White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/171834/

Version: Accepted Version

Article:
Yeardley, A.S., Bellinghausen, S., Milton, R. et al. (2021) Efficient global sensitivity-based model calibration of a high-shear wet granulation process. Chemical Engineering Science, 238. 116569. ISSN: 0009-2509

https://doi.org/10.1016/j.ces.2021.116569

© 2021 Elsevier. This is an author produced version of a paper subsequently published in Chemical Engineering Science. Uploaded in accordance with the publisher's self-archiving policy. Article available under the terms of the CC-BY-NC-ND licence (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Efficient Global Sensitivity-Based Model Calibration of a High-Shear Wet Granulation Process

Aaron S. Yeardley, Stefan Bellinghausen, Robert Milton, James D. Litster, Solomon F. Brown*

Department of Chemical and Biological Engineering, The University of Sheffield, Sheffield S10 2TN

Abstract

Model-driven design requires a well-calibrated model and therefore needs efficient workflows to achieve this. This efficiency can be achieved with the identification of the critical process parameters (CPPs) and the most impactful modelling parameters followed by a targeted experimental campaign to prioritise the calibration of these. To identify these parameters it is essential to perform a global sensitivity analysis (GSA).

Here, an efficient GSA is applied to a wet granulation case study with the Sobol' indices used to identify the CPPs and impactful modelling parameters. The population balance, mechanistic model that is used requires considerable computational effort for a GSA so a Gaussian Process surrogate is utilised to interrogate the underlying model. These key results reduce the input-space by 80% enabling the proposal of a targeted experimental design and model calibration workflow. This substantially improves the ability to deploy model-based design to determine the impactful parameter values, reducing the experimental effort by 42.1% compared to a conventional experimental design.

Keywords: Gaussian Process, Sobol' Indices, Global Sensitivity Analysis, Model Calibration, Granulation, Experimental Design

1. Introduction

A common experimental design approach is the factorial design of experiments (DOE). In a factorial DOE, the experimental space is covered by identifying a high and a low level for every input factor (Montgomery and Runger, 2014). Every combination of input factor levels is tested experimentally and so, as a consequence, the experimental effort increases exponentially with the number of factors. In order to reduce the required experimental effort, the most critical process parameters (CPPs) should be identified *a priori* so only they are used as factors. However, the CPPs are commonly identified heuristically which can be unreliable if the process-specific experience is limited, e.g. due to recent process modifications or new formulations. Gaining experience through a rigorous experimental

*Corresponding author.
E-mail address: s.f.brown@sheffield.ac.uk

investigation of all process parameters requires a very high experimental effort. To identify the CPPs, a sensitivity analysis is proposed by using a predictive model that reduces the need for heuristics. However, models for particulate processes and product design are coupled with many parameters and degrees of freedom. Experimentally testing every combination of input variables, across 3 to 5 length scales makes designing particulate products a costly process in terms of time, money and materials. Consequently, researchers using experimental design methods overcame these issues using DOE and began developing more sophisticated methods such as the sequential approach. These methods are much more adaptive, as they offer a dynamic class of experimental design by incorporating system knowledge through progressive steps (Garud et al., 2017). However, these more sophisticated methods do not allow the opportunity to incorporate uncertainty using of a substantially reduced number of experiments. Therefore, this work focuses on the development of an application of a model-driven design workflow to identify the CPPs and thus reduce, and better target, the number of experiments to be performed using high shear wet granulation (HSWG) as a case study.

By applying a model-driven design workflow, the most important operating ranges are identified by evaluating model predictions to reduce the experimental effort further (Wang et al., 2019; Bellinghausen, 2020). The process model is based on a population balance modelling (PBM) framework. Mechanistic understanding of the rate processes is incorporated through appropriate kernels. An essential part of model-driven design is model calibration workflows, which typically involve a combination of designed characterisation tests, parameter estimation from lab-scale experiments and reasonable assumptions to determine all modelling parameters. It is important, therefore, to ensure that impactful modelling parameters are identified upfront, so that an efficient model calibration workflow can be applied (Bellinghausen, 2020). This is because the impactful parameters need to be determined more accurately, while a lower accuracy is acceptable for the remaining parameters. As such, by focussing on the impactful parameters, the experimental or computational effort to determine parameters is reduced advancing towards improved model predictions.

While there are a number of types of process models, in this study we will focus on PBM as it is the most frequently used method for particulate processes such as crystallisation (Costa et al., 2007; Sulttan and Rohani, 2019), polymerisation (Sood et al., 2016; Brunier et al., 2017), granulation (Meyer et al., 2015; Shirazian et al., 2019), milling (Kumar Akkisetty et al., 2010; Capece et al., 2011), and mixing (Sen et al., 2012; Boukouvala et al., 2012). PBM keeps track of particle properties over time using population balance equations as a process scaling approach. Often, models are validated by comparing simulation results to the same experimental results used for the parameter estimation study and so they are not critically assessed. For process design and scale up purposes it it essential that models are fully validated using data that has not been used to train the model (Chaudhury et al., 2014). Therefore, the

need to develop efficient methods to reduce the cost of DOE for HSWG is constrained by the lack of understanding of both the modelling parameters and the experimental parameters. Hence, a model-driven design workflow that will identify and further the understanding of the CPPs is essential.

Critical modelling decisions in model-driven design need to be based on a good understanding of the process model. Additional insight is given by a sensitivity analysis that characterises the relationship between the model's inputs and outputs. Thus, the uncertainty in the outcome can be apportioned to the different sources of uncertainty in the input. Sensitivity analysis has been implemented in many different research fields (for example Saltelli et al., 2005; Rohmer and Foerster, 2011; Al et al., 2019) as it is widely acknowledged as a good practise to better understand model behaviour. In particular, the use of a global sensitivity analysis (GSA) quantifies the variation of the model response in the entire parameter domain fully exploring the input space to identify the CPPs of interest. Sensitivity studies have been employed to advance particulate processes (Van Bockstal et al., 2018; Mortier et al., 2014) as it is an effective tool to rank and prioritise the process variables and the modelling parameters helping to focus experimental and model calibration efforts. Prioritising CPPs allow a reduction from a high dimensional model, hence further tools such as optimisation (Wang et al., 2017) can be applied to aid the development of particulate processing for design.

Focussing on granulation, sensitivity analysis studies have been applied to help determine the most important factors affecting the granulation output. For instance, Cryer and Scherer (2003) conducted a sensitivity analysis on fluid-bed granulation, beginning with a $1/2$ fraction factorial statistical DOE. The research determined the important factors affecting the granule size to be bed bowl charge, binder spray rate, air flow rate, and input air temperature. Then a PBM model sensitivity analysis was used to further the understanding available from the experimental DOE alone, suggesting 65% of the predicted variance is accounted for from the binder spray rate. Similarly, Metta et al. (2019) applied a GSA to further understand the complex interplay between process wide CPPs and critical quality attributes using a flowsheet model. Wet granulation was included in the continuous tablet manufacturing process and modelled using PBM in the model flowsheet. The GSA was achieved using the Morris method and the variance based method of Sobol' indices, both agreeing that the liquid feed rate to the granulator to be a CPP which affects the tablet properties. Interestingly, the work also showed significantly less samples were needed for the Morris method but it did not provide the detailed and quantitative information that Sobol' indices does. This research successfully identifies the CPPs throughout the continuous tablet manufacturing process but these studies were constrained by a computationally expensive flowsheet. The existing research has many problems in representing a high resolution for the twin screw wet granulation model, predicting only steady state outputs.

