Cambridge Centre for Computational Chemical Engineering

University of Cambridge

Department of Chemical Engineering

Preprint ISSN 1473 – 4273

Statistical approximation of the inverse problem in multivariate population balance modelling

Andreas Braumann [1] , Peter L. W. Man [1] , Markus Kraft [1]

released: 4 August 2009

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

Preprint No. 78

c4e

Key words and phrases: granulation, parameter estimation, uncertainties, multivariate population balance, inverse problem

Edited by
Cambridge Centre for Computational Chemical Engineering
Department of Chemical Engineering
University of Cambridge
Cambridge CB2 3RA
United Kingdom.

Fax: + 44 (0)1223 334796
E-Mail: c4e@cheng.cam.ac.uk
World Wide Web: http://www.cheng.cam.ac.uk/c4e/

Abstract

This paper deals with the estimation of model parameters and their uncertainties encountered in granulation modelling. A multivariate, detailed population balance model of a high shear granulation process is locally approximated by first and second order response surfaces, allowing a fast computation of the model response. The response surfaces are used in three different objective functions—moment matching, expected least squares and expected weighted least squares—in order to estimate ranges for the rate constants for particle coalescence, particle compaction, particle breakage, and reaction, which appear as free parameters in the granulation model. Firstly, second order response surfaces for the population balance model are constructed and used as approximation of the model in the objective functions for the numerical solution of the inverse problem. Secondly, the choice of objective function is investigated. It is found that the uncertainties of the model predictions differ for the three objective functions only slightly. The estimates for the intervals of the model parameters either overlap or are very close. However, the moment matching objective function is recommended because the number of estimated parameters and experimental data sets can be chosen independently.

Contents

1 Introduction

Granulation is an important and widely used process in many industries. In order to improve the process, e. g. produce a product which is cheaper and/or closer to the required specification, the process needs to be better understood, for instance, through additional experiments and/or modelling of the process. Detailed modelling of the granulation processes, such as by population balances, captures the complex structure of the granules and the various transformations in a granulator such as nucleation, coalescence and breakage, to name but a few, thereby aiming to bridge between the micro and macro scales. The rates with which these processes occur are normally not known and have to be estimated using experimental results.

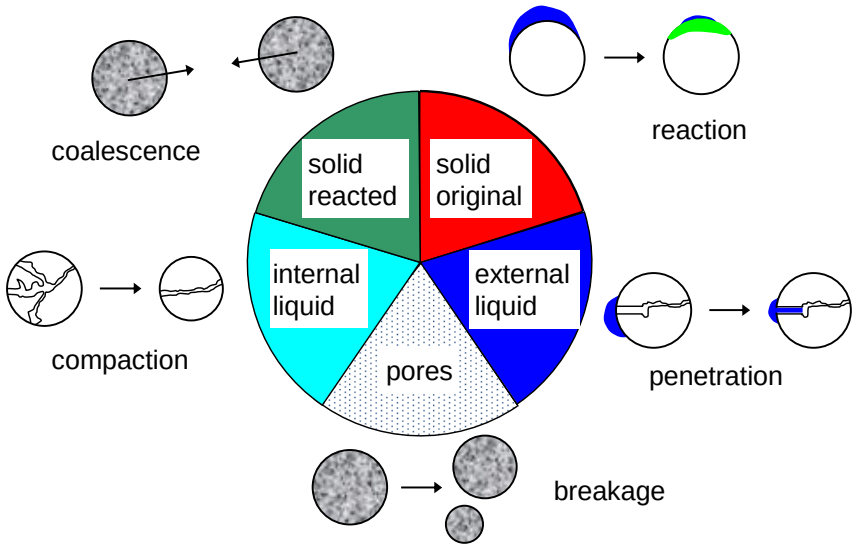
This so called "inverse problem" is encountered in virtually all modelling applications. Ramkrishna and Mahoney [18] emphasis the importance of the inverse problem in population balance modelling. Population balance models are not only used in granulation [1, 2, 5, 6, 8, 17], but also in other fields such as precipitation and crystallisation [9, 12, 20], combustion [7, 15, 16, 23], polymerisation [13, 26] and modelling of biological systems [11]. In terms of parameter estimation, rates, such as the growth rate in a precipitation [3] or the aggregation rate in a granulation process [19], are very often the centre of interest. This parameter identification can be performed in different ways. For instance, Vikhansky and Kraft [24] developed a stochastic algorithm that can be used to calculate the sensitivities of the parameters of a population balance for coagulation only processes. Based on this method, they solve the inverse problem for a liquid-liquid extractor [25].

Following a technique developed by Sheen *et al.* [21] for parameter estimation in combustion problems, Braumann and Kraft [4] study the inverse problem occurring in granulation described by a multidimensional population balance model, in which the particles are characterised by five independent variables, incorporating the transformations coalescence, compaction, breakage, penetration and reaction. In this approach linear response surfaces [4] are employed to avoid the time consuming evaluation of the five dimensional population balance model when minimising the objective function given by the inverse problem. The choice of this objective function is not unique and has not been studied in this context.

The purpose of the current paper is twofold. Firstly, the methodology proposed by Braumann and Kraft [4] is extended to second order response surfaces. Secondly, new objective functions will be introduced and compared with each other by using experimental data from Simmons *et al.* [22]. As a result of the optimisation procedure intervals of the rate constants for particle coalescence, particle compaction, particle breakage, and reaction will be determined.

2 Multivariate population balance model for granulation

In this work we model the wet granulation of small beads in a high shear mixer. A powder is agitated in a mixer and starts forming granules after being exposed to binder which can



transformation	rate
coalescence	$R_{\text{coal}} = \widehat{K}_0 \cdot f_{\text{coal}}(S, P, C)$
compaction	$R_{\text{comp}} = k_{\text{porred}} \cdot f_{\text{comp}}(S, P, C)$
breakage	$R_{\text{breakage}} = \widehat{k}_{\text{att}} \cdot f_{\text{breakage}}(S, P, C)$
penetration	$R_{\text{pen}} = \widehat{k}_{\text{pen}} \cdot f_{\text{pen}}(S, P)$
reaction	$R_{\text{reac}} = k_{\text{reac}} \cdot f_{\text{reac}}(S, P)$

S	particle states
P	material properties
C	process conditions

Figure 1: *Particle model, transformations and rates in the multivariate granulation model.*

be added either as liquid or as solid flakes. The powder transforms into nearly spherical particles which are held together by binder between the beads and exhibit also binder on the outer surface. Chemical reactions within the granules lead to the formation of an additional component. In addition, attrition of the granules caused by the mixing process can be observed.

In this study, a population balance model for granulation, which has been described in detail in [6], shall be used. The model, which describes the dynamics in a batch mixer, is outlined in terms of the type space of the population members, the transformations which describe how the population changes, and the rates with which this happens. These key elements of the model are shown in **Figure 1**. The population consists of granules, which in turn have five internal coordinates: solid original, solid reacted, liquid external, liquid internal and pores. The particles in the granulator are subject to various transformations. These are particle coalescence, compaction, breakage, liquid penetration and reaction within a particle. Each of these transformations takes place with a rate that depends on the states of the particles, material properties, process conditions and rate constants. For instance, the rate for particle coalescence is proportional to the collision rate constant $\widehat{K}_0$. More precisely, this rate constant is part of the coalescence kernel K, being composed of,

$$K = n_{\text{impeller}}\, \widehat{K}_0\, \widetilde{K}\,, \tag{1}$$

where n_{impeller} is the rotational speed of the impeller and $\widetilde{K}$ the collision efficiency. The latter is determined using the Stokes criterion [10]. Furthermore, particles undergo compaction, whereby the porosity of the particles is reduced. The rate of the compaction step is determined by the compaction rate constant k_{porred}, whereby the model for this transformation follows the findings of [14]. Hence, the reduction of the porosity of a single particle $\Delta\varepsilon$ equates as,

$$\Delta\varepsilon = k_{\text{porred}}\, f(\varepsilon, n_{\text{impeller}}, \varepsilon_{\text{min}})\,, \tag{2}$$

with ε being the porosity of the particle and ε_{min} the minimum porosity. Breakage as the antagonistic transformation to particle coalescence is also considered in the model framework. The rate with which the process happens is determined by the breakage frequency g. As such, the breakage of a granule is caused by impacts on the particles. The energy for this is thought to be introduced by the rotary movement of the impeller (a chopper is not considered in the current model). Hence, the breakage frequency is proportional to the introduced kinetic energy, and is modelled by,

$$g = \widehat{k}_{\text{att}}\, U_{\text{imp}}^2\, (\varepsilon\, \Psi(s_{\text{r}}) + \chi)\, v\,, \tag{3}$$

where $\widehat{k}_{\text{att}}$ is the rate constant for breakage, U_{imp} the impact velocity of the particles, Ψ a function of the amount of reacted solid s_{r} accounting for strengthening of the particle due to reaction, χ the volume fraction of the external liquid and v the particle volume. Within a granule, two kind of liquids are distinguished—external on the outer surface and internal inside the pores. Due to capillary effects liquid will migrate into empty pores, which is called penetration. The rate with which the liquid migrates within a particle is proportional to the penetration rate constant $\widehat{k}_{\text{pen}}$, leading to,

$$r_{\text{pen}} = \widehat{k}_{\text{pen}}\, \eta^{-1/2}\, l_{\text{e}}\, (p - l_{\text{i}})\,, \tag{4}$$

with η as the viscosity of the liquids, l_{e} and l_{i} as the amount of external and internal liquid, and p as the pore volume. The combination of different materials in a single granule can result in reactions within the granules, such as hydration or neutralisation reactions, resulting in the formation of reacted solid s_{r}. The rate for the reaction transformation is proportional to the reaction rate constant k_{reac}. With the reaction on the outer surface being depended on the amount of external liquid l_{e}, the rate law is given by,

$$r_{\text{reac}} = k_{\text{reac}}\, a_{\text{e}}\, \frac{l_{\text{e}}}{l_{\text{e}} + s_{\text{r}}}\,, \tag{5}$$

where a_{e} is the external surface area of the particle. An identical law is applied for the reaction inside the pores, where the amount of internal liquid l_{i} drives the reaction.

