016/j.powtec.2009.09.014.

[9] A. Braumann, M. Kraft, and W. Wagner. Numerical study of a stochastic particle algorithm solving a multidimensional population balance model for high shear granulation. *Journal of Computational Physics*, 229(20):7672–7691, 2010. doi:10.1016/j.jcp.2010.06.021.

[10] A. Braumann, P. L. W. Man, and M. Kraft. Statistical approximation of the inverse problem in multivariate population balance modeling. *Industrial & Engineering Chemistry Research*, 49(1):428–438, 2010. doi:10.1021/ie901230u.

[11] A. Braumann, P. L. W. Man, and M. Kraft. The inverse problem in granulation modelling – two different statistical approaches. *AIChE Journal*, in press, 2011. doi:10.1002/aic.12526.

[12] M. Goodson, M. Kraft, S. Forrest, and J. Bridgwater. A multi-dimensional population balance model for agglomeration. In *PARTEC 2004 - International Congress for Particle Technology*, 2004.

[13] P. Holm, T. Schaefer, and C. Larsen. End-point detection in a wet granulation process. *Pharmaceutical Development and Technology*, 6(2):181–192, 2001. doi:10.1081/PDT-100000739.

[14] A. Johansen and T. Schaefer. Effects of physical properties of powder particles on binder liquid requirement and agglomerate growth mechanisms in a high shear mixer. *European Journal of Pharmaceutical Sciences*, 14(2):135–147, 2001. doi:10.1016/S0928-0987(01)00164-6.

[15] A. Johansen and T. Schaefer. Effects of interactions between powder particle size and binder viscosity on agglomerate growth mechanisms in a high shear mixer. *European Journal of Pharmaceutical Sciences*, 12(3):297–309, 2001. doi:10.1016/S0928-0987(00)00182-2.

[16] J. R. Jones and J. Bridgwater. A case study of particle mixing in a ploughshare mixer using Positron Emission Particle Tracking. *International Journal of Mineral Processing*, 53(1–2):29–38, 1998. doi:10.1016/S0301-7516(97)00054-9.

[17] C. A. Kastner, G. P. E. Brownbridge, S. Mosbach, and M. Kraft. Impact of powder characteristics on a particle granulation model. *Chemical Engineering Science*, 97: 282–295, 2013. doi:10.1016/j.ces.2013.04.032.

[18] P. C. Knight, T. Instone, J. M. K. Pearson, and M. J. Hounslow. An investigation into the kinetics of liquid distribution and growth in high shear mixer agglomeration. *Powder Technology*, 97(3):246–257, 1998. doi:10.1016/S0032-5910(98)00031-X.

[19] H. G. Kristensen, P. Holm, and T. Schaefer. Mechanical properties of moist agglomerates in relation to granulation mechanisms Part II. Effects of particle size distribution. *Powder Technology*, 44(3):239–247, 1985. doi:10.1016/0032-5910(85)85005-1.

[20] J. D. Litster, K. P. Hapgood, J. N. Michaels, A. Sims, M. Roberts, S. K. Kameneni, and T. Hsu. Liquid distribution in wet granulation: dimensionless spray flux. *Powder Technology*, 114(1–3):32–39, 2001. doi:10.1016/S0032-5910(00)00259-X.

[21] M. B. Mackaplow, L. A. Rosen, and J. N. Michaels. Effect of primary particle size on granule growth and endpoint determination in high-shear wet granulation. *Powder Technology*, 108(1):32–45, 2000. doi:10.1016/S0032-5910(99)00203-X.

[22] P. L. W. Man, A. Braumann, and M. Kraft. Resolving conflicting parameter estimates in multivariate population balance models. *Chemical Engineering Science*, 65(13):4038–4045, 2010. doi:10.1016/j.ces.2010.03.042.

[23] C. Mangwandi, M. Adams, M. Hounslow, and A. Salman. An investigation of the influence of process and formulation variables on mechanical properties of high shear granules using design of experiment. *International Journal of Pharmaceutics*, 427 (2):328–336, 2012. doi:10.1016/j.ijpharm.2012.02.029.

[24] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL http://www.R-project.org/. ISBN 3-900051-07-0.

[25] N. Rahmanian, M. Ghadiri, and X. Jia. Seeded granulation. *Powder Technology*, 206(1–2):53–62, 2011. doi:10.1016/j.powtec.2010.07.011.

[26] R. Ramachandran, M. A. Ansari, A. Chaudhury, A. Kapadia, A. V. Prakash, and F. Stepanek. A quantitative assessment of the influence of primary particle size polydispersity on granule inhomogeneity. *Chemical Engineering Science*, 71:104–110, 2012. doi:10.1016/j.ces.2011.11.045.

[27] T. Schaefer. Melt pelletization in a high shear mixer. X. Agglomeration of binary mixtures. *International Journal of Pharmaceutics*, 139(1–2):149–159, 1996. doi:10.1016/0378-5173(96)04615-7.

[28] T. Schaefer, D. Johnsen, and A. Johansen. Effects of powder particle size and binder viscosity on intergranular and intragranular particle size heterogeneity during high shear granulation. *European Journal of Pharmaceutical Sciences*, 21(4):525–531, 2004. doi:10.1016/j.ejps.2003.12.002.

[29] US Food and Drug Administration. *Guidance for Industry: Q8 Pharmaceutical Development*. Rockville, MD, May 2006.

[30] K. Zuurman, K. A. Riepma, G. K. Bolhuis, H. Vromans, and C. F. Lerk. The relationship between bulk density and compactibility of lactose granulations. *International Journal or Pharmaceutics*, 102(1–3):1–9, 1994. doi:10.1016/0378-5173(94)90033-7.