Therefore, this work will address the shortcomings from previous studies by concentrating on the HSWG case study as a single unit operation model. Although there are many sensitivity analysis techniques available (such as the

Morris method), this work requires a detailed understanding of the importance of each parameter and their interactions. Consequently, this work will use a variance-based decomposition GSA method known as the Sobol' sensitivity indices which are considered the benchmark for GSA methods (Sumner et al., 2012; Kontoravdi et al., 2005; Xie et al., 2019). However, as noted by Metta et al. (2019), the calculations require significant amounts of data to ensure convergence of integrals to a satisfactory precision level. Due to the complexity of the underlying model, the traditional method of computation using Monte Carlo techniques (Kucherenko et al., 2009) are computationally impracticable for this research. Various surrogate modelling techniques reduce the computational burden by allowing the indirect interrogation of the model so that a significant reduction in the number of simulations can be used to evaluate the high-dimensional integrals. Examples of surrogate modelling techniques include polynomial chaos expansion (Brown et al., 2013; Sudret, 2008), artificial neural networks (Li et al., 2016), and Gaussian Processes (GPs) (Marrel et al., 2009; Yeardley et al., 2020a,b). Here, based on our previous work, the surrogate model will be developed using GPs as they are a widely used tool for Bayesian nonlinear regression and provide an approach that predicts a distribution allowing for uncertainties for each prediction. Furthermore, GP regression (also known as Kriging) has been proven for predictive modelling of pharmaceutical processes (Jia et al., 2009; Boukouvala et al., 2010). The novel approach will allow us to:

- identify of the most impactful modelling parameters for the HSWG process model,
- understand which of the operating parameters for the process are the CPPs,
- determine the parameters that will have a sufficient impact to influence the DOE independently and from interactions,
- investigate the more beneficial production conditions and a reduction in experimental effort.

These goals are achieved by developing a population balance based HSWG model that is directly emulated by a GP surrogate model enabling a reduction in the computational effort required to analytically calculate the Sobol' indices for a GSA.

2. Modelling Tools

To perform the GSA of the HSWG, we use PBM to model the HSWG, then GP regression acts as a surrogate model allowing the calculation of Sobol' indices. This section will describe the background behind each modelling tool. The nomenclature of the mathematical notation shown in the paper follows that of standard mathematics, where a bold lower case variable represents a vector and a bold upper case variable represents a matrix.

2.1. *High-Shear Wet Granulation Process Model*

This section summarises the process model for HSWG proposed by (Bellinghausen, 2020), as shown by the process schematic in Figure 1.

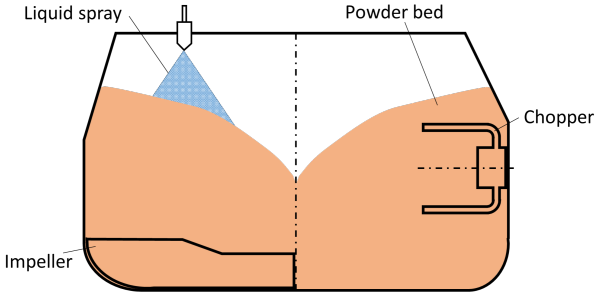


Figure 1: HSWG process schematic

The model is based on a 1-D, 1-compartment population balance modelling (PBM) framework (Ramkrishna and Mahoney, 2002) applying the lumped parameter approach (Hounslow et al., 2001). The particle size distribution is tracked by determining the particle number density n for every size bin i over time t:

$$\frac{\partial V n(v,t)}{\partial t} + \frac{\partial}{\partial v}\left[V n(v,t)\left(\dot{G}_{lay} + \dot{G}_{cons}\right)\right] = V\left[\dot{b}_{nuc}(v) + \dot{b}_{coal}(v) + \dot{b}_{br}(v) - \dot{d}_{coal}(v) - \dot{d}_{br}(v)\right] \quad (1)$$

where n is the volume-specific number density of particles, v is the particle volume, V is the control volume, $\dot{G}_{lay}$ and $\dot{G}_{cons}$ are the rate of change due to layering and consolidation respectively, $\dot{b}_{nuc}$, $\dot{b}_{coal}$ and $\dot{b}_{br}$ are the birth rates due to nucleation, coalescence and breakage, and $\dot{d}_{coal}$ and $\dot{d}_{br}$ are the death rates due to coalescence and breakage. Rate expressions for nucleation, consolidation, coalescence and breakage are incorporated in the model using inputs that determine the effect of each of these rate processes on the particle size. These rate expressions and underlying modelling assumptions are described in detail by (Bellinghausen, 2020). The model is implemented in gFormulate v1.5 (Process Systems Enterprise Ltd.).

2.2. *Gaussian Process (GP) Regression*

GP regression is used as a surrogate model for the PBM. This non-parametric machine learning technique enables direct interrogation of the PBM to be replaced by a reduced model encapsulating the systems behaviour in a cheaper, simpler framework.

The GP takes a $(1 \times d)$ row vector of inputs $\mathbf{x}$ and returns a Gaussian random variable through calculations using the predictive equations shown in Yeardley et al. (2020a). At the heart of this lies the kernel function $k\colon \mathbb{R}^{i+d} \times \mathbb{R}^{j+d} \rightarrow \mathbb{R}^{i} \times \mathbb{R}^{j}$, expressing the correlation between responses to input samples of sizes $(i \times d)$ and $(j \times d)$. This work

exclusively uses the automatic relevance determination (ARD) kernel (Wipf and Nagarajan, 2008):

$$k(\mathbf{x}', \mathbf{x}) := \sigma_f^2 \exp\left(-\frac{(\mathbf{x}-\mathbf{x}')\Lambda^{-2}(\mathbf{x}-\mathbf{x}')^\intercal}{2}\right) \tag{2}$$

where Λ is a $(d \times d)$ *diagonal* positive definite lengthscale matrix. The GP surrogate model is learnt from mapping the training inputs $\mathbf{X}$ to the observed responses $\mathbf{y}$, assuming the training data takes the form $\mathbf{y} = f(\mathbf{X}) + \mathbf{e}$ where $\mathbf{e}$ is an independent and identically distributed random error term. Regression uses the learned model to make predictions and so requires the optimisation of $d + 2$ hyperparameters, constituting of Λ, σ_f, and σ_e, through the maximum marginal likelihood $p[\mathbf{y}|\mathbf{X}]$ using the ROMCOMMA software library (ROMCOMMA, 2019).

2.3. Sobol' Indices

This work implements the use of the variance based GSA using a GP surrogate model to calculate the Sobol' Indices. The calculation of the Sobol' indices follows Jin et al. (2004) by substituting the true simulation model with the mean of the conditional GP resulting in semi-analytic Sobol' indices.

Sobol' indices are calculated as a variance-based GSA method, which describes how the variance of the output can be decomposed into terms that are dependent on the input factors (Iooss and Lemaître, 2015). Each input has two Sobol' indices calculated (Sobol, 1993, 2001). The first-order Sobol' index, S_i measures the contribution to the variance solely attributable to x_i. Whereas, the total Sobol' index, S_i^T expresses the whole effect of an input on the output, including interactions with all other inputs (Saltelli and Homma, 1996). Thus, the effect that interactions of an input variable has with other inputs on the variance on the output is calculated by the difference between S_i^T and S_i. Both indices are interpreted as the amount of variance ascribable by x_i and so the larger the value of the index, the more influential the input factor is.

The calculation of Sobol' indices are determined through a decomposition method presented by Sobol (1993), which evaluates each term through multidimensional integrals which requires a large sampling cost (Saltelli et al., 2008). Many alternative designs to calculate an estimation of Sobol' indices have been derived (Marrel et al., 2009; McKay et al., 2000). However, in this work we have calculated the Sobol' indices of the predicted value from a GP so that $y = \bar{f}(\mathbf{x})$, as shown by the mathematical details described in (Yeardley et al., 2020a).

3. Methodology

A suitable workflow was produced by applying a model-driven design (see Figure 2). GSA is the key to the model calibration and so both the modelling parameters and operating parameters undergo a GSA.

Figure 2 shows both types of parameters as separate pathways and so variations can be seen in both. Initially, modelling parameters may be known from physics or literature so if they do, those values can be used. If not a GSA

is needed to discover the most impactful modelling parameters. Whereas, it is always important to identify the CPPs within the operating parameters straight away using a GSA. Figure 2 presents the GSA methodology as one box for both the modelling parameters and the operating parameters because they are both implemented with the same method. After both GSA's a criterion is applied to distinguish between parameters with low or high impact dependent on the parameters Sobol' index. Determining which category each modelling and operating parameters falls into then helps determine the final steps for the model calibration. These final steps help determine the procedures necessary to efficiently calibrate the process model.