3 Model approximation by second order response surfaces

The multivariate population balance model for granulation that is used in the current study is quite complex. This means, the model contains some 20 parameters, some of them are

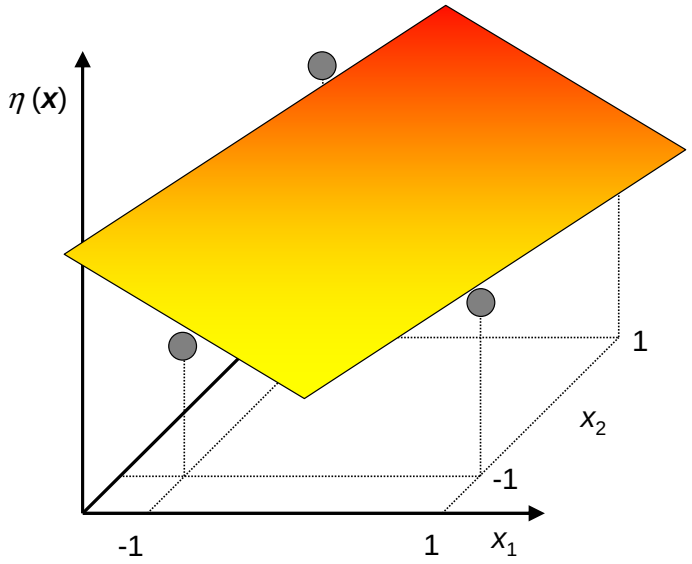


Figure 2: *An experimental design with two coded variables x_1 and x_2 and a linear response surface. The circles represent the evaluations of the full model at the corner points.*

often unknown such as the rate constants. These unknown model parameters can be found by solving the inverse problem, i. e., comparing model predictions with experimental results. In order to perform this task, the model has to be evaluated several times. Since the numerical solving of the complex granulation model is computationally expensive, the full model shall locally be approximated by second order response surfaces. As result, the model evaluations for solving the inverse problem can be obtained very quickly.

The response surfaces are dependent on the model parameters, so is the full granulation model. These variables of the responses surfaces shall be denotes by

$$\boldsymbol{x} = (x_1, \ldots, x_K)\,. \tag{6}$$

Often this set of parameters is a subset of the full parameter set of a complex model. In order to construct the response surfaces, $\boldsymbol{x}$ are usually used in the dimensionless form, and hence they are called coded variables. The uncoded counterpart of such coded variable is the respective non-dimensionless parameter in the full granulation model. A second order response surface as a function of $\boldsymbol{x}$ takes following form,

$$\eta(\boldsymbol{x}) = \beta_0 + \sum_{k=1}^{K} \beta_k\, x_k + \sum_{k=1}^{K} \sum_{l \geq k}^{K} \beta_{kl}\, x_k\, x_l\,. \tag{7}$$

For the establishment of the coefficients β_0, β_k and β_{kl}, a so-called experimental design needs to be set up. This means, the full granulation model needs to be evaluated at different combinations of the parameters $\boldsymbol{x}$. As an example, the construction of the response surface as a function of two variables ($K = 2$), x_1 and x_2, is visualised in **Figure 2**. At first, the limits for the coded variables have to be chosen. It is common to choose an experimental design such that the coded variables become

$$x_k \in [-1, 1]\,. \tag{8}$$

Table 1: *Unknown parameters in the non-dimensionless and the coded form.*

transformation	rate constant	coded variable
coalescence	$\widehat{K}_0$	x_1
compaction	k_{porred}	x_2
breakage	$\widehat{k}_{\text{att}}$	x_3
reaction	k_{reac}	x_4

In general, the coded variables span a K-dimensional hypercube with 2^K corner points. For the corner points of the chosen design with $K = 2$, i. e., for the different parameter combinations of x_1 and x_2, the full model is evaluated. This set of model responses $f_{\text{sim,full}}$ can then be used in a fitting procedure in order to obtain the parameters of the response surface $\eta(\boldsymbol{x})$,

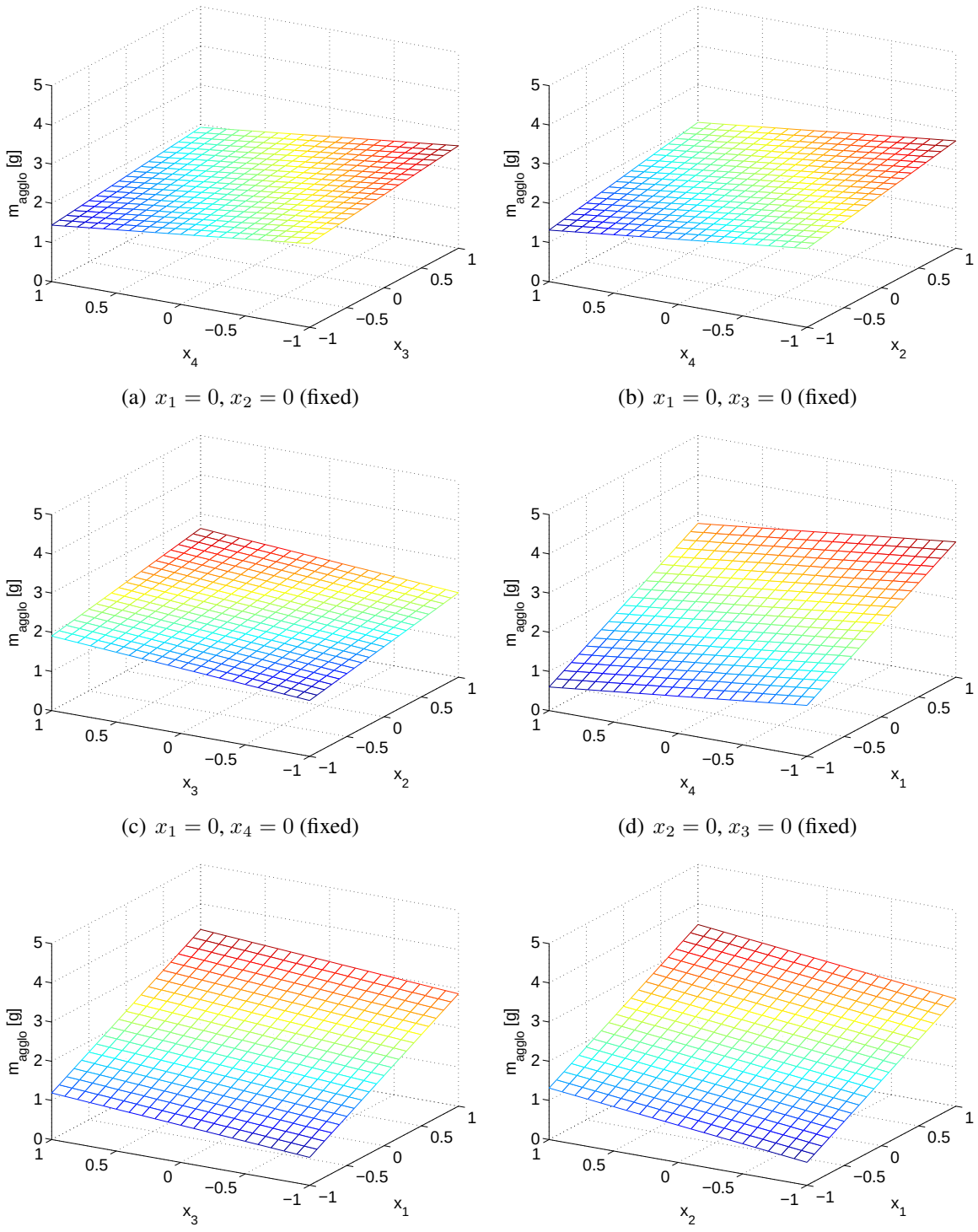
$$\eta(\boldsymbol{x}) = f_{\text{sim,full}}(\boldsymbol{x}) + \varepsilon(\boldsymbol{x})\,, \tag{9}$$

where ε is the approximation error. In the case of a linear response surface with two variables, as depicted in Figure 2, three parameters determine the approximation of the full model. While the parameter β_0 gives the value of the response surface at (0,0), the two coefficients β_1 and β_2 provide the slopes of the surface. In fact, they are the sensitivities of the approximated model with respect to the parameters x_1 and x_2.

