Journal of Process Control 166 (2026) 103804

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Sparse identification of physically plausible aggregation kernels for wet granulation processes☆,☆☆

Stefan R. Tölle [a,*], Lorenz Dörschel [a], Stefan Klinken-Uth [b], Alana Delvos [b], Jörg Breitkreutz [b], Heike Vallery [a,c], Sebastian Stemmler [a]

[a] Institute of Automatic Control, RWTH Aachen University, Aachen, Germany
[b] Institute of Pharmaceutics and Biopharmaceutics, Heinrich Heine University, Düsseldorf, Germany
[c] Department of Biomechanical Engineering, Delft University of Technology, Delft, The Netherlands

ARTICLE INFO

Dataset link: https://doi.org/10.18154/RWTH-2026-01122

Keywords:
Chemical process control
Pharmaceutical systems
Population balance models
Sparse identification of nonlinear dynamics
Wet granulation
Continuous manufacturing

ABSTRACT

Population balance models provide a mathematical framework for describing the dynamics of particulate systems. For aggregation processes, such as twin-screw wet granulation, population balance models critically depend on accurate aggregation kernels, yet first-principles derivation is often impractical, and data-driven methods can lack physical interpretability. This work presents a sparse identification framework for learning physically plausible aggregation kernels directly from data. The approach enforces physical constraints, exploits structural properties, applies a systematic scaling strategy, and extends naturally to actuated systems. Validation on experimental data from a continuous twin-screw wet granulation process confirms the method's robustness and its ability to recover physically meaningful aggregation kernels from real-world measurement data. The identified population balance model was able to predict the median particle size with a mean absolute error of 134.8 µm (10 %).

1. Introduction

Wet granulation is a core operation in pharmaceutical manufacturing, where fine powders are aggregated into larger granules to improve flowability, compressibility, and product uniformity [1]. Controlling this process is critical to ensure consistent product quality and efficient operation [2,3]. However, granulation dynamics are inherently complex and difficult to capture with conventional lumped-parameter models.

While wet granulation is a prominent example, particulate systems arise in many other fields, including chemical, biological, environmental, and materials engineering. The state of a particle is determined by both its internal coordinates (e.g., size, shape, or composition) and its external coordinates (e.g., spatial position), making the modeling task inherently challenging. Population balance models (PBMs) use integro-partial differential equations to represent the time evolution of particle populations along these coordinates. A comprehensive review is given by Ramkrishna and Singh [4].

Population balance equations (PBEs) provide a mechanistic modeling framework for key particulate phenomena such as growth, aggregation, and breakage. In wet granulation, aggregation is the dominant mechanism. Aggregation involves particles binding to form larger granules, with the aggregation kernel defining the rate and likelihood of these interactions. Deriving this aggregation kernel from first principles requires detailed knowledge on particle interaction and often involves empirical elements [5]. While such models are valuable for understanding underlying physics, in many practical engineering settings, it is more feasible to identify the aggregation kernel directly from data. Existing approaches to this inverse problem either (1) fit unknown parameters in predefined mechanistic aggregation kernels (e.g., [6,7]) or (2) construct flexible kernel functions from local basis functions [8]. The first approach is limited by the chosen model structure, which

☆ This article is part of a Special issue entitled: 'TC 6.1 Chemical Process Control (IFAC WC 2026)' published in Journal of Process Control.

☆☆ The authors appreciate the financial support of the German Research Foundation (DFG) received within the research program SPP 2364 "Autonomous processes in particle technology" for the project "Model-based quality control in continuous manufacturing of pharmaceutical granules and tablets (QC4CM)" (Nr. 504702251). We would like to acknowledge LB Bohle for their support and for granting access to the QbCon® 1. The authors would also like to thank DFE Pharma, JRS Pharma, Janssen Pharmaceutica, and Ashland Industries Deutschland for providing Pharmatose® 300M, Vivapur® 102, Ibuprofen, and Plasdone™ K-29/32.

* Corresponding author.
E-mail address: S.Toelle@irt.rwth-aachen.de (S.R. Tölle).

https://doi.org/10.1016/j.jprocont.2026.103804
Received 31 October 2025; Received in revised form 16 April 2026; Accepted 25 July 2026
Available online 6 August 2026

might has limited applicability. The second is more flexible but often less interpretable and prone to overfitting due to the large number of free parameters.