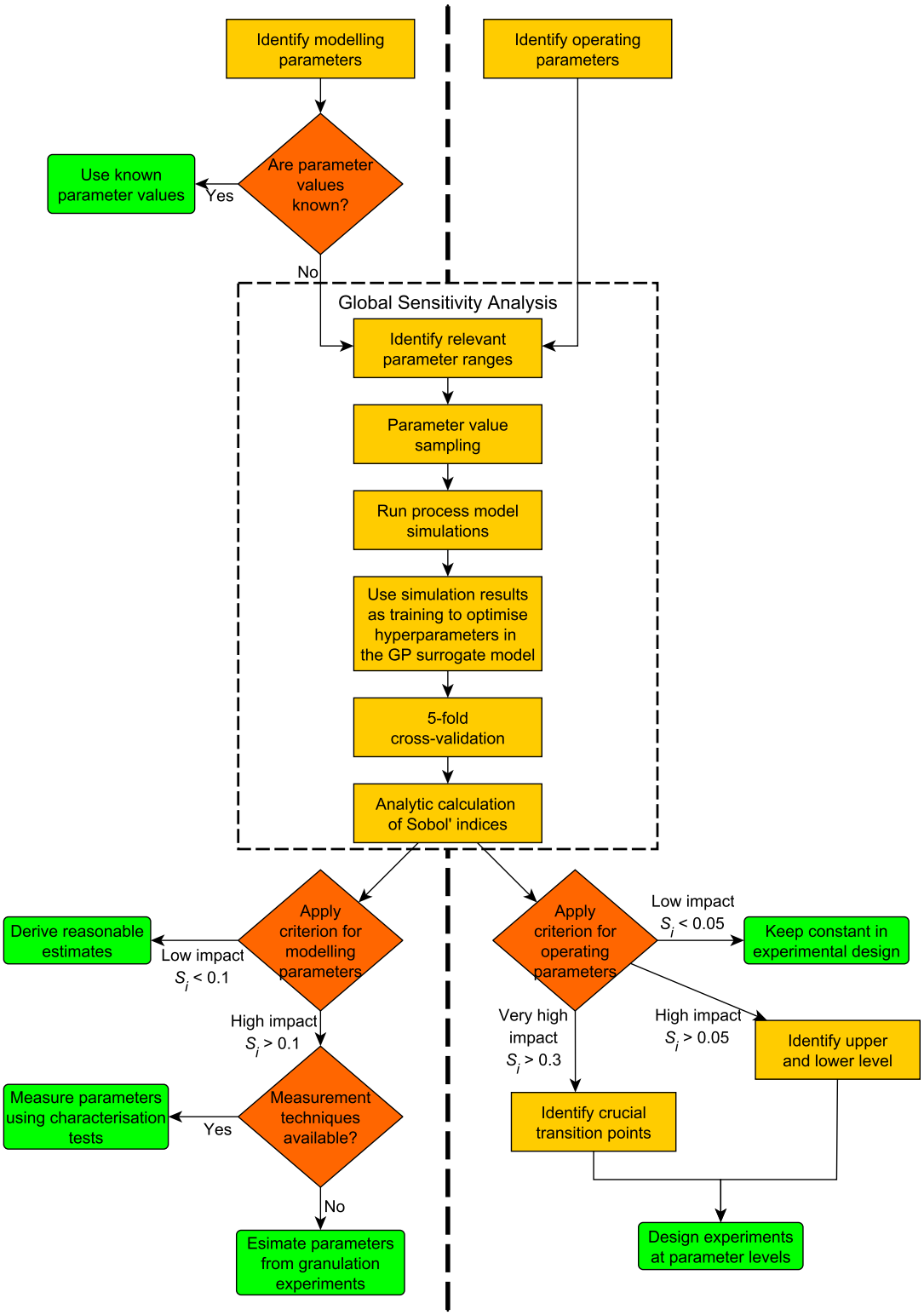


Figure 2: Model-driven design workflow using GSA with Sobol' indices criteria.

3.1. Parameter Identification and Sampling

For a GSA, representative model outputs need to be selected to assess the process performance. Particle size distribution and porosity are the most relevant product properties of wet granulation (Mittal, 2017). For this reason, the three properties of the particle size distribution, mass-median particle diameter $D50$, fines mass fraction $W\,(< 90\,\mu\mathrm{m})$ and coarse mass fraction $W\,(> 1\,\mathrm{mm})$, are selected as well as granule porosity ε.

The HSWG process model has been experimentally verified by Bellinghausen (2020) through an assessment of modelling assumptions and a comparison of model results to experimental trends. Individual wet granulation rate expressions used have been experimentally validated previously (Pohlman and Litster, 2015; Davis, 2016; Sayin, 2016; Bellinghausen et al., 2019). The results confirmed that the model is accurate across a wide range of conditions and process scales. In additional to the parameter estimation, model selection or validation is required to ensure the model is capable of predicting experimental results.

Therefore, initially a GSA is conducted as a screening tool using model runs with different combinations of modelling parameters. Then the model is run with different combinations of operating parameters for the second GSA. The parameter values for the process model simulations were produced using Quasi-random (Sobol') sampling with a normal distribution so that the input variables were based on reasonable uncertainty ranges. The GSA tool of gFormulate v1.5 is used for sampling and simulations. The modelling parameters and the rate process that they control are shown in Table 1. Additionally, their mean value and their standard deviations (STD) are listed with reasoning for each uncertainty range.

Table 2 shows the operating parameters mean and STD used for sampling the 10 l high-shear mixer filled with 2 kg of dry powder case study. Engineering rules for high-shear wet granulation were applied to derive the operating parameter ranges which are specific to the process scale investigated.

Table 1: The input parameters used for the Modelling Parameter GSA

Rate Process	Modelling Parameter	Mean	STD	Source
Nucleation	Nuclei-to-Drop Diameter Ratio K_d $[-]$	1.4	0.1	Hapgood et al. (2009)
Consolidation	Consolidation Coefficient k_{cons} $[-]$	1.5	0.25	Assumption
Consolidation/Coalescence	Dynamic Contact Angle θ [rad]	0.2	0.1	Pohlman and Litster (2015)
	Strength Parameter A $[-]$	7.0	0.5	Smith (2007)
	Strength Parameter B $[-]$	220	9	Smith (2007)
	Strength Parameter n $[-]$	0.59	0.01	Smith (2007)
	Surface Tension γ^{lv} $[\mathrm{mN\,m^{-1}}]$	73.9	0.1	Pallas and Harrison (1990)
Coalescence	Critical Pore Saturation S_{crit} $[-]$	0.16	0.01	Assumption
	Collision Coefficient $k_{I/II}$ $[\log_{10}\mu\mathrm{m}^{1.5}]$	13.5	0.5	Pohlman (2015)
	Elasticity Parameter k_E $[-]$	24.9	1.8	Pohlman and Litster (2015)
	Elasticity Parameter p_E $[-]$	0.17	0.01	Pohlman and Litster (2015)
	Elasticity Parameter q_E $[-]$	-6.9	0.6	Pohlman and Litster (2015)
	Elasticity Parameter r_E $[-]$	-1.50	0.015	Pohlman and Litster (2015)
	Height of Asperities h_a [µm]	1.51	0.2	Kastner et al. (2013)
	Poisson Ratio ν $[-]$	0.03	0.002	Pohlman and Litster (2015)
Breakage	Impact Energy $E_{m,kin}$ $[\mathrm{J\,kg^{-1}}]$	1000	50	Meier et al. (2008)
	Material Strength Parameter f_{Mat} $[\mathrm{kg\,m\,J^{-1}}]$	1.0	0.1	Meier et al. (2008)
	Breakage Coefficient k_{br} $[-]$	0.0035	0.001	Assumption
	Minimum Fragment Diameter $d_{j,min}$ [µm]	15	5	Vogel and Peukert (2005)
	Power Law Exponent q $[-]$	1.2	0.1	Assumption