In the current study, the response surfaces shall be used as approximation of the multidimensional population balance model. Four parameters of this model shall be determined by comparison of the model predictions with experimental results, i. e., solve the inverse problem. The unknown parameters are the rate constants for particle coalescence, particle compaction, particle breakage and reaction, which are converted into coded variables in order to set up the response surfaces. **Table 1** lists the connection between the coded variables and the unknown parameters. In a previous study [6], the model has been applied to the granulation of nonpareils with aqueous polyethylene glycol (PEG) 4000 solutions in a bench scale mixer. In order to construct the response surfaces, a lower and upper boundary for each of the four unknown parameters were chosen and linked to the coded variables. Subsequently, evaluations of the full model for a process time of 80 s, with an impeller speed of 900 rpm and a binder with a water/PEG4000 ratio of 50:50, are carried out for the corner points of the experimental design, i. e., for different parameter combinations, so that a linear response surface for this case can be set up. Projections of this surface are shown in **Figure 3**, with two of the four coded variables being fixed at zero. In order to construct second order response surfaces, additional points other than the corner points need to be evaluated, so that the curvature of the response surface can be determined. In the current study these four points are

$$\begin{aligned} \boldsymbol{x}_1 &= (0.5, 0, 0, 0) \quad , \quad \boldsymbol{x}_2 = (0, 0.5, 0, 0) \\ \boldsymbol{x}_3 &= (0, 0, 0.5, 0) \quad , \quad \boldsymbol{x}_4 = (0, 0, 0, 0.5)\,. \end{aligned} \tag{10}$$

This means, for the same design and process conditions another response surface is obtained, this time of second order (**Figure 4**). Compared to the linear response surface the second order response surface possesses some curvature.



(a) $x_1 = 0$, $x_2 = 0$ (fixed)

(b) $x_1 = 0$, $x_3 = 0$ (fixed)

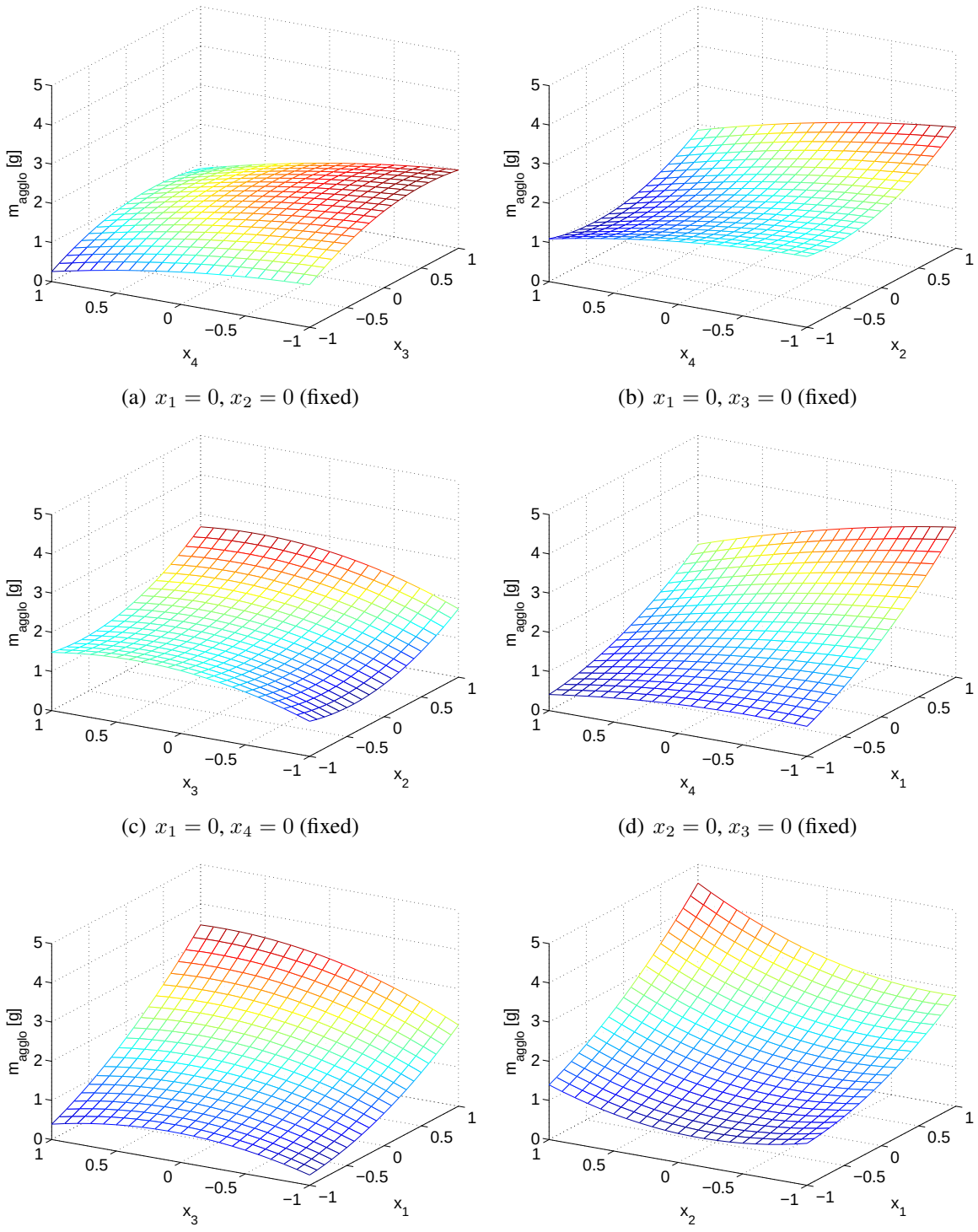
(c) $x_1 = 0$, $x_4 = 0$ (fixed)

(d) $x_2 = 0$, $x_3 = 0$ (fixed)

(e) $x_2 = 0$, $x_4 = 0$ (fixed)

(f) $x_3 = 0$, $x_4 = 0$ (fixed)

Figure 3: *Linear response surfaces with two fixed variables for 900 rpm and a water/PEG4000 ratio of 50:50 after 80 s.*



(a) $x_1 = 0$, $x_2 = 0$ (fixed)

(b) $x_1 = 0$, $x_3 = 0$ (fixed)

(c) $x_1 = 0$, $x_4 = 0$ (fixed)

(d) $x_2 = 0$, $x_3 = 0$ (fixed)

(e) $x_2 = 0$, $x_4 = 0$ (fixed)

(f) $x_3 = 0$, $x_4 = 0$ (fixed)

Figure 4: *Second order response surfaces with two fixed variables for 900 rpm and a water/PEG4000 ratio of 50:50 after 80 s.*

4 Parameter estimation and error propagation

4.1 Theory

In order to complete the model we need to estimate the four unknown rate constants introduced above. This task is often done by fitting the model to a set of experimental data. The fitting procedure is carried out as an optimisation which requires many model evaluations and hence can lead to prohibitive computational demand. Because of this, the model is approximated by the response surfaces. Furthermore, experimental observations have uncertainties or errors associated with them. This means, data can be used to determine the range the unknown parameters lie in. Hence, an unknown parameter, x, will have some mean, x_0, and a standard deviation, c. If it is further assumed that the parameter is Gaussian distributed, x can be expressed by

$$x = x_0 + c\,\xi\,, \tag{11}$$

where ξ is standard normally distributed. The uncertainty in the parameters $\boldsymbol{x}$ carries over into the model, in the current case the response surfaces. Hence eq. (7) turns into,

$$\begin{aligned}\eta\left(\boldsymbol{x_0}, \boldsymbol{c}, \boldsymbol{\xi}\right) = & \underbrace{\beta_0 + \sum_{k=1}^{K} \beta_k\, x_{0,k} + \sum_{k=1}^{K}\sum_{l\geq k}^{K} \beta_{kl}\, x_{0,k}\, x_{0,l}}_{\eta\left(\boldsymbol{x_0}\right)} \\ & + \sum_{k=1}^{K} \beta_k\, c_k\, \xi_k + \sum_{k=1}^{K}\sum_{l\geq k}^{K} \beta_{kl}\, c_k\, c_l\, \xi_k\, \xi_l \\ & + \sum_{k=1}^{K}\sum_{l\geq k}^{K} \beta_{kl}\left(x_{0,k}\, c_l\, \xi_l + x_{0,l}\, c_k\, \xi_k\right)\,.\end{aligned} \tag{12}$$

It is apparent that η can take a range of values for given $\boldsymbol{x_0}$ due to the randomness introduced by $\boldsymbol{\xi}$. However, the mean and the variance of eq. (12) as characteristic measures can be computed. The mean equates as,

$$\mu(\boldsymbol{x_0}, \boldsymbol{c}) = \mathbb{E}\left(\eta(\boldsymbol{x_0}, \boldsymbol{c}, \boldsymbol{\xi})\right) = \eta(\boldsymbol{x_0}) + \sum_{k=1}^{K} \beta_{kk}\, c_k^2\,, \tag{13}$$

whereas the variance $\sigma^2(\boldsymbol{x_0}, \boldsymbol{c})$ can be written as,

$$\begin{aligned}
\sigma^2(\boldsymbol{x_0}, \boldsymbol{c}) &= \mathrm{Var}\left(\eta\left(\boldsymbol{x_0}, \boldsymbol{c}, \boldsymbol{\xi}\right)\right) \\
&= \sum_{k=1}^{K} \beta_k^2\, c_k^2 + \sum_{k=1}^{K}\sum_{l>k}^{K} \beta_{kl}^2\, c_k^2\, c_l^2 + 2\sum_{k=1}^{K}(\beta_{kk}\, c_k^2)^2 \\
&\quad + \sum_{k=1}^{K}\sum_{l>k}^{K} \beta_{kl}^2 \left[(x_{0,k}\, c_l)^2 + (x_{0,l}\, c_k)^2\right] \\
&\quad + \sum_{k=1}^{K}\sum_{\substack{s=1\\ s\neq k}}^{K}\sum_{\substack{l>k\\ l>s}}^{K} \beta_{kl}\, \beta_{sl}\, x_{0,k}\, c_l^2\, x_{0,s} + \sum_{s=1}^{K}\sum_{k>s}^{K}\sum_{l>k}^{K} \beta_{kl}\, \beta_{sk}\, x_{0,l}\, c_k^2\, x_{0,s} \\
&\quad + \sum_{k=1}^{K}\sum_{l>k}^{K}\sum_{t>l}^{K} \beta_{kl}\, \beta_{lt}\, x_{0,k}\, c_l^2\, x_{0,t} + \sum_{k=1}^{K}\sum_{l>k}^{K}\sum_{\substack{t>k\\ t\neq l}}^{K} \beta_{kl}\, \beta_{kt}\, x_{0,l}\, c_k^2\, x_{0,t} \\
&\quad + \sum_{k=1}^{K}(2\,\beta_{kk}\, x_{0,k}\, c_k)^2 + 2\sum_{k=1}^{K}\sum_{l>k}^{K} \beta_{kl}\left(\beta_l\, x_{0,k}\, c_l^2 + \beta_k\, x_{0,l}\, c_k^2\right) \\
&\quad + 4\sum_{k=1}^{K} \beta_k\, \beta_{kk}\, x_{0,k}\, c_k^2 + 4\sum_{k=1}^{K}\sum_{l>k}^{K} x_{0,k}\, x_{0,l}\, \beta_{kl}\left(c_k^2\, \beta_{kk} + c_l^2 \beta_{ll}\right)\,,
\end{aligned} \tag{14}$$

Further details about the derivation of the variance of $\eta(\boldsymbol{x_0}, \boldsymbol{c}, \boldsymbol{\xi})$ can be found in appendix A.