Tiong et al. [9] recently proposed the use of the sparse identification of nonlinear dynamics (SINDy) framework for kernel identification. SINDy [10] discovers governing equations from data by selecting a sparse combination of functions from a predefined library, thereby avoiding exhaustive combinatorial searches. Tiong et al. [9] demonstrated that this framework can recover canonical PBMs from synthetic data with noise.

While Tiong et al. [9] demonstrated the feasibility of combining SINDy with PBMs, their study focused primarily on showcasing the conceptual framework rather than establishing a practically deployable methodology. A major drawback of the approach is that the identified kernels are not necessarily physically plausible. Their proposed remedy, adding an additional cost term to the optimization problem, substantially increases algorithmic complexity and tuning effort. Moreover, inputs as considered in SINDy with control [11] were not incorporated in their formulation.

Building on this foundation, the presented work advances the approach in two key directions. First, we introduce several methodological refinements that elevate the previously conceptual framework to real-world applicability. In particular:

(1) Reformulating the identification problem such that a physically plausible model structure is inherently preserved.
(2) Introducing a systematic scaling and weighting strategy for SINDy applied to PBMs.
(3) Physically constraining the sparse identification problem to improve stability.
(4) Extending the framework to actuated particulate systems, enabling the incorporation of process inputs.

Second, these refinements enable the application of sparse identification to real particulate process data. One such application example is continuous twin-screw wet granulation. By successfully applying the proposed method to experimental data from twin-screw wet granulation, we demonstrate both practical relevance and robustness of the proposed approach.

Our contribution aims to make PBMs more accessible for real-world engineering applications, enabling automated identification of physically consistent PBEs directly from data. The framework can operate without detailed process knowledge, as the candidate aggregation kernel library can be defined a priori and applied to different aggregation processes. When expert knowledge is available, it can be seamlessly integrated to further refine the model. The extension to actuated systems also makes the approach particularly well-suited for control applications, such as model predictive control in particulate processes.

The paper is organized as follows. Section 2 introduces the granulation process under investigation and derives a simplified PBE describing the process. Section 3 details the proposed methodological advancements. Section 4 describes the procedure to validate the approach. The results of the aggregation kernel identification scheme are presented in chapter 5 and discussed in chapter 6. The outlook in Section 7 concludes the paper.

2. Problem formulation and modeling

2.1. Process description and modeling assumptions

In this paper we demonstrate the proposed aggregation kernel identification algorithm on a continuous twin-screw wet granulation process. The process is schematically illustrated in Fig. 1. A continuous stream of particles enters the twin-screw wet granulator (TSG) and is transported along the spatial dimension z by two co-rotating screws

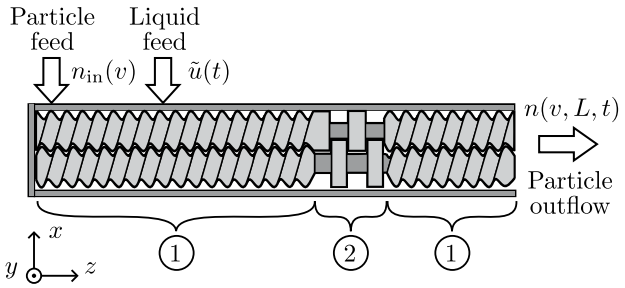


Fig. 1. Schematic of the continuous twin-screw granulator with (1) screw zone and (2) kneading zone.

with length L. Wetting with a granulation liquid occurs toward the inlet of the barrel. The screw zone (1) is interrupted by a kneading zone (2) intended to improve granule formation. At the outlet located at $z = L$, the granules leave the system for further downstream processing.

The quantity of interest when modeling the particle size distribution (PSD) is the number density function $n(v, \boldsymbol{r}, t)$. It describes the density of particles at time t along the external coordinates $\boldsymbol{r} = [x, y, z]^{\mathrm{T}}$ and the internal coordinate v, here representing the particle volume. The particles enter the granulator with a given number density function n_{in}. The liquid-to-solid (L/S) ratio $\tilde{u}(t)$ has substantial influence on the resulting PSD and is thus treated as the control input of the system.

Furthermore, the following modeling assumptions were made:

A1 The remaining process parameters, screw speed ω_{screw} and the dry-solid volume flow rate q, are constant.
A2 The granulation dynamics can be adequately represented by pure aggregation.
A3 In the screw zone, the particle movement can be described by plug-flow, and no agglomeration occurs due to the absence of relative particle movement.
A4 In the kneading zone, the particles are well mixed.