Table 2: The input parameters used for the Operating Parameter GSA

Operating Parameter	Mean	STD	Design approach
Liquid to Solid Ratio L/S $[\mathrm{kg\,kg^{-1}}]$	0.17	0.015	Induction growth regime (Iveson et al., 2001)
Impeller Frequency n_{imp} $[\mathrm{min^{-1}}]$	321	32	Roping flow regime (Tran, 2015)
Kneading Time t_{kn} [s]	239	85	Induction time (Iveson et al., 2001)
Liquid Spray Rate $\dot{V}$ $[\mathrm{g\,min^{-1}}]$	71	9	Drop-controlled regime (Hapgood et al., 2003)

3.2. Application of GP Surrogate Modelling for the Calculation of Sobol' Indices

GP regression is used to encapsulate the HSWG process model in a much simpler framework enabling the semi-analytic calculation of Sobol' indices as shown by the modelling tools in Section 2. In this case study, the process model has four selected outputs and so the GP surrogate model captures each output as four independent GP regression predictive tools. In the first surrogate model twenty modelling parameters are used as a (1×20) row vector of inputs to predict the distribution of the four process outputs. Similarly, the four operating parameters are used in a second surrogate model. These predictions are made using a formula that is learnt from training data produced through sampling. GP regression involves optimising the hyperparameters using the ROMCOMMA software library (ROMCOMMA, 2019). From there, the predictive equation is used to calculate Sobol' indices but first cross-validation ensures inaccuracies in the predictive capabilities of the GP surrogate models are not carried through to the GSA. Therefore, this research tested the GP surrogate model using 5-fold cross-validation, whereby, the training data is randomly split into five subsets. Then four of the five subsets are used as training data, while the remaining is used to test the GP surrogate models predictions. This procedure is repeated so that every single sampling point produced is used for testing.

3.3. Parameter Criterion

The threshold for each GSA criteria will be set by the average total Sobol' index value for the four outputs, $\hat{S}_i^T$. Firstly, the modelling parameters GSA acts a screening that uses a threshold of $\hat{S}_i^T > 0.1$. Any modelling parameters that fail the criteria are not impactful and can have reasonable values from literature used.

Similarly, the threshold for the operating parameters GSA is set so that CPPs have a value of $\hat{S}_i^T > 0.05$. Then these CPPs are used in the experimental design with various levels dependent on how much of an impact each has. The focus of the experimental design should be on operating parameters with a very decisive impact ($\hat{S}_i^T > 0.3$). This is achieved by introducing multiple levels for these parameters. The operating parameters with $\hat{S}_i^T < 0.05$ are kept constant in the experimental design. The different thresholds for the modelling parameters GSA and operating parameters GSA are due to the difference in the number of variables. There are a total of twenty modelling parameters in comparison to the four operating parameters and so the threshold for the modelling parameter GSA needs to be larger ensuring an efficient reduction in parameters.

3.4. Experimental Design Considerations

Based on the GSA results, an experimental design for model calibration and validation is proposed. Impactful modelling parameters that can be directly derived from physical properties need to be determined by characterisation tests. The remaining impactful modelling parameters need to be estimated from experimental data. To reduce the

number of experiments and understand the impact of the operating conditions, all CPPs are varied in the experimental design but the very impactful parameters ($\hat{S}_i^T > 0.3$) are studied in more depth through smaller step changes. This more rigorous investigation of the very impactful CPPs is also needed to accurately estimate and validate the more impactful modelling parameters.

4. Results

4.1. Modelling Parameter Screening

4.1.1. GP Validation

The GP surrogate model was used to predict the outputs of the HSWG model using the twenty modelling parameters shown in Table 1. To be able to confidently calculate the Sobol' indices using a surrogate model, the predictions produced by the GPs must be fully validated against the true simulations to ensure inaccuracies are not inherently produced in the GSA. The GP surrogate model was tested using the 5-fold cross-validation technique (Hastie, 2009) and the resulting diagnostic values are presented in Table 3. The residuals of each prediction are plotted in Figure 3, comparing the true values with the predicted mean values showing a correlation that follows the red $y = x$ trend line. This is further exemplified by the coefficient of determination (r^2) and the root mean squared error (RMSE) values presented in Table 3 for each output. Both diagnostics measure a skill score, corresponding to the accuracy the predicted mean (Al-Taweel, 2018), but with different scales. Further, the predictive distribution is analysed to ensure the the GPs are not predicting with over confidence. This is shown by counting the outliers for any prediction where it's true standardised value is outside of the predictions 95% uncertainty distribution. Table 3 presents the outliers of all four outputs to be between 8.80% and 11.5% inferring the residuals are not Gaussian as the outliers are greater than the expected 5%. This shows the GPs are overconfident due some the predicted STD's being to small making the true validation values outside the predicted distribution. The prediction results here are in good agreement, and have shown the GP surrogate model is sufficiently accurate for the purposes of this work.

Table 3: Resulting diagnostic values from the modelling parameters prediction GPs

Output	**Correlation Coefficient,** r^2	**RMSE**	**Outliers at** 2 **STD's**
D50	0.661	0.622	11.5%
Fines Fraction	0.886	0.340	8.80%
Coarse Fraction	0.741	0.509	10.1%
Granule Porosity	0.919	0.285	9.90%

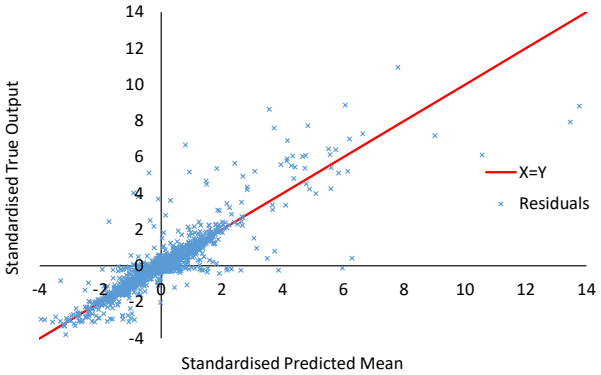


Figure 3: The residuals showing the observed standardised values against the predicted standardised values for all of the test predictions using cross-validation.

The results from the cross-validation for the modelling parameters are satisfactory but do not show accuracy where the GPs predict exact test values. This is because of an issue with surrogate models known as the curse of dimensionality, where the surrogate requires an exponentially increasing number of output results throughout the input space as the number of parameters increases.

4.1.2. Sobol' Indices

The results are the total Sobol' indices for the modelling GSA calculated for each output are shown in in Figure 4 where each bar corresponds to the first-order Sobol' indices plus the remaining indices due to interactions with other input variables. The split in the bars are to illustrate the first-order value as the bottom and the interactions as the top. Instantly, it is clear to see in Figure 4 that interactions have a large effect on the four outputs. Additionally, each output has different input variables influencing it with great variability. For example, the collision coefficient produces over 80% of the total variance for D50 but only 1% of the variance for the granule porosity. Most importantly, the total Sobol' index values for only eight of the twenty modelling parameters are large enough to be shown clearly in Figure 4 and so the remaining are grouped into the "remaining input variables" category.

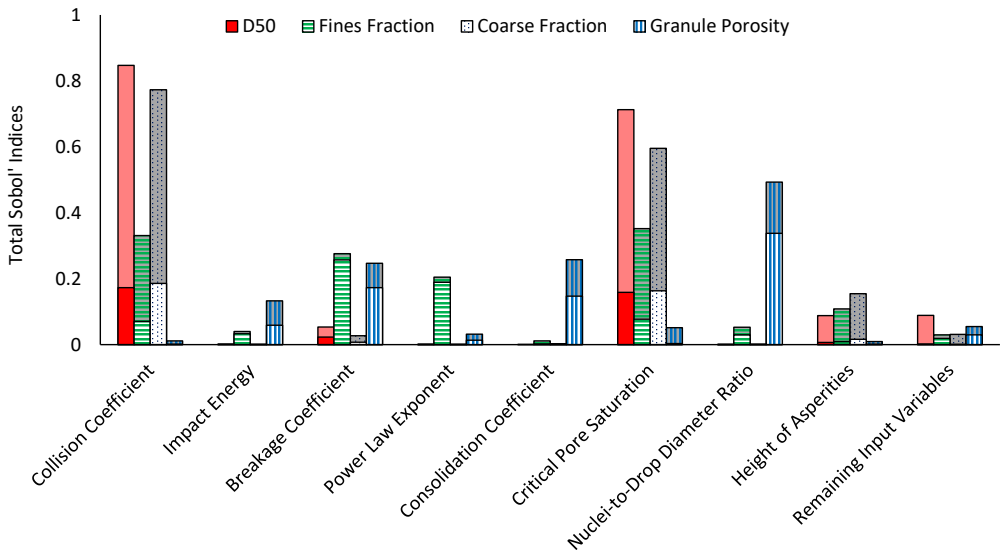


Figure 4: The total Sobol' indices for the modelling parameters with respect to each output of the HSWG model. Each bar is split to show the first-order Sobol' indices (bottom) and the interactions (top)

For each modelling parameter, further analysis calculating the average Sobol' index value per output ($\hat{S}_i^T$) shows which are the most impactful modelling parameters to be included in a parameter estimation. The criteria, set in Section 3, states the $\hat{S}_i^T > 0.1$ for i to be an impactful parameter. All the Sobol' index values shown in Figure 4 have there $\hat{S}_i^T$ presented in Table 4 clearly showing which output the modelling parameter affects the most by the colour of the shading.