The ultimate aim of this study is the estimation of the unknown parameters and their uncertainties using experimental results. These results usually consist of a mean value η_0^{exp} and the experimental uncertainty σ^{exp}, so that the experimental result are reported as,

$$\eta^{\mathrm{exp}} = \eta_0^{\mathrm{exp}} \pm \sigma^{\mathrm{exp}}\,. \tag{15}$$

The values of the unknown model parameters shall be chosen in such a way as to maximise the agreement between the experimental results and the corresponding model predictions. Hence one possible objective function, Φ, can be defined as

$$\Phi(\boldsymbol{x_0}, \boldsymbol{c}) = \sum_{i=1}^{N}\left(\left[\eta_{0,i}^{\mathrm{exp}} - \mu_i(\boldsymbol{x_0}, \boldsymbol{c})\right]^2 + \left[\sigma_i^{\mathrm{exp}} - \sigma_i(\boldsymbol{x_0}, \boldsymbol{c})\right]^2\right)\,, \tag{16}$$

where N is the number of experimental sets under consideration. We shall call eq. (16) the *moment matching objective function*. The best matching set of parameters $(\boldsymbol{x}^*, \boldsymbol{c}^*)$ is obtained by employing a suitable method, i. e., an appropriate optimisation routine, so that

$$(\boldsymbol{x_0^*}, \boldsymbol{c}^*) = \underset{\boldsymbol{x_0},\, \boldsymbol{c}}{\mathrm{argmin}}\left\{\Phi(\boldsymbol{x_0}, \boldsymbol{c})\right\}\,. \tag{17}$$

4.2 Results

The outlined methodology is applied to a granulation process in a batch operated mixer. Simmons *et al.* [22] performed such a process in a bench scale mixer using nonpareils as

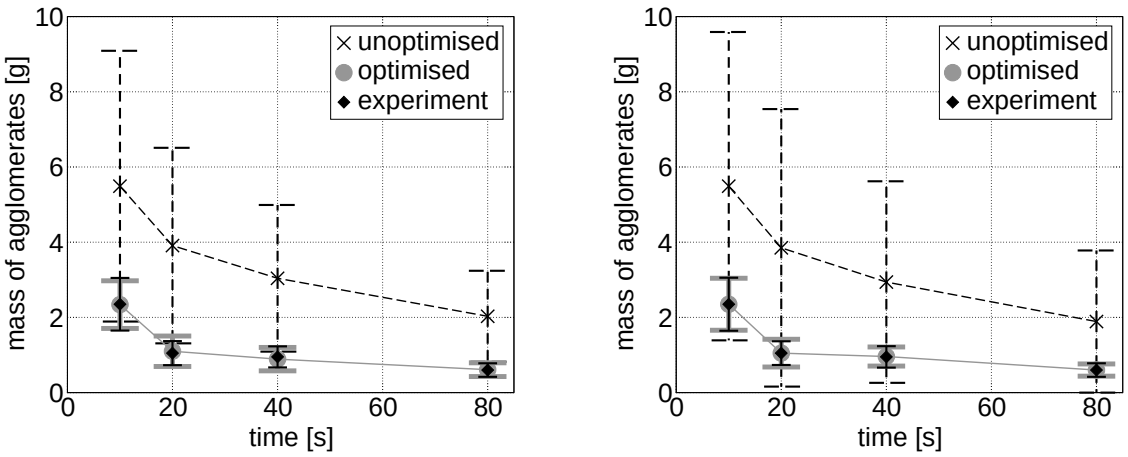
solid material and aqueous polyethylene glycol (PEG) 4000 solutions as binder. Three different compositions of binder were used, i. e., with 10, 25 and 50 % PEG4000 in the solution. The mixer was operated at three different speeds, namely at 600, 900 and 1200 rpm, with samples being taken at four different times when the binder with 50 % PEG4000 was used. The samples were screened with respect to the mass of agglomerates, i. e., particles that exceeded a certain size. Using the multivariate population balance model introduced above, it is possible to study the temporal evolution of a particle ensemble in a granulator. However, in order to do so, several parameters must be known/estimated. Hence, the current methodology shall be used in order to determine four unknown rate constants, namely for particle collision, $\widehat{K}_0$, compaction, k_{porred}, breakage, $\widehat{k}_{\mathrm{att}}$, and reaction, k_{reac}, employing the experimental results from Simmons *et al.* [22]. More specifically, the mass of agglomerates obtained from the experiments with a binder composition of PEG4000/water of 50:50 are used, because samples at four different times were taken for these cases.

In a first step, the experimental data from the setup with the medium impeller speed of 900 rpm and a binder composition of PEG4000/water of 50:50 shall be used. Further on, these data will be referred to as sets 1–4. In order to estimate the unknown parameters, the response surfaces for the process conditions (and times) of the respective experiments need to be constructed. As described earlier, the complex granulation model is therefore evaluated at specific points of the design, allowing the construction of linear and second order response surfaces. These local approximations of the full granulation model are then used in the following optimisation step, i. e., the estimation of the unknown parameters, in order to predict the mass of agglomerates. Upon choosing the initial guesses for $\boldsymbol{x_0}$ and $\boldsymbol{c}$, the model predictions, namely the mean and the variance, for the unoptimised cases can be computed. As initial value we choose

$$\boldsymbol{x_0}^{(0)} = \mathbf{0} \quad , \quad \boldsymbol{c}^{(0)} = \mathbf{1} \, , \tag{18}$$

leading to the unoptimised results shown in **Figure 5(a)**, when the linear response surfaces are used for the computation of the model responses. After bringing in the four experimental sets with an impeller speed of 900 rpm and performing the parameter estimation, the model predictions match with the experimental results very well. This means, there is a set of model parameters that allows the model to take very similar values as the experimental ones. Furthermore, the uncertainties of the model predictions are reduced significantly. As such, the proposed methodology shows very similar results for the mass of agglomerates when second order response surfaces are chosen to provide the model responses (**Figure 5(b)**), i. e., the *a priori* uncertainties in the model predictions are rather large at the beginning, but can be reduced subsequently.

Despite these very similar model predictions when using linear and second order response surfaces, i. e., different model approximations, it is worthwhile looking at the values of the unknown parameters. These parameters of the complex population balance model are obtained after retransformation of the solution $(\boldsymbol{x_0^*}, \boldsymbol{c^*})$ and listed in **Table 2**. The parameter values obtained from using linear and second order response surfaces do not differ massively from each other. The mean deviation being between 6 and 14 %. With respect to the uncertainties of the single parameters it is interesting to note that they increase as well as decrease when second order response surfaces are used instead of linear response surfaces. However, whilst the relative increase in uncertainty for the coalescence rate



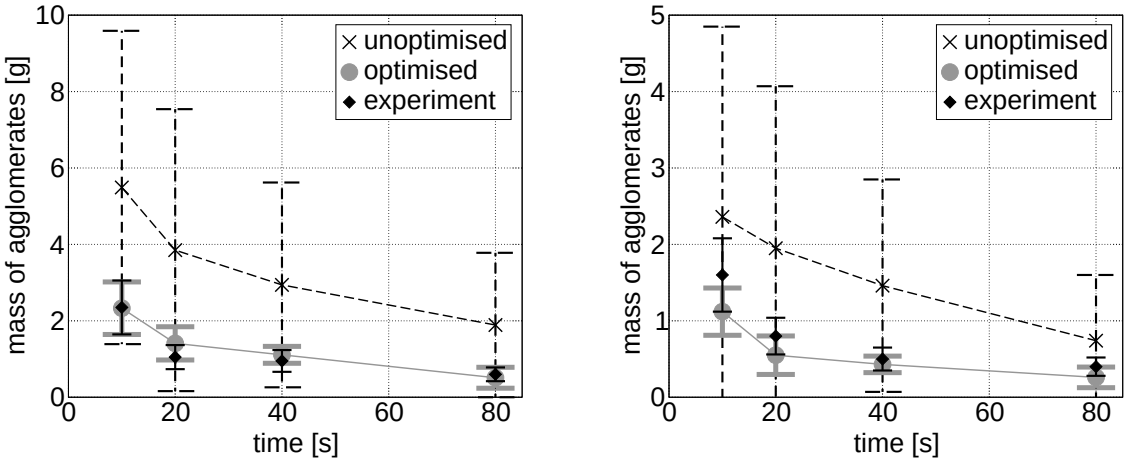
(a) Linear response surfaces as model

(b) 2nd order response surface as model

Figure 5: *Mass of agglomerates in the unoptimised and optimised model, and experiment at impeller speed of 900 rpm using sets 1–4 in moment matching objective function.*