Assumptions A3 and A4 are necessary, as it is only possible to measure the PSD outside of the TSG. This makes fitting and validating the dependence of the number density function on the spatial coordinate intractable without these simplifications. Experimental support for assumptions A3 and A4 is provided by the study performed by Chan Seem et al. [12].

2.2. Population balance formulation

The distributed characteristics of the process at hand make it necessary to use population balances to capture the dynamics of the particle population. The general PBE reads as

$$\frac{\partial}{\partial t} n + \nabla_v \cdot \dot{V} n + \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \dot{\boldsymbol{R}} n = h(v, \boldsymbol{r}, \tilde{u}), \tag{1}$$

with the particle flux $\dot{V} n$ along the internal coordinate, the particle flux $\dot{\boldsymbol{R}} n$ along the external coordinates, and the source term h, which models particle birth and death due to aggregation and breakage. For a pure aggregation process (A2), $\dot{V} n = 0$ holds. Hence, process modeling ultimately amounts to modeling the particle movement (i.e. finding $\dot{\boldsymbol{R}} n$) and agglomeration (i.e. finding $h(v, \boldsymbol{r}, \tilde{u})$).

The particle flux along the external coordinates can be modeled using an axial dispersion model in z-direction. Particle flux differs in the screw and kneading zone of the TSG:

$$\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \dot{\boldsymbol{R}} n = \begin{cases} v_1 \frac{\partial n}{\partial z} - D_1 \frac{\partial^2 n}{\partial z^2}, & \text{screw zone} \\ v_2 \frac{\partial n}{\partial z} - D_2 \frac{\partial^2 n}{\partial z^2}, & \text{kneading zone} \end{cases}, \tag{2}$$

with the velocities v_1, v_2 in z-direction and the dispersion coefficients D_1, D_2.

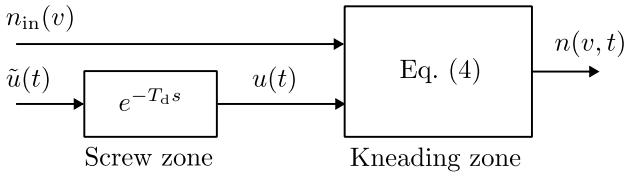


Fig. 2. Simplified dynamic modeling in the screw zone and the kneading zone.

Due to A3 and A4, we can immediately identify ($D_1 = 0$) and ($v_2 = 0$, $D_2 \to \infty$). This greatly simplifies the analysis as only the PSD inside the kneading zone is relevant for system description. In this case, the screw zones can be combined and described by a time-delay with delay T_{d}. Furthermore, n is not a function of the external coordinates anymore because the particle population is uniformly distributed in space for the kneading zone (A4).

For well-mixed open systems, the general PBE (1) can be simplified [13]. The stationary particle stream enters the spatial domain $\Omega_{\boldsymbol{r}}$ of the kneading zone through the left boundary $\partial\Omega_{\boldsymbol{r},\text{in}}$ and exits through the right boundary $\partial\Omega_{\boldsymbol{r},\text{out}}$. Integrating (1) over the domain $\Omega_{\boldsymbol{r}}$ of the external coordinates with volume $V_{\boldsymbol{r}}$ and using the divergence theorem yields:

$$\frac{\partial}{\partial t}\underbrace{\int_{\Omega_{\boldsymbol{r}}} n\mathrm{d}V_{\boldsymbol{r}}}_{=V_{\boldsymbol{r}}n} + \underbrace{\int_{\partial\Omega_{\boldsymbol{r},\text{in}}} \dot{\boldsymbol{R}}n\cdot\mathrm{d}\boldsymbol{A}_{\boldsymbol{r}}}_{=-qn_{\text{in}}} + \underbrace{\int_{\partial\Omega_{\boldsymbol{r},\text{out}}} \dot{\boldsymbol{R}}n\cdot\mathrm{d}\boldsymbol{A}_{\boldsymbol{r}}}_{=qn} = \underbrace{\int_{\Omega_{\boldsymbol{r}}} h(v,\tilde{u}(t-T_{\text{d}}))\mathrm{d}V_{\boldsymbol{r}}}_{=V_{\boldsymbol{r}}h(v,\tilde{u}(t-T_{\text{d}}))}. \tag{3}$$

By rearranging the above equation, we derive the following simplified PBE:

$$\frac{\partial}{\partial t}n(v,t) - \frac{1}{T}\left(n_{\text{in}}(v) - n(v,t)\right) = h(v,u,t), \tag{4}$$

with the mean residence time $T = V_{\boldsymbol{r}}/q$ and the delayed input $u(t) = \tilde{u}(t-T_{\text{d}})$. Overall, the simplified model structure shown in Fig. 2 results.