Table 4: The impact each modelling parameter has on the four outputs where green is highly impactful and red is negligible. Summarised in the last column by the average Sobol' index value for each modelling parameter, $\hat{S}_i^T$

	D50 S_i^T	**Fines Fraction** S_i^T	**Coarse Fraction** S_i^T	**Granule Porosity** S_i^T	**Average** $\hat{S}_i^T$
Collision Coefficient					0.49
Impact Energy					0.04
Breakage Coefficient					0.15
Power Law Exponent					0.06
Consolidation Coefficient					0.07
Critical Pore Saturation					0.43
Nuclei-to-Drop Diameter Ratio					0.14
Height of Asperities					0.09
Remaining Input Variables					0.05

Table 4 shows an $\hat{S}_i^T$ value greater than 0.1 for just four of the twenty modelling parameters. Therefore, they can be reduced so that in a characterisation test or a parameter estimation analysis the collision coefficient, breakage coefficient, critical pore saturation and nuclei-to-drop diameter ratio are estimated accurately. Whereas the remaining modelling parameters can be fixed using default values from literature recommended previously as the means used in the GSA sampling.

Interestingly, the two most impactful parameters, collision coefficient and critical pore saturation are the most impactful coalescence parameters followed by the height of asperities which has a $\hat{S}_i^T$ just below the threshold set. The coalescence rate is proportional to the collision coefficient and the critical pore saturation impacts the coalescence criterion. Thus, both impactful modelling parameters impact the three particle size properties the most. Whereas, the impact the collision coefficient and the critical pore saturation has on the granule porosity is negligible as shown in Table 4 by the red shading. The strong interaction between the two parameters shows that coalescence depends heavily on both parameters. Additionally, the high impact of both coalescence parameters provides evidence that coalescence is the most dominant rate process for particle size.

The breakage coefficient dominates the impact energy and the power law exponent as the most most impactful breakage parameter. Overall, the breakage coefficient has a high impact on the granule porosity but less of an impact on the D50 and coarse fraction showing that breakage has a negligible effect on coalescence. Yet, it has a high impact

on fines fraction because the increase in breakage results in more fine particles.

The most impactful nucleation parameters is the nuclei-to-drop diameter ratio which determines the initial granule size and porosity. Due to the high impact of other rate processes like coalescence, the impact which nuclei-to-drop diameter has on size is negligible. Therefore, it mainly has a large impact on granule porosity.

4.2. Operating Parameters

4.2.1. GP Validation

A 5-fold cross-validation of the GPs used within the operating parameters GP surrogate model was performed. The GPs predicted each of the four HSWG model outputs using the four operating parameter inputs. A plot of the residuals shown in Figure 5 compares the observed values to the predicted mean values. The four GPs have a high predictive quality as the datapoints lie closely to the red $y = x$ trend line. Table 5 shows the three diagnostic skill scores showing accurate mean predictions through the r^2 and RMSE values. Further the predictions are neither overconfident nor underconfident as the predictions are shown to be normally distributed with the outliers now much closer to 5%. Overall, the high predictive accuracy shown from the GP surrogate models enables confidence in calculation of the Sobol' indices required in the GSA.

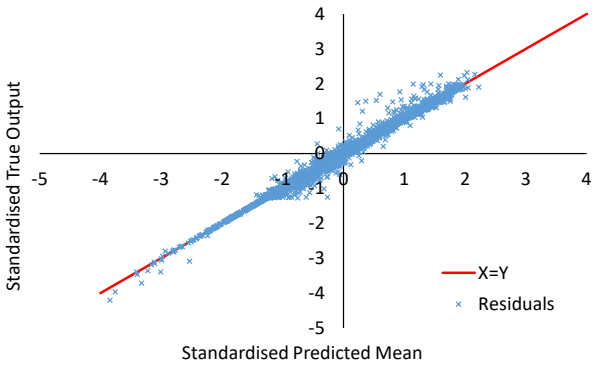


Figure 5: The residuals showing the observed standardised values against the predicted standardised values for all of the test predictions using cross-validation.

Table 5: Resulting diagnostic values from the operating parameters prediction GPs

Output	**Correlation Coefficient,** r^2	**RMSE**	**Outliers at** 2 **STD's**
D50	0.992	0.0869	6.00%
Fines Fraction	0.991	0.0959	6.05%
Coarse Fraction	0.996	0.0662	5.25%
Granule Porosity	0.999	0.0302	8.45%

4.2.2. Sobol' Indices

The Sobol' indices for each of the four operating parameters were calculated with respect to the four model outputs. The results are the total Sobol' indices which are presented in Figure 6 where each bar corresponds to the first-order Sobol' indices plus the remaining indices due to interactions with other input variables. Clearly, the liquid to solid ratio dominates all four model outputs, consistent with previous literature (Ameye et al., 2002). Both the impellar frequency and kneading time have a small influence. The liquid spray rate has negligible total Sobol' indices for all four outputs.

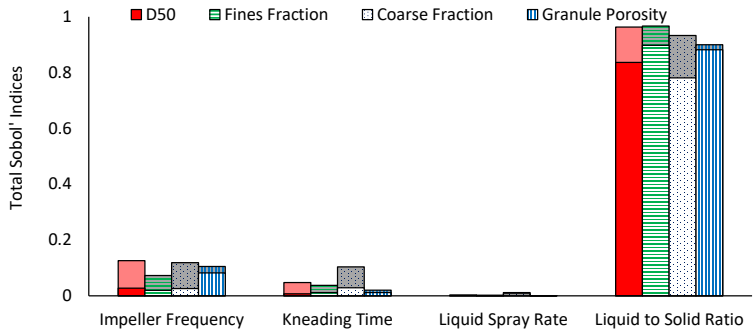


Figure 6: The Sobol' indices for the operating parameters with respect to each output.

The average Sobol' index value per output ($\hat{S}_i^T$) for each operating parameter are shown in Table 6 clearly showing the CPP's and the impact each operating parameter has on each model output from the colour of the shading.

Table 6: The impact each operating parameter has on the four outputs where green is highly impactful and red is negligible. Summarised in the last column by the average Sobol' index value for each modelling parameter, $\hat{S}_i^T$

	D50	**Fines Fraction**	**Coarse Fraction**	**Granule Porosity**	**Average**
	S_i^T	S_i^T	S_i^T	S_i^T	$\hat{S}_i^T$
Impeller Frequency					0.11
Kneading Time					0.05
Liquid Spray Rate					0.004
Liquid to Solid Ratio					0.94

An $\hat{S}_i^T = 0.94$ shows that the liquid to solid ratio has a very high impact and so Figure 2 shows the CPP should have its crucial transition points identified. Clearly the liquid to solid ratio is dominant throughout the HSWG process and so it is vital that it should be accurately studied and be the focus of the experimental design. Therefore, it is recommended to have four levels in smaller step changes to identify potential critical levels of the liquid to solid ratio,

The value for both the impeller frequency and the kneading time is $0.05 < \hat{S}_i^T < 0.3$ and so the experimental design must identify the upper and lower level of these CPPs. Interestingly, Figure 6 shows both the impeller frequency and kneading time to have 63% and 71%, respectively, of their $\hat{S}_i^T$ controlled by their interactions. These results are consistent with an experimental observation (Iveson and Litster, 1998) showing that at high liquid to solid ratio, the impeller frequency and kneading time have a higher impact because the high liquid to solid ratio aids in coalescence of particles, ensuring major impact from the particle size.