Table 2: *Rate constants of granulation model with their uncertainties using sets 1–4 and moment matching objective function.*

response surface	$\widehat{K}_0 \cdot 10^{10}$ $[\mathrm{m}^3]$	k_{porred} $[\mathrm{s\,m}^{-1}]$	$\widehat{k}_{\mathrm{att}} \cdot 10^{-7}$ $[\mathrm{s\,m}^{-5}]$	$k_{\mathrm{reac}} \cdot 10^{9}$ $[\mathrm{m\,s}^{-1}]$
1st order	1.18 ± 0.00	0.20 ± 0.04	4.8 ± 1.0	3.6 ± 0.0
2nd order	1.04 ± 0.07	0.2136 ± 0.0005	4.2 ± 0.4	3.28 ± 0.01



(a) 900 rpm (b) 1200 rpm

Figure 6: *Mass of agglomerates in the unoptimised and optimised model, and experiment for setups with different impeller speed using second order response surfaces of sets 1–8 and the moment matching objective function.*

constant $\widehat{K}_0$ and the reaction constant k_{reac} is moderate when we switch to second order response surfaces, the relative uncertainties for the compaction rate constant k_{porred} and the breakage rate constant $\widehat{K}_{\mathrm{att}}$ are reduced in both cases from 20 % to less than 10 % or even lower. This suggests that the second order response surfaces are better suited as approximation to the response of the complex granulation model, resulting in smaller relative uncertainties for the unknown parameters.

So far, only the estimation of the unknown model parameters for the same process conditions, i. e., binder composition and impeller speed, but for different process times has been considered. However, the full granulation model is structured in such a way that changes in impeller speed and binder composition are dealt within the process model rather than in the rate constants. This means, for different impeller speeds and binder compositions, there should be one single set of parameters that fits all conditions. In order to study the effect of different process conditions in the experiments (and in the model), second order response surfaces for setups with an impeller speed of 1200 rpm and a binder composition for PEG4000/water of 50:50 (as before) were constructed in the aforementioned manner. Altogether eight different experimental sets with impeller speeds of 900 (sets 1–4) and 1200 rpm (sets 5–8) are used then in order to estimate the four unknown parameters and their uncertainties. The mass of agglomerates for these cases is shown in **Figure 6**. As before the model predictions along with their uncertainties for the unoptimised cases can be plotted. The uncertainties for every single set reduces significantly in the optimised state, although not necessarily to the experimental uncertainty of that particular set. The big difference to previous considerations is that the model predictions by and large do not coincide with the experimental value. For the 900 rpm setup (**Figure 6(a)**) the model predictions of the optimised cases lie within the experimental uncertainty range of these sets, so that the model can be judged as being appropriate for this setup. For the 1200 rpm setup the situation seems to be not so clear. **Figure 6(b)** suggests that the model predictions are

Table 3: *Rate constants of granulation model with their uncertainties using second order response surfaces and moment matching objective function.*

sets used	$\widehat{K}_0 \cdot 10^{10}$ $[\mathrm{m}^3]$	k_{porred} $[\mathrm{s\,m}^{-1}]$	$\widehat{k}_{\mathrm{att}} \cdot 10^{-7}$ $[\mathrm{s\,m}^{-5}]$	$k_{\mathrm{reac}} \cdot 10^{9}$ $[\mathrm{m\,s}^{-1}]$
1–4	1.04 ± 0.07	0.2136 ± 0.0005	4.2 ± 0.4	3.28 ± 0.01
1–8	1.0 ± 0.0	0.36 ± 0.04	7.7 ± 0.4	3.5 ± 0.0

just within the experimental confidence bands for the 10, 20 and 80 s case. This means, these cases should be revisited by the scientist in two ways. On the one hand it seems that the model deserves slight improvement in order to better capture the conditions of the considered setup, but on the other hand it is also advisable to check the experimental result, maybe even confirming the result by a repetition of the experiment where possible.

The parameter values being obtained from using four or eight sets in the routine respectively, can differ significantly (**Table 3**). Whilst the collision rate constant $\widehat{K}_0$ hardly changes when the 1200 rpm cases (sets 5–8) are brought into the methodology, the compaction rate constant k_{porred} and the breakage rate constant k_{att} increase significantly. The less than 10 % increase in the reaction rate constant is rather moderate. The uncertainties of the parameter estimates do not change significantly when eight instead of four sets are used for their estimation.

5 Different objective functions

The parameter estimations so far were based on using the *moment matching objective function*. In this, the aim is to match the model predictions with the experimental results for each experiment/scenario. However, since the experimental results as well as the model predictions carry uncertainties, a new random variable expressed as the squared difference of the experimental results and the model prediction can be defined,

$$\left[\eta^{\mathrm{exp}}(\xi) - \eta(x_0, c, \xi)\right]^2 \,, \tag{19}$$

where ξ is standard normally distributed. The expression in eq. (19) is just a function of the single model parameter $x = x_0 + c\,\xi$, but obviously a model response with a number of parameters, $\boldsymbol{x}$, should be considered. Furthermore, the expression in eq. (19) is a random variable, so that the expectation of it should be taken. Hence, an objective function Φ can be written as,

$$\Phi(\boldsymbol{x_0}, \boldsymbol{c}, \boldsymbol{\xi}) = \mathbb{E}\left\{\sum_{i=1}^{N} \left[\eta_i^{\mathrm{exp}}(\boldsymbol{\xi}) - \eta_i(\boldsymbol{x_0}, \boldsymbol{c}, \boldsymbol{\xi})\right]^2\right\} \,, \tag{20}$$

with N being the number of experiments. Whilst the model response η_i is a function of the set of random variables $\boldsymbol{\xi}$, with the size of $\boldsymbol{\xi}$ being K, the experimental results are just functions of the single random variable ξ_i. This means, the number of experiments N used in the objective function (eq. (20)) cannot be bigger than the number of uncertain variables $\boldsymbol{x}$, i. e., $N = K$. After a few steps the following expression for the objective

function is obtained,

$$\Phi(\boldsymbol{x_0}, \boldsymbol{c}) = \sum_{i=1}^{N} \left[\eta_{0,i}^{\text{exp}} - \mu_i(\boldsymbol{x_0}, \boldsymbol{c})\right]^2 + \sum_{i=1}^{N} (\sigma_i^{\text{exp}})^2 + \sigma_i^2(\boldsymbol{x_0}, \boldsymbol{c}) \\ - \sum_{i=1}^{N} 2\,\sigma_i^{\text{exp}} \left(\beta_{i,i}\, c_i + \sum_{k=1}^{i} \beta_{ki,i}\, x_{0,k}\, c_i + \sum_{l \geq i}^{K} \beta_{il,i}\, x_{0,l}\, c_i\right) . \tag{21}$$

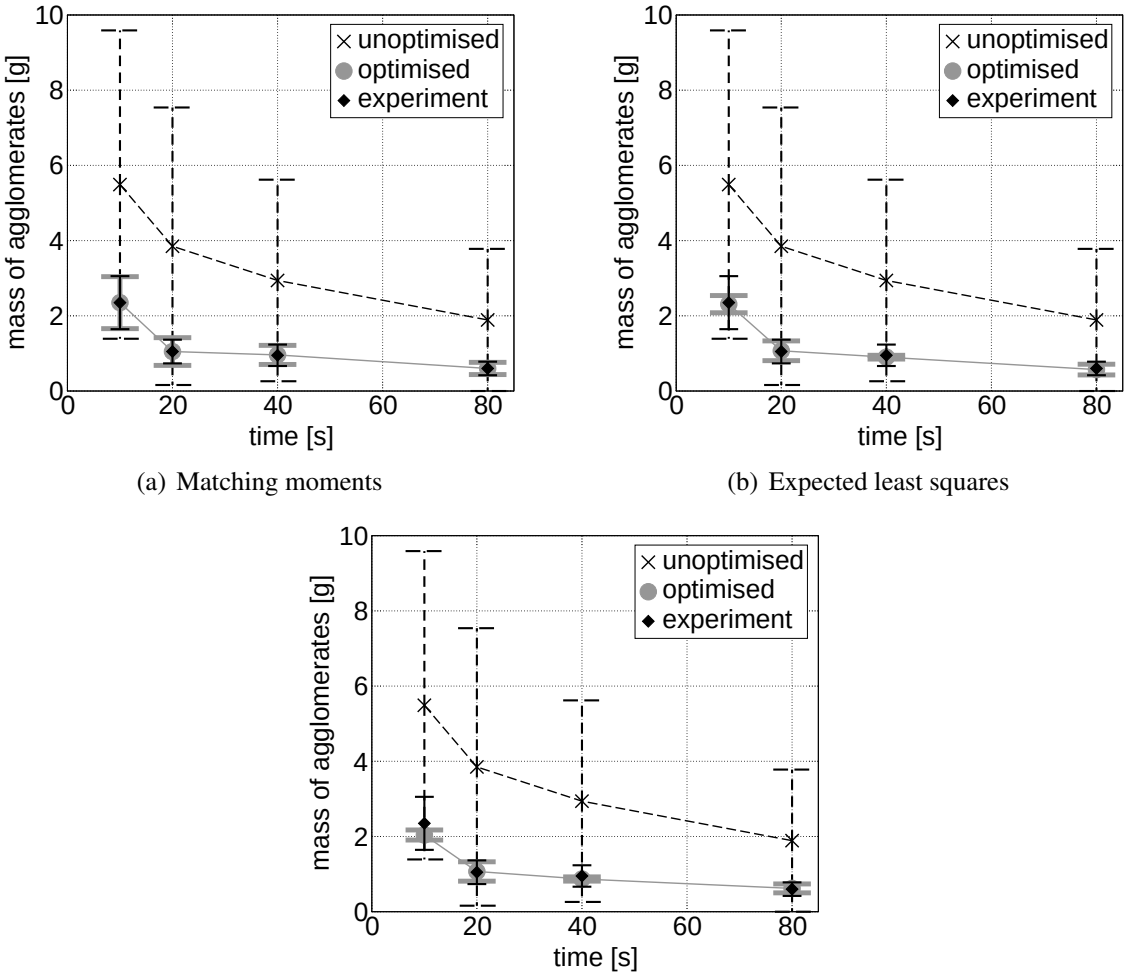
This objective function shall be called *expected least squares objective function*. Eq. (21) is now applied in the methodology in order to estimate the four unknown parameters. As experimental results, sets 1–4 (900 rpm) are used. The resulting model predictions and their associated uncertainties are plotted in **Figure 7(b)**. As for the moment matching objective function the model predictions are very close to the experimental results, although it has to be said that a perfect match cannot be quite established for $t = 40\,\text{s}$. Moreover the *a priori* uncertainties in the model predictions get significantly reduced. In contrast to the result from the moment matching objective function (**Figure 7(a)**) the uncertainty in the model predictions using the expected least squares objective function can become smaller than the experimental uncertainties. This is perfectly possible due to the structure of the objective function, being the expectation over the sum of the differences between experimental results and (uncertain) model predictions. In fact, smaller uncertainty bounds are better when it comes to model judgement, because the model can be trusted more. The derived values of the model parameters of interest are listed in **Table 4**. When the expected least squares instead of the moment matching objective function is used for the parameter estimation, the parameter estimates do not differ significantly from each other, with the biggest relative change being just 4 %. However, with respect to the parameter uncertainties, a more significant change can be observed. This means, while the uncertainty in the breakage rate constant vanishes, the uncertainty in the reaction rate constant increases substantially.