The residence time distribution (RTD) $E(t)$ of the TSG for the present case is given by

$$E(t) = \begin{cases} 0, & \text{for } t < T_{\text{d}} \\ \frac{1}{T}\exp\left(-\frac{t-T_{\text{d}}}{T}\right), & \text{for } t \geq T_{\text{d}} \end{cases}. \tag{5}$$

Particle movement is therefore characterized by the two unknown parameters T and T_{d}. To find suitable parameters, model (5) has to be fitted to experimental residence time data.

Due to A2, the r.h.s. of (4) is given by the Smoluchowski coagulation equation:

$$h(v,u,t) = \underbrace{\frac{1}{2}\int_0^v \beta(v-v',v',u)n(v-v',t)n(v',t)\mathrm{d}v'}_{\text{Particle birth}} - \underbrace{\int_0^\infty \beta(v,v',u)n(v,t)n(v',t)\mathrm{d}v'}_{\text{Particle death}}, \tag{6}$$

with the unknown aggregation kernel β. Since the control input u influences the rate at which particles coagulate, the aggregation kernel is also considered a function of the input.

2.3. Spatial and temporal discretization

To formulate a sparse parameter fitting problem, the PBE (4) has to be discretized in the temporal and internal coordinates. We define the uniform temporal grid $\mathcal{T} = \{t_1, t_2, \ldots, t_{n_t}\}$ and the not necessarily uniform grid of particle sizes $\mathcal{V} = \{v_1, v_2, \ldots, v_{n_v}\}$. Thus the number of sample points is given by n_t and n_v, respectively. The number of particles per unit volume with a particle size in the interval $[v_i, v_{i+1})$ is given by:

$$N_i(t) = \int_{v_i}^{v_{i+1}} n(v,t)\mathrm{d}v. \tag{7}$$

Integrating (4) over the interval $[v_i, v_{i+1})$ and approximating the time derivative using central differences yields the discretized PBE:

$$\frac{N_i(t_{j+1}) - N_i(t_{j-1})}{2T_{\text{s}}} - \frac{N_{\text{in}}(v_i) - N_i(t_j)}{T} = \int_{v_i}^{v_{i+1}} h(v,u,t_j)\mathrm{d}v, \tag{8}$$

with the sample time T_{s} and the number of particles N_{in} at the inlet, which is defined analogous to (7).

For the discretization of the Smoluchowski coagulation Eq. (6), different approaches are available in literature. We use the fixed pivot technique proposed by Kumar and Ramkrishna [14], as this discretization exactly preserves the integral properties particle number and mass compared to the continuous PBE. Using the fixed pivot technique, the r.h.s. of (8) for a given aggregation kernel β is given by the expression

$$\begin{aligned} H(v_i,u,t_j,\beta) &:= \int_{v_i}^{v_{i+1}} h(v,u,t_j)\mathrm{d}v \\ &= \sum_{\substack{k,l \\ v_{i-1}\leq(v_k+v_l)\leq v_{i+1}}}^{k\geq l}\left(1-\frac{1}{2}\delta_{k,l}\right)\eta\beta(v_k,v_l,u)N_k(t_j)N_l(t_j) \\ &\quad - N_i(t_j)\sum_{k=1}^{n_v}\beta(v_i,v_k,u)N_k(t_j), \end{aligned} \tag{9}$$

with η defined as

$$\eta = \begin{cases} \frac{v_{i+1}-v}{v_{i+1}-v_i}, & \text{for } v_i \leq v \leq v_{i+1} \\ \frac{v-v_{i-1}}{v_i-v_{i-1}}, & \text{for } v_{i-1} \leq v \leq v_i \end{cases} \tag{10}$$

and the Kronecker delta $\delta_{k,l}$. For more details, the reader is referred to Kumar and Ramkrishna [14].

2.4. SINDy for population balance models

With the system equations discretized, the SINDy framework can be employed to identify the aggregation kernel from data.