The liquid spray rate was only varied within the recommended drop-controlled regime. Within this narrow range, the liquid spray rate has been proven to have negligible impact on the output of the HSWG model. These findings are in agreement with data in the literature (Smrčka et al., 2015). Therefore, the liquid spray rate can be assumed constant because good experimental design recommendations are already available.

4.3. Experimental Design Proposal

The two GSA's have been used for the HSWG case study in an attempt to improve the experimental design. Here, we show the considerations that have to be made and how the analysis has reduced the work considerably. This section is in essence a demonstration of concept to show the practicality of this research.

A conventional factorial DOE is not very useful for parameter estimation. A better experimental design is needed that allows the identification of critical conditions. In this way, modelling parameters that depend on critical operating conditions can be accurately estimated in an efficient way. Therefore, the results presented from both GSA's helps guide the experimental design that would lead to a well-calibrated HSWG model. Only four of the twenty modelling parameters have a significant impact and so only the values for the impactful modelling parameters have to be determined accurately. This is done through the use of characterisation tests to derive the physical properties and parameter estimations for empirical parameter using lab-scale experimental data. An experimental investigation of operating parameters with a negligible impact is not beneficial for model calibration. Therefore, the proposal of a reduced experimental design only based on the CPPs will significantly reduce the experimental effort and still reach the goal of a well-calibrated process model.

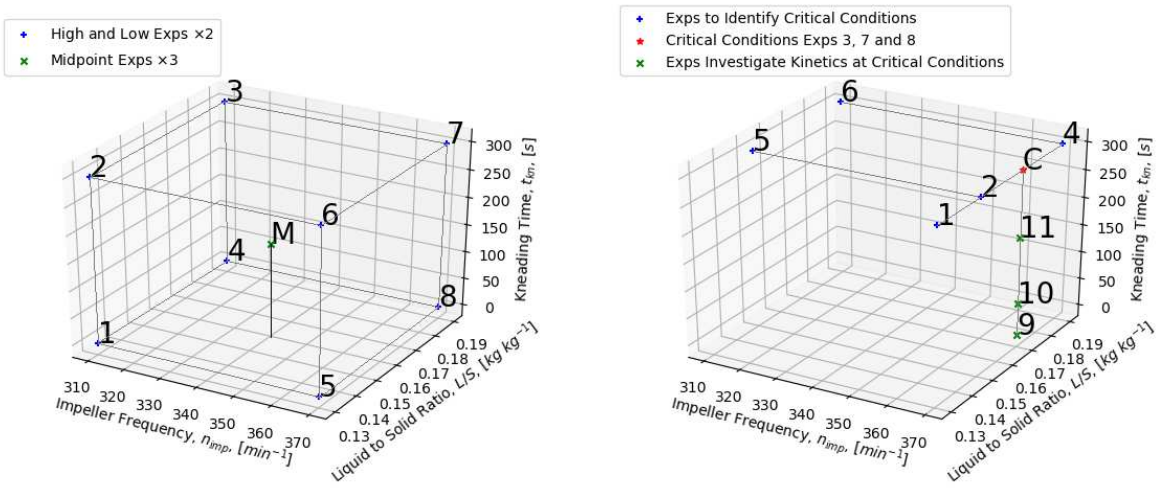
Using these results an experimental design proposal based on the GSA results was developed for a 10 l granulator which is shown in Table 7. The ranges for impeller frequency and liquid to solid ratio depend on the equipment scale and formulation, respectively. In Exp $1-4$, the liquid to solid ratio of the critical granulation conditions is determined. The critical granulation conditions are defined as the lowest liquid to solid ratio at which the critical pore saturation is reached and rapid growth occurs (Iveson et al., 2001). Rapid growth is identified by a significant increase in particle size $D50$ and the formation of particles that are much greater than 1 mm. The empirical minimum porosity correlation (Bellinghausen, 2020) is fitted using the results of Exp $1-6$. At critical conditions, triplicates are performed to assess

the reproducibility (Exp $7-8$). In Exp $9-11$, the kneading time is reduced at critical conditions to determine the critical time and investigate the process kinetics in more detail at critical conditions. This is limited to the critical conditions because the kneading time is expected to have the highest impact at these operating conditions. In this scenario, it is recommended to use the results at critical conditions (Exp 3 and $7-11$) to estimate the three impactful modelling parameters: collision and breakage coefficient, and critical pore saturation. The remaining experiments are available to validate the estimates and assess the predictive power of the model.

Table 7: Recommended experimental design for parameter estimation

Exp	n_{imp} $\left[\mathrm{min}^{-1}\right]$	L/S $\left[\mathrm{kg\,kg}^{-1}\right]$	t_{kn} [min]	Explanation
$1-4$	370 (high)	0.13, 0.15, 0.17, 0.19	5 (high)	Vary L/S to determine critical L/S
$5-6$	310 (low)	0.15, 0.19	5 (high)	and estimate the porosity parameters
$7-8$	370 (high)	0.17 (critical)	5 (high)	Triplicates to assess reproducibility
$9-11$	370 (high)	0.17 (critical)	0, 1, 3	Reduce t_{kn} to determine critical t_{kn}

Additionally, Figure 7 compares the proposed experimental design to a conventional experimental design that would have been conducted without the GSA results. The conventional experimental design is a 2^4 factorial DOE with three midpoint triplicates. The 3D scatter plot shown in Figure 7a shows eight high and low experiments with a midpoint experiment shown in the middle. It must be noted that there are three midpoint experiments and another eight high and low experiments for when the liquid spray rate value has changed. This conventional design therefore includes a total of 19 experiments. Whereas, the proposed experimental design shown in Figure 7b, consists of eight initial experiments to determine the critical liquid to solid ratio including three triplicates to asses reproducibility. Then an additional three experiments to determine the critical kneading time. Therefore, the proposed design of experiments reduces the conventional design from 19 experiments to 11 experiments. The workflow proposed reduces the experimental effort by 42.1%. Further, the proposed method has additional benefits as it will determine the critical pore saturation, an impactful parameter, whereas, the conventional method is not a very promising method with only two levels of each factor. Consequently, the proposed experimental design is a very promising approach to efficiently calibrate a HSWG process model.



(a) Conventional (b) Proposed

Figure 7: The experimental designs for the HSWG case study.

5. Conclusion

Designing a new product via particulate processes involves testing every combination of input variables and so an experimental study for high shear wet granulation (HSWG) is often a costly process in time, money and materials. To help, an appropriate process model is used within a model-driven design workflow. Therefore, a model calibration workflow is proposed based on the identified critical process parameters (CPPs) and most impactful modelling parameters. Thus better targeted experimental design can be derived with the goals to increase the usability for model calibration and reduce the experimental effort.

To do this, an efficient global sensitivity analysis (GSA) is proposed to calculate the Sobol' indices for the HSWG model. In order to make the computational times practical, a Gaussian Process (GP) surrogate model is utilised to avoid the need of a very significant number of model simulations for the evaluation of high-dimensional integrals. The novel approach is applied using a suitable workflow so that both the modelling and operating parameters undergo individual GSA's, as both sets of parameters need to be considered for the experimental design.