When using the experimental results in the parameter estimation procedure, one may want to give different weights to different experimental results. For instance, one might not be so certain about a particular result, and hence would like to give it less importance in the procedure. This means, the squared difference between experiment and model should be weighted. In principle any kind of weight can be chosen, but for obvious reasons one may want to chose the inverse value of the experimental uncertainty as weight, so that the objective function takes following form,

$$\Phi\left(\boldsymbol{x_0}, \boldsymbol{c}, \boldsymbol{\xi}\right) = \mathbb{E}\left\{\sum_{i=1}^{N} \left[\frac{\eta_i^{\text{exp}}(\boldsymbol{\xi}) - \eta_i(\boldsymbol{x_0}, \boldsymbol{c}, \boldsymbol{\xi})}{\sigma_i^{\text{exp}}}\right]^2\right\} . \tag{22}$$

This objective function shall be called *expected weighted least squares objective function*. After a few steps eq. (22) turns into,

$$\Phi(\boldsymbol{x_0}, \boldsymbol{c}) = \sum_{i=1}^{N} \left[\frac{\eta_{0,i}^{\text{exp}} - \mu_i(\boldsymbol{x_0}, \boldsymbol{c})}{\sigma_i^{\text{exp}}}\right]^2 + N + \sum_{i=1}^{N} \left[\frac{\sigma_i(\boldsymbol{x_0}, \boldsymbol{c})}{\sigma_i^{\text{exp}}}\right]^2 \\ - \sum_{i=1}^{N} \frac{2}{\sigma_i^{\text{exp}}} \left(\beta_{i,i}\, c_i + \sum_{k=1}^{i} \beta_{ki,i}\, x_{0,k}\, c_i + \sum_{l \geq i}^{K} \beta_{il,i}\, x_{0,l}\, c_i\right) . \tag{23}$$



(a) Matching moments

(b) Expected least squares

(c) Expected weighted least squares

Figure 7: *Mass of agglomerates in the unoptimised and optimised model, and experiment using second order response surfaces and different objective functions (impeller speed of 900 rpm).*

The application of this objective function in the methodology leads to the predicted mass of agglomerates as shown in **Figure 7(c)**. The difference between the results from the expected least squares and the expected weighted least squares objective function is fairly small, except for the prediction at $t = 10\,\mathrm{s}$, where the model prediction does not match exactly with the experimental mass, but is still within the uncertainty bounds. As can be expected from these findings, the estimates for the unknown parameters are not very different from the estimates found in previous arrangements, except for the reaction rate constant (Table 4). Using the expected weighted least squares objective function means that the reaction constant is approximately 20 % smaller than for the other two objective functions. However, the uncertainties for the reaction constant are rather large, so that the ranges for all three parameter estimates overlap. If a greater confidence about a parameter estimate is required, more experimental data should be included into the procedure. This means, the moment matching objective function should be preferred over the other two

Table 4: *Rate constants of granulation model with their uncertainties derived from sets 1–4 using different objective functions.*

case	$\widehat{K}_0 \cdot 10^{10}$ $[\mathrm{m}^3]$	k_{porred} $[\mathrm{s\,m}^{-1}]$	$\widehat{k}_{\mathrm{att}} \cdot 10^{-7}$ $[\mathrm{s\,m}^{-5}]$	$k_{\mathrm{reac}} \cdot 10^{9}$ $[\mathrm{m\,s}^{-1}]$
moment matching	1.04 ± 0.07	0.2136 ± 0.0005	4.2 ± 0.4	3.28 ± 0.01
expected least squares	1.05 ± 0.04	0.2052 ± 0.0	4.1 ± 0.0	3.2 ± 0.8
exp. weighted least squares	1.01 ± 0.0	0.2 ± 0.0	4.10 ± 0.09	2.6 ± 0.6

objective functions, because the number of estimates parameters can be chosen independently from the number of experiments.

6 Conclusions

In this paper we discussed the estimation of parameters and their uncertainties for granulation processes, being modelled by multivariate population balances. A response surface methodology for the local approximation of the complex granulation model for the subsequently applied methodology is introduced. The mean and variance for model approximation of up to second order are given and used in a moment matching objective function, extending previous work with linear response surfaces. It was shown that the second order response surfaces are better suited for the model approximation than the first order response surface, leading to smaller relative uncertainties in the uncertainties of the four unknown model parameters. Furthermore, two other objective functions for the parameter estimation are studied. This so called expected least squares and expected weighted least squares objective functions yield estimates for the unknown model parameters that are not too different from the estimates using the moment matching objective function. However, the uncertainties in the model predictions are smaller than the one resulting from the moment matching function. This means, the predictions from the model are more accurate. However, in order to choose the number of estimated parameters independently from the number of experiments considered, the moment matching objective function needs to be used. It is conceivable that the presented methodology can be extended in further studies in such a way that a global optimum for the unknown parameters can be found. This would remove the current constraint that the solution is dependent on the *a priori* guess of the model parameters and their uncertainties.

Nomenclature

Cov	Covariance of the random variables X and Y	$[X\,Y]$
c	uncertainty factor	$[x]$

$\mathbb{E}$	expectation of the random variable X	$[X]$
g	breakage frequency	s^{-1}
K	number of model parameters	-
$\widehat{K}_0$	rate constant for coalescence	m^3
$\widehat{k}_{\mathrm{att}}$	rate constant for breakage	$\mathrm{s\,m}^{-5}$
k_{porred}	rate constant for compaction	$\mathrm{s\,m}^{-1}$
k_{reac}	rate constant for reaction	$\mathrm{m\,s}^{-1}$
l_{e}	external liquid	m^3
l_{i}	internal liquid	m^3
N	number of experimental observations	-
n_{impeller}	impeller speed	s^{-1}
p	pores	m^3
s_{o}	original solid	m^3
s_{r}	reacted solid	m^3
Var	variance of the random variable X	$[X^2]$
X	random variable	$[X]$
x	model parameter (treated as random)	$[x]$
Y	random variable	$[Y]$
Z	random variable	$[Z]$

Greek letters

β	parameter of response surface	$[\eta]$
ε	porosity	-
η	model response	$[\eta]$
η^{exp}	experimental response	$[\eta^{\mathrm{exp}}]$
μ	model prediction	$[\eta]$
ξ	random variable	-
σ	uncertainty	$[\eta]$
Φ	objective function	$[\Phi]$

Superscripts

*	optimum
(0)	initial value
exp	experiment

Subscripts

0	base value
i	counting variable
k	counting variable
l	counting variable

s counting variable
t counting variable

Acknowledgments

The authors thank the EPSRC (grants EP/E01724X/1 and EP/E01772X/1) for funding their research.