Criteria for both the modelling parameters and the operating parameters were set so that the average total Sobol' index value for the four outputs, $\hat{S}_i^T$, were used to reduce experimental effort. The most impactful modelling parameters were determined using a threshold of $\hat{S}_i^T > 0.1$ reducing the modelling parameters from twenty to just four. The collision coefficient, breakage coefficient, critical pore saturation and nuclei-to-drop diameter ratio were recognised as the most impactful modelling parameters where the values need to be estimated in a characterisation test or through parameter estimation. The operating parameters used a threshold of $\hat{S}_i^T > 0.05$ to determine the CPPs. The calculated Sobol' indices found the liquid to solid ratio to dominate the HSWG process and so it should be the

focus of the experimental design. Both the impeller frequency and kneading time were found to have an impact, primarily controlled by their interactions. Whereas, the liquid spray rate has negligible impact on the output of the process so it can be excluded from the experimental design by keeping the values constant from literature. Hence, we have recommended various levels for each operating parameter as factors in an experimental design in an attempt to estimate the most impactful modelling parameters that require parameter estimation for the model calibration of a HSWG process model. The results show the experimental design should be focused on the liquid to solid ratio and so four levels are recommended for this factor. Whereas, impeller frequency and kneading time can be analysed using just two levels. Hence we have shown that the HSWG process model can be efficiently calibrated through the use of a GP based GSA. Overall, the research has successfully completed the goals set out in Section 1 by identifying both the most impactful modelling parameters and the CPPs. From this, the impact of each parameter on a DOE has been analysed ensuring a proposal of a more efficient experimental design for the model calibration of a HSWG process. The proposal substantially improves the ability to quickly deploy model-based design by determining the critical pore saturation value while reducing the experimental effort of 42.1% in comparison to a conventional experimental design.

During this research we have applied the methodology and proposed an efficient experimental design that can now be experimentally validated. Conducting this parameter estimation with experimental data will fully calibrate the PBM model to a 10 l high-shear mixer filled with 2 kg of dry powder case study. Consequently, we have proposed a much more efficient experimental design that will indicate any unidentified subprocesses. On a more practical level, it would be beneficial to extend the work carried out in this research to any particulate processes, further reducing the experimental costs of designing new products. Additionally, the GP based sensitivity analysis method will be developed further to enable optimisation of the parameter values.

References

Al, R., Behera, C.R., Zubov, A., Gernaey, K.V., Sin, G., 2019. Meta-modeling based efficient global sensitivity analysis for wastewater treatment plants – An application to the BSM2 model. Computers & Chemical Engineering 127, 233–246. URL: https://doi.org/10.1016/j.compchemeng.2019.05.015, doi:10.1016/j.compchemeng.2019.05.015.

Al-Taweel, Y., 2018. Diagnostics and Simulation-Based Methods for Validating Gaussian Process Emulators. Phd thesis. University of Sheffield.

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

Bellinghausen, S., 2020. Modelling and scaling rules for high-shear wet granulation of pharmaceuticals. Ph.D. thesis. The University of Sheffield.

Bellinghausen, S., Gavi, E., Jerke, L., Ghosh, P.K., Salman, A.D., Litster, J.D., 2019. Nuclei size distribution modelling in wet granulation. Chemical Engineering Science: X 4, 100038.

Boukouvala, F., Dubey, A., Vanarase, A., Ramachandran, R., Muzzio, F.J., Ierapetritou, M., 2012. Computational approaches for studying the granular dynamics of continuous blending processes, 2–population balance and data-based methods. Macromolecular Materials and Engineering 297, 9–19.

Boukouvala, F., Muzzio, F.J., Ierapetritou, M.G., 2010. Predictive Modeling of Pharmaceutical Processes with Missing and Noisy Data. AIChE Journal 56, 2860–2872. doi:https://doi.org/10.1002/aic.12203, arXiv:0201037v1.

Brown, S., Beck, J., Mahgerefteh, H., Fraga, E.S., 2013. Global sensitivity analysis of the impact of impurities on CO2 pipeline failure. Reliability Engineering and System Safety 115, 43–54. URL: http://dx.doi.org/10.1016/j.ress.2013.02.006, doi:10.1016/j.ress.2013.02.006.

Brunier, B., Sheibat-Othman, N., Othman, S., Chevalier, Y., Bourgeat-Lami, E., 2017. Modelling particle growth under saturated and starved conditions in emulsion polymerization. The Canadian Journal of Chemical Engineering 95, 208–221.

Capece, M., Bilgili, E., Dave, R., 2011. Identification of the breakage rate and distribution parameters in a non-linear population balance model for batch milling. Powder Technology 208, 195–204.

Chaudhury, A., Barrasso, D., Pandey, P., Wu, H., Ramachandran, R., 2014. Population balance model development, validation, and prediction of CQAs of a high-shear wet granulation process: towards QbD in drug product pharmaceutical manufacturing. Journal of Pharmaceutical Innovation 9, 53–64.

Costa, C.B.B., Maciel, M.R.W., Maciel Filho, R., 2007. Considerations on the crystallization modeling: Population balance solution. Computers & Chemical Engineering 31, 206–218.

Cryer, S.A., Scherer, P.N., 2003. Observations and Process Parameter Sensitivities in Fluid-Bed Granulation. AIChE Journal 49, 2802–2809. doi:10.1002/aic.690491113.

Davis, N.J., 2016. Mechanical dispersion of semi-solid binders in high-shear granulation. Ph.D. thesis. Purdue University.

Garud, S.S., Karimi, I.A., Kraft, M., 2017. Design of computer experiments: A review. Computers and Chemical Engineering 106, 71–95. doi:10.1016/j.compchemeng.2017.05.010.

Hapgood, K.P., Litster, J.D., Smith, R., 2003. Nucleation regime map for liquid bound granules. AIChE Journal 49, 350–361.

Hapgood, K.P., Tan, M.X.L., Chow, D.W.Y., 2009. A method to predict nuclei size distributions for use in models of wet granulation. Advanced Powder Technology 20, 293–297.

Hastie, T., 2009. The elements of statistical learning : data mining, inference, and prediction. Springer series in statistics. 2nd ed. ed., Springer, New York.

Hounslow, M.J., Pearson, J.M.K., Instone, T., 2001. Tracer studies of high-shear granulation: II. population balance modeling. AIChE Journal 47, 1984–1999.

Iooss, B., Lemaître, P., 2015. A review on global sensitivity analysis methods. Operations Research/ Computer Science Interfaces Series 59, 101–122. doi:10.1007/978-1-4899-7547-8_5, arXiv:1404.2405.

Iveson, S.M., Litster, J.D., 1998. Fundamental studies of granule consolidation Part 2: Quantifying the effects of particle and binder properties. Powder Technology 99, 243–250.

Iveson, S.M., Wauters, P.A.L., Forrest, S., Litster, J.D., Meesters, G.M.H., Scarlett, B., 2001. Growth regime map for liquid-bound granules: further development and experimental validation. Powder Technology 117, 83–97.

Jia, Z., Davis, E., Muzzio, F.J., Ierapetritou, M.G., 2009. Predictive modeling for pharmaceutical processes using kriging and response surface. Journal of Pharmaceutical Innovation 4, 174–186. doi:10.1007/s12247-009-9070-6.

Jin, R., Chen, W., Sudjianto, A., 2004. Analytical metamodel-based global sensitivity analysis and uncertainty propagation for robust design. SAE Transactions Journal of Materials & Manufacturing doi:10.4271/2004-01-0429.

Kastner, C.A., Brownbridge, G.P.E., Mosbach, S., Kraft, M., 2013. Impact of powder characteristics on a particle granulation model. Chemical Engineering Science 97, 282–295.

Kontoravdi, C., Mantalaris, A., Asprey, S., Pistikopoulos, E., 2005. Application of the sobol global sensitivity analysis method to a dynamic model

of mab-producing mammalian cell cultures. Proceedings of the IASTED International Conference on Modelling, Identification, and Control, MIC , 361–366.

Kucherenko, S., Rodriguez-Fernandez, M., Pantelides, C., Shah, N., 2009. Monte Carlo evaluation of derivative-based global sensitivity measures. Reliability Engineering and System Safety 94, 1135–1148. doi:10.1016/j.ress.2008.05.006.

Kumar Akkisetty, P., Lee, U., Reklaitis, G.V., Venkatasubramanian, V., 2010. Population balance model-based hybrid neural network for a pharmaceutical milling process. Journal of Pharmaceutical Innovation 5, 161–168.

Li, S., Yang, B., Qi, F., 2016. Accelerate global sensitivity analysis using artificial neural network algorithm: Case studies for combustion kinetic model. Combustion and Flame 168, 53–64. doi:10.1016/j.combustflame.2016.03.028.