A Variance of expanded second order response surface

The model response as a function of $\boldsymbol{x}$, $\boldsymbol{c}$, $\boldsymbol{\xi}$ is given by

$$\eta\left(\boldsymbol{x_0}, \boldsymbol{c}, \boldsymbol{\xi}\right) = \eta\left(\boldsymbol{x_0}\right) + \sum_{k=1}^{K} \beta_k \, c_k \, \xi_k + \sum_{k=1}^{K} \sum_{l \geq k}^{K} \beta_{kl} \, c_k \, c_l \, \xi_k \, \xi_l + \sum_{k=1}^{K} \sum_{l \geq k}^{K} \beta_{kl} \left(x_{0,k} \, c_l \, \xi_l + x_{0,l} \, c_k \, \xi_k\right) . \quad (12)$$

Taking the variance of eq. (12) results in,

$$\mathrm{Var}\left(\eta\right) = \mathrm{Var}\left(\eta(\boldsymbol{x_0}) + \sum_{k=1}^{K} \beta_k \, c_k \, \xi_k + \sum_{k=1}^{K} \sum_{l \geq k}^{K} \beta_{kl} \, c_k \, c_l \, \xi_k \, \xi_l + \sum_{k=1}^{K} \sum_{l \geq k}^{K} \beta_{kl} \, \left(x_{0,k} \, c_l \, \xi_l + x_{0,l} \, c_k \, \xi_k\right) \right) \quad . \quad (24)$$

With $\mathrm{Var}\left(\eta(\boldsymbol{x_0})\right) = 0$ and $\mathrm{Cov}\left(X, a\right) = 0$, where a is a constant, eq. (24) can be simplified and rewritten as

$$\mathrm{Var}\left(\eta\right) = \mathrm{Var}\left(\underbrace{\sum_{k=1}^{K} \beta_k \, c_k \, \xi_k}_{A} + \underbrace{\sum_{k=1}^{K} \sum_{l>k}^{K} \beta_{kl} \, c_k \, c_l \, \xi_k \, \xi_l}_{B} + \underbrace{\sum_{k=1}^{K} \beta_{kk} \, c_k^2 \, \xi_k^2}_{C} + \underbrace{\sum_{k=1}^{K} \sum_{l>k}^{K} \beta_{kl} \left(x_{0,k} \, c_l \, \xi_l + x_{0,l} \, c_k \, \xi_k\right)}_{D} + \underbrace{\sum_{k=1}^{K} 2 \, \beta_{kk} \, x_{0,k} \, c_k \, \xi_k}_{E} \right) \quad . \quad (25)$$

In the following we use the relationships:

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}\left(X_i, X_j\right) \tag{26}$$

$$= \sum_{i=1}^{n} \mathrm{Var}\left(X_i\right) + 2\sum_{i=1}^{n}\sum_{j>i}^{n} \mathrm{Cov}\left(X_i, X_j\right) \quad \text{with} \tag{27}$$

$$\mathrm{Cov}\left(X_i, X_j\right)_{i=j} = \mathrm{Var}\left(X_i\right) \;\;,\;\; \mathrm{Cov}\left(X_i, X_j\right) = 0 \;\text{ if } X_i, X_j \text{ i.i.d.} \tag{28}$$

$$\mathrm{Cov}\left(X, V+W\right) = \mathrm{Cov}\left(X, V\right) + \mathrm{Cov}\left(X, W\right) \tag{29}$$

$$\mathrm{Cov}\left(X, b\,Z\right) = b\,\mathrm{Cov}\left(X, Z\right) \quad . \tag{30}$$

Hence eq. (25) takes following form,

$$\begin{aligned}\mathrm{Var}\left(\eta\right) = {} & \mathrm{Var}\left(A\right) + \mathrm{Var}\left(B\right) + \mathrm{Var}\left(C\right) + \mathrm{Var}\left(D\right) + \mathrm{Var}\left(E\right) \\ & + 2\,\mathrm{Cov}\left(A, B\right) + 2\,\mathrm{Cov}\left(A, C\right) + 2\,\mathrm{Cov}\left(A, D\right) + 2\,\mathrm{Cov}\left(A, E\right) \\ & + 2, \mathrm{Cov}\left(B, C\right) + 2\,\mathrm{Cov}\left(B, D\right) + 2\,\mathrm{Cov}\left(B, E\right) \\ & + 2\,\mathrm{Cov}\left(C, D\right) + 2\,\mathrm{Cov}\left(C, E\right) + 2\,\mathrm{Cov}\left(D, E\right) \quad .\end{aligned} \tag{31}$$

In detail the different terms equate as

$$\mathrm{Var}\left(\sum_{k=1}^{K} \beta_k\, c_k\, \xi_k\right) = \sum_{k=1}^{K} \beta_k^2\, c_k^2\, \mathrm{Var}\left(\xi_k\right) = \sum_{k=1}^{K} \beta_k^2\, c_k^2 \quad , \tag{32}$$

$$\mathrm{Var}\left(\sum_{k=1}^{K}\sum_{l>k}^{K} \beta_{kl}\, c_k\, c_l\, \xi_k\, \xi_l\right) = \sum_{k=1}^{K}\sum_{l>k}^{K} \beta_{kl}^2\, c_k^2\, c_l^2\, \mathrm{Var}\left(\xi_k\, \xi_l\right) = \sum_{k=1}^{K}\sum_{l>k}^{K} \beta_{kl}^2\, c_k^2\, c_l^2 \quad , \tag{33}$$

because ξ_k and ξ_l are independent for $k \neq l$, $\mathrm{Var}\,(\xi_k\,\xi_l) = E\,(\xi_k^2\,\xi_l^2) - [E(\xi_k)\,E(\xi_l)]^2$ and $\mathrm{Cov}\,(\xi_k\,\xi_l, \xi_k\,\xi_m) = 0$ if k, l, m are all different . We note that $\mathrm{Var}\,(\chi_n^2 = X_1^2, \ldots, X_n^2) = 2\,n$ for $X_i \sim \mathcal{N}(0,1)$ $(i = 1, \ldots, n)$.

$$\mathrm{Var}\left(\sum_{k=1}^{K} \beta_{kk}\,c_k^2\,\xi_k^2\right) = \sum_{k=1}^{K} (\beta_{kk}\,c_k^2)^2\,\mathrm{Var}\left(\xi_k^2\right) = 2\,\sum_{k=1}^{K} (\beta_{kk}\,c_k^2)^2 \quad , \tag{34}$$

$$\begin{aligned}
&\mathrm{Var}\left(\sum_{k=1}^{K}\sum_{l>k}^{K} \beta_{kl}\,(x_{0,k}\,c_l\,\xi_l + x_{0,l}\,c_k\,\xi_k)\right) \\
&\quad = \sum_{k=1}^{K}\sum_{l>k}^{K} \beta_{kl}^2\,\left[x_{0,k}^2\,c_l^2\,\mathrm{Var}\,(\xi_l) + x_{0,l}^2\,c_k^2\,\mathrm{Var}\,(\xi_k)\right] \\
&\qquad + \sum_{k=1}^{K}\sum_{l>k}^{K}\sum_{s=1}^{K}\sum_{t>s}^{K} \beta_{kl}\,\beta_{st}\,\mathrm{Cov}\,(x_{0,k}\,c_l\,\xi_l + x_{0,l}\,c_k\,\xi_k, x_{0,s}\,c_t\,\xi_t + x_{0,t}\,c_s\,\xi_s) \\
&\quad = \sum_{k=1}^{K}\sum_{l>k}^{K} \beta_{kl}^2\,\left[x_{0,k}^2\,c_l^2 + x_{0,l}^2\,c_k^2\right] \\
&\qquad + \sum_{k=1}^{K}\sum_{\substack{s=1\\s\neq k}}^{K}\sum_{\substack{l>k\\l>s}}^{K} \beta_{kl}\,\beta_{sl}\,x_{0,k}\,c_l^2\,x_{0,s} + \sum_{s=1}^{K}\sum_{k>s}^{K}\sum_{l>k}^{K} \beta_{kl}\,\beta_{sk}\,x_{0,l}\,c_k^2\,x_{0,s} \\
&\qquad + \sum_{k=1}^{K}\sum_{l>k}^{K}\sum_{t>l}^{K} \beta_{kl}\,\beta_{lt}\,x_{0,k}\,c_l^2\,x_{0,t} + \sum_{k=1}^{K}\sum_{l>k}^{K}\sum_{\substack{t>k\\t\neq l}}^{K} \beta_{kl}\,\beta_{kt}\,x_{0,l}\,c_k^2\,x_{0,t}\,.
\end{aligned} \tag{35}$$

$$\mathrm{Var}\left(\sum_{k=1}^{K} 2\,\beta_{kk}\,x_{0,k}\,c_k\,\xi_k\right) = \sum_{k=1}^{K} (2\,\beta_{kk}\,x_{0,k}\,c_k)^2 \quad . \tag{36}$$

For the computation of the covariances a few more relationships are needed,

$$\mathrm{Cov}\,(X, Y) = E(X\,Y) - E(X)\,E(Y)\,. \tag{37}$$

Let X, Y, Z be i.i.d. and $\sim \mathcal{N}(0,1)$ then

$$\mathbb{E}(X\,Y\,Z) = 0 \;\;,\;\; \mathbb{E}(X^2\,Y) = 0 \;\;,\;\; \mathbb{E}(X\,Y^2) = 0 \;\;,\;\; \mathbb{E}(X^3) = 0 \quad . \tag{38}$$

The covariances equate subsequently as,

$$\mathrm{Cov}\,(A, B) = 0\,, \tag{39}$$

$$\mathrm{Cov}\,(A, C) = 0\,, \tag{40}$$

$$\mathrm{Cov}\,(A, D) = \sum_{k=1}^{K}\sum_{l>k}^{K} \beta_{kl}\,\left(\beta_l\,x_{0,k}\,c_l^2 + \beta_k\,x_{0,l}\,c_k^2\right)\,, \tag{41}$$

$$\mathrm{Cov}\,(A, E) = \sum_{k=1}^{K} 2\,\beta_k\,\beta_{kk}\,x_{0,k}\,c_k^2\,, \tag{42}$$

$$\mathrm{Cov}\,(B, C) = 0\,, \tag{43}$$
$$\mathrm{Cov}\,(B, D) = 0\,, \tag{44}$$
$$\mathrm{Cov}\,(B, E) = 0\,, \tag{45}$$
$$\mathrm{Cov}\,(C, D) = 0\,, \tag{46}$$
$$\mathrm{Cov}\,(C, E) = 0\,, \tag{47}$$
$$\mathrm{Cov}\,(D, E) = \sum_{k=1}^{K} \sum_{l>k}^{K} 2\, x_{0,k}\, x_{0,l}\, \beta_{kl}\, (c_k^2\, \beta_{kk} + c_l^2\, \beta_{ll})\,. \tag{48}$$

Subsequently these calculations yield the variance of the model response,

$$\begin{aligned}
\sigma^2(\boldsymbol{x_0}, \boldsymbol{c}) &= \mathrm{Var}\,(\eta\,(\boldsymbol{x_0}, \boldsymbol{c}, \boldsymbol{\xi})) \\
&= \sum_{k=1}^{K} \beta_k^2\, c_k^2 + \sum_{k=1}^{K} \sum_{l>k}^{K} \beta_{kl}^2\, c_k^2\, c_l^2 + 2 \sum_{k=1}^{K} (\beta_{kk}\, c_k^2)^2 \\
&\quad + \sum_{k=1}^{K} \sum_{l>k}^{K} \beta_{kl}^2 \left[(x_{0,k}\, c_l)^2 + (x_{0,l}\, c_k)^2 \right] \\
&\quad + \sum_{k=1}^{K} \sum_{\substack{s=1 \\ s\neq k}}^{K} \sum_{\substack{l>k \\ l>s}}^{K} \beta_{kl}\, \beta_{sl}\, x_{0,k}\, c_l^2\, x_{0,s} + \sum_{s=1}^{K} \sum_{k>s}^{K} \sum_{l>k}^{K} \beta_{kl}\, \beta_{sk}\, x_{0,l}\, c_k^2\, x_{0,s} \\
&\quad + \sum_{k=1}^{K} \sum_{l>k}^{K} \sum_{t>l}^{K} \beta_{kl}\, \beta_{lt}\, x_{0,k}\, c_l^2\, x_{0,t} + \sum_{k=1}^{K} \sum_{l>k}^{K} \sum_{\substack{t>k \\ t\neq l}}^{K} \beta_{kl}\, \beta_{kt}\, x_{0,l}\, c_k^2\, x_{0,t} \\
&\quad + \sum_{k=1}^{K} (2\, \beta_{kk}\, x_{0,k}\, c_k)^2 + 2 \sum_{k=1}^{K} \sum_{l>k}^{K} \beta_{kl} \left(\beta_l\, x_{0,k}\, c_l^2 + \beta_k\, x_{0,l}\, c_k^2 \right) \\
&\quad + 4 \sum_{k=1}^{K} \beta_k\, \beta_{kk}\, x_{0,k}\, c_k^2 + 4 \sum_{k=1}^{K} \sum_{l>k}^{K} x_{0,k}\, x_{0,l}\, \beta_{kl} \left(c_k^2\, \beta_{kk} + c_l^2 \beta_{ll} \right)\,.
\end{aligned} \tag{14}$$

References

[1] A. A. Adetayo, J. D. Litster, S. E. Pratsinis, and B. J. Ennis. Population balance modelling of drum granulation of material with wide size distribution. *Powder Technology*, 82:37–49, 1995. doi:10.1016/0032-5910(94)02896-V.

[2] R. Boerefijn and M. J. Hounslow. Studies of fluid bed granulation in an industrial R&D context. *Chemical Engineering Science*, 60:3879–3890, 2005. doi:10.1016/j.ces.2005.02.021.

[3] A. S. Bramley, M. J. Hounslow, and R. L. Ryall. Aggregation during precipitation from solution: A method for extracting rates from experimental data. *Journal of Colloid and Interface Science*, 183:155–165, 1996. doi:10.1006/jcis.1996.0530.

[4] A. Braumann and M. Kraft. Incorporating experimental uncertainties into multivariate granulation modelling, 2009. Technical Report 66, c4e-Preprint Series. Submitted to Chemical Engineering Science for publication.

[5] 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*, 62:4717–4728, 2007. doi:10.1016/j.ces.2007.05.028.

[6] A. Braumann, M. Kraft, and P. R. Mort. Applying response surface methodology to multidimensional granulation modelling, 2008. Technical Report 56, c4e-Preprint Series. Submitted to Powder Technology for publication.

[7] M. Celnik, R. Patterson, M. Kraft, and W. Wagner. Coupling a stochastic soot population balance to gas-phase chemistry using operator splitting. *Combustion and Flame*, 148:158–176, 2007. doi:10.1016/j.combustflame.2006.10.007.

[8] A. Darelius, H. Brage, A. Rasmuson, I. N. Björn, and S. Folestad. A volume-based multi-dimensional population balance approach for modelling high shear granulation. *Chemical Engineering Science*, 61:2482–2493, 2006. doi:10.1016/j.ces.2005.11.016.

[9] M. P. Elsner, G. Ziomek, and A. Seidel-Morgenstern. Simultaneous preferential crystallization in a coupled, batch operation mode—Part I: Theoretical analysis and optimization. *Chemical Engineering Science*, 62:4760–4769, 2007. doi:10.1016/j.ces.2007.05.035.

[10] B. J. Ennis, G. Tardos, and R. Pfeffer. A microlevel-based characterization of granulation phenomena. *Powder Technology*, 65:257–272, 1991. doi:10.1016/0032-5910(91)80189-P.

[11] A. G. Fredrickson and N. V. Mantzaris. A new set of population balance equations for microbial and cell cultures. *Chemical Engineering Science*, 57:2265–2278, 2002. doi:10.1016/S0009-2509(02)00116-1.

[12] A. Gerstlauer, A. Mitrović, S. Motz, and E.-D. Gilles. A population model for crystallization processes using two independent particle properties. *Chemical Engineering Science*, 56:2553–2565, 2001. doi:10.1016/S0009-2509(00)00448-6.

[13] C. D. Immanuel and F. J. Doyle III. Computationally efficient solution of population balance models incorporating nucleation, growth and coagulation: application to emulsion polymerization. *Chemical Engineering Science*, 58:3681–3698, 2003. doi:10.1016/S0009-2509(03)00216-1.

[14] S. M. Iveson, J. D. Litster, and B. J. Ennis. Fundamental studies of granule consolidation, part 1: Effects of binder content and binder viscosity. *Powder Technology*, 88:15–20, 1996. doi:10.1016/0032-5910(96)03096-3.

[15] N. Morgan, M. Kraft, M. Balthasar, D. Wong, M. Frenklach, and P. Mitchell. Numerical simulations of soot aggregation in premixed laminar flames. *Proceedings of the Combustion Institute*, 31:693–700, 2007. doi:10.1016/j.proci.2006.08.021.

[16] R. I. A. Patterson and M. Kraft. Models for the aggregate structure of soot particles. *Combustion and Flame*, 151:160–172, 2007. doi:10.1016/j.combustflame.2007.04.012.

[17] J. M.-H. Poon, C. D. Immanuel, F. J. Doyle III, and J. D. Litster. A three-dimensional population balance model of granulation with a mechanistic representation of the nucleation and aggregation phenomena. *Chemical Engineering Science*, 63:1315–1329, 2008. doi:10.1016/j.ces.2007.07.048.

[18] D. Ramkrishna and A. W. Mahoney. Population balance modeling. Promise for the future. *Chemical Engineering Science*, 57:595–606, 2002. doi:10.1016/S0009-2509(01)00386-4.

[19] C. F. W. Sanders, A. W. Willemse, A. D. Salman, and M. J. Hounslow. Development of a predictive high-shear granulation model. *Powder Technology*, 138:18–24, 2003. doi:10.1016/j.powtec.2003.08.046.

[20] H.-C. Schwarzer, F. Schwertfirm, M. Manhart, H.-J. Schmid, and W. Peukert. Predictive simulation of nanoparticle precipitation based on the population balance equation. *Chemical Engineering Science*, 61:167–181, 2006. doi:10.1016/j.ces.2004.11.064.

[21] D. A. Sheen, X. You, H. Wang, and T. Løvås. Spectral uncertainty quantification, propagation and optimization of a detailed kinetic model for ethylene combustion. *Proceedings of the Combustion Institute*, 32(1):535–542, 2009. doi:10.1016/j.proci.2008.05.042.

[22] T. Simmons, R. Turton, and P. Mort. An investigation into the effects of time and shear rate on the spreading of liquids in coating and granulation processes. In *Fifth World Congress on Particle Technology*, 2006.

[23] J. Singh, M. Balthasar, M. Kraft, and W. Wagner. Stochastic modeling of soot particle size and age distributions in laminar premixed flames. *Proceedings of the Combustion Institute*, 30:1457–1465, 2005. doi:10.1016/j.proci.2004.08.120.

[24] A. Vikhansky and M. Kraft. A Monte Carlo methods for identification and sensitivity analysis of coagulation processes. *Journal of Computational Physics*, 200:50–59, 2004. doi:10.1016/j.jcp.2004.03.006.

[25] A. Vikhansky, M. Kraft, M. Simon, S. Schmidt, and H.-J. Bart. Droplets population balance in a rotating disc contactor: An inverse problem approach. *AIChE Journal*, 52(4):1441–1450, 2006. doi:10.1002/aic.10735.

[26] M. Wulkow. The simulation of molecular weight distributions in polyreaction kinetics by discrete galerkin methods. *Macromolecular Theory and Simulations*, 5: 393–416, 1996. doi:10.1002/mats.1996.040050303.