Marrel, A., Iooss, B., Laurent, B., Roustant, O., 2009. Calculations of Sobol indices for the Gaussian process metamodel. Reliability Engineering and System Safety 94, 742–751. doi:10.1016/j.ress.2008.07.008.

McKay, M.D., Beckman, R.J., Conover, W.J., 2000. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 42, 55–61. doi:10.1080/00401706.2000.10485979.

Meier, M., John, E., Wieckhusen, D., Wirth, W., Peukert, W., 2008. Characterization of the grinding behaviour in a single particle impact device: studies on pharmaceutical powders. European Journal of Pharmaceutical Sciences 34, 45–55.

Metta, N., Ghijs, M., Schäfer, E., Kumar, A., Cappuyns, P., Van Assche, I., Singh, R., Ramachandran, R., De Beer, T., Ierapetritou, M., Nopens, I., 2019. Dynamic flowsheet model development and sensitivity analysis of a continuous pharmaceutical tablet manufacturing process using the wet granulation route. Processes 7, 1–35. doi:10.3390/pr7040234.

Meyer, K., Bück, A., Tsotsas, E., 2015. Dynamic multi-zone population balance model of particle formulation in fluidized beds. Procedia engineering 102, 1456–1465.

Mittal, B., 2017. Chapter 4 - pharmaceutical unit operations, in: Mittal, B. (Ed.), How to Develop Robust Solid Oral Dosage Forms from Conception to Post-Approval. Academic Press, pp. 69 – 95. URL: http://www.sciencedirect.com/science/article/pii/B9780128047316000042, doi:https://doi.org/10.1016/B978-0-12-804731-6.00004-2.

Montgomery, D.C., Runger, G.C., 2014. Applied statistics and probability for engineers. John Wiley and Sons.

Mortier, S.T.F.C., Gernaey, K.V., Beer, T.D., Nopens, I., 2014. Global Sensitivity Analysis Applied to Drying Models for One or a Population of Granules. AIChE Journal 60, 1700–1717. doi:https://doi.org/10.1002/aic.14383, arXiv:0201037v1.

Pallas, N.R., Harrison, Y., 1990. An automated drop shape apparatus and the surface tension of pure water. Colloids and Surfaces 43, 169–194.

Pohlman, D.A., 2015. Multi-scale modeling of high-shear granulation. Ph.D. thesis. Purdue University.

Pohlman, D.A., Litster, J.D., 2015. Coalescence model for induction growth behavior in high shear granulation. Powder Technology 270, 435–444.

Ramkrishna, D., Mahoney, A.W., 2002. Population balance modeling. Promise for the future. Chemical Engineering Science 57, 595–606.

Rohmer, J., Foerster, E., 2011. Global sensitivity analysis of large-scale numerical landslide models based on Gaussian-Process meta-modeling. Computers and Geosciences 37, 917–927. URL: http://dx.doi.org/10.1016/j.cageo.2011.02.020, doi:10.1016/j.cageo.2011.02.020.

ROMCOMMA, 2019. https://github.com/C-O-M-M-A/rom-comma. [Online]. [Accessed on 25 November 2019].

Saltelli, A., Homma, T., 1996. Importance measures in global sensitivity analysis of model output. Reliab. Eng. Sys. Safety 52, 1–17.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global Sensitivity Analysis. The primer ed., John Wiley & Sons Ltd.

Saltelli, A., Ratto, M., Tarantola, S., Campolongo, F., 2005. Sensitivity Analysis for Chemical Models. Chemical Reviews 105, 2811–2828. URL: https://pubs.acs.org/doi/10.1021/cr040659d, doi:10.1021/cr040659d.

Sayin, R., 2016. Mechanistic studies of twin screw granulation. Ph.D. thesis. Purdue University.

Sen, M., Singh, R., Vanarase, A., John, J., Ramachandran, R., 2012. Multi-dimensional population balance modeling and experimental validation of continuous powder mixing processes. Chemical Engineering Science 80, 349–360.

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

Smith, R.M., 2007. Wet Granule Breakage in High Shear Mixer Graunlators. Ph.D. thesis. The University of Queensland.

Smrčka, D., Dohnal, J., Štěpánek, F., 2015. Effect of process scale-up on the dissolution of granules with a high content of active pharmaceutical ingredient. Powder Technology 285, 88–95.

Sobol, I.M., 1993. Sensitivity analysis for nonlinear mathematical models. Mathematical Modelling Computational Experiments 1, 407–414.

Sobol, I.M., 2001. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. Mathematics and Computers in Simulation , 271–280doi:https://doi.org/10.1016/S0378-4754(00)00270-6.

Sood, A., Awasthi, A., Bharti, R., 2016. A population balance model for butyl acrylate emulsion polymerization. Indian Chemical Engineer 58, 40–60.

Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions. Reliability Engineering and System Safety 93, 964–979. doi:10.1016/j.ress.2007.04.002.

Sulttan, S., Rohani, S., 2019. Coupling of CFD and population balance modelling for a continuously seeded helical tubular crystallizer. Journal of Crystal Growth 505, 19–25.

Sumner, T., Shephard, E., Bogle, I.D.L., 2012. A methodology for global-sensitivity analysis of time-dependent outputs in systems biology modelling. Journal of The Royal Society Interface 9, 2156–2166. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2011.0891, doi:10.1098/rsif.2011.0891, arXiv:https://royalsocietypublishing.org/doi/pdf/10.1098/rsif.2011.0891.

Tran, A.L.H., 2015. Powder flow in bertical high shear mixer granulators. Ph.D. thesis. The University of Queensland.

Van Bockstal, P.J., Mortier, S.T.F., Corver, J., Nopens, I., Gernaey, K.V., De Beer, T., 2018. Global Sensitivity Analysis as Good Modelling Practices tool for the identification of the most influential process parameters of the primary drying step during freeze-drying. European Journal of Pharmaceutics and Biopharmaceutics 123, 108–116. URL: https://doi.org/10.1016/j.ejpb.2017.12.006, doi:10.1016/j.ejpb.2017.12.006.

Vogel, L., Peukert, W., 2005. From single particle impact behaviour to modelling of impact mills. Chemical Engineering Science 60, 5164–5176.

Wang, L.G., Morrissey, J.P., Sousani, M., Barrasso, D., Slade, D., Hanley, K., Ooi, J.Y., Litster, J.D., 2019. Model driven design in particulate products manufacturing, in: International Granulation Workshop, Lausanne, Switzerland.

Wang, Z., Escotet-Espinoza, M.S., Ierapetritou, M., 2017. Process analysis and optimization of continuous pharmaceutical manufacturing using flowsheet models. Computers and Chemical Engineering 107, 77–91. URL: https://doi.org/10.1016/j.compchemeng.2017.02.030, doi:10.1016/j.compchemeng.2017.02.030.

Wipf, D., Nagarajan, S., 2008. A new view of automatic relevance determination, in: Platt, J., Koller, D., Singer, Y., Roweis, S. (Eds.), Advances in Neural Information Processing Systems, Curran Associates, Inc.. pp. 1625–1632. URL: https://proceedings.neurips.cc/paper/2007/file/9c01802ddb981e6bcfbec0f0516b8e35-Paper.pdf.

Xie, X., Schenkendorf, R., Krewer, U., 2019. Efficient sensitivity analysis and interpretation of parameter correlations in chemical engineering. Reliability Engineering and System Safety 187, 159–173. URL: http://www.sciencedirect.com/science/article/pii/S0951832018300541, doi:https://doi.org/10.1016/j.ress.2018.06.010. sensitivity Analysis of Model Output.

Yeardley, A.S., Bugryniec, P.J., Milton, R.A., Brown, S.F., 2020a. A Study of the Thermal Runaway of Lithium-ion Batteries : A Gaussian Process

based Global Sensitivity Analysis. Journal of Power Sources doi:10.1016/j.jpowsour.2020.228001.

Yeardley, A.S., Roberts, D., Milton, R., Brown, S.F., 2020b. An Efficient Hybridization of Gaussian Processes and Clustering for Electricity Price Forecasting, in: 30th European Symposium on Computer Aided Process Engineering, Elsevier B.V.