According to the SINDy algorithm, an unknown scalar function

$$f(\boldsymbol{x}) = \boldsymbol{\varphi}(\boldsymbol{x})^{\text{T}}\boldsymbol{\xi} \tag{11}$$

is modeled as a linear combination of typically a large number K of nonlinear candidate functions φ_k. The library of candidate functions is given by

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \begin{bmatrix}\varphi_1(\boldsymbol{x}) & \ldots & \varphi_K(\boldsymbol{x})\end{bmatrix}^{\text{T}} \tag{12}$$

and the parameter vector is given by

$$\boldsymbol{\xi} = \begin{bmatrix}\xi_1 & \xi_2 & \ldots & \xi_K\end{bmatrix}^{\text{T}}. \tag{13}$$

The entries ξ_i of the parameter vector are the aggregation rate constants associated to the candidate kernel φ_i.

The parameter vector is then fitted to a set of measured training data by solving a least-squares problem. Due to the large number of possible basis functions, this approach would normally lead to a model that suffers from over-fitting. To overcome this problem, the SINDy algorithm regularizes the least-squares problem, such that the number of non-zero elements of the parameter vector $\boldsymbol{\xi}$ is penalized.

The key property that makes the SINDy algorithm suitable for kernel identification is that the source term h is a linear function of β. As (11) is linear in $\boldsymbol{\xi}$, the resulting problem is linear in the decision variables, enabling a least-square formulation as exploited by Tiong et al. [9].

For setting up such a sparse parameter fitting problem, the measurement of the number function $N_i(t)$ is used to construct the snapshot matrix

$$\boldsymbol{N} := \begin{bmatrix} N_1(t_1) & \cdots & N_1(t_{n_t}) \\ \vdots & \ddots & \vdots \\ N_{n_v}(t_1) & \cdots & N_{n_v}(t_{n_t}) \end{bmatrix}. \tag{14}$$

Then the l.h.s. of (8) for all discrete time steps t_j and discrete particle sizes v_i can be summarized in the matrix

$$\boldsymbol{Y} := \begin{bmatrix} y_1(t_1) & \cdots & y_1(t_{n_t}) \\ \vdots & \ddots & \vdots \\ y_{n_v}(t_1) & \cdots & y_{n_v}(t_{n_t}) \end{bmatrix}, \tag{15}$$

with $y_i(t_j) = \frac{N_i(t_{j+1}) - N_i(t_{j-1})}{2T_s} - \frac{N_{\text{in}}(v_i) - N_i(t_j)}{T}$.

In similar fashion, (9) can be evaluated for every snapshot and every candidate function φ_k, finally leading to the structure

$$\sum_{k=1}^{K} \boldsymbol{X}_k \xi_k \approx \boldsymbol{Y}. \tag{16}$$

The concrete formula for $\boldsymbol{X}_k$ will depend on specific choices and our version will significantly differ from the one by Tiong et al. [9].

3. Methodological advances in kernel identification

3.1. Structural exploitation of birth and death terms

Tiong et al. [9] used the SINDy algorithm to separately identify the birth and death term in (6). The separate identification neglects that particle birth and death always occur together in the pure aggregation process. We thus propose a new SINDy formulation for kernel identification that exploits the special structure of the Smoluchowski coagulation equation and thereby guarantees that the identified PBE is physically plausible.

The source term h is a linear function of ξ if the kernel from (11) is used. Substituting the aggregation kernel from (6) with (11) yields:

$$h(v,u,t) = \left[\frac{1}{2} \int_0^v \boldsymbol{\varphi}(v-v', v', u)^{\mathrm{T}} n(v-v', t) n(v', t) \mathrm{d}v' - \int_0^{\infty} \boldsymbol{\varphi}(v, v', u)^{\mathrm{T}} n(v, t) n(v', t) \mathrm{d}v' \right] \boldsymbol{\xi}. \tag{17}$$

Thus, the linear parameter ξ describes both the birth and the death due to aggregation with the same kernel. This is in contrast to the formulation used by Tiong et al. [9], where an additional term in the cost function and even manual user decisions are required to obtain physically plausible solutions. Our solution, on the other hand, structurally guarantees that particle birth and death due to aggregation occur only together, which substantially simplifies the algorithm.

Following our approach, we can write

$$\boldsymbol{X}_k := \begin{bmatrix} x_{1,k}(t_1) & \cdots & x_{1,k}(t_{n_t}) \\ \vdots & \ddots & \vdots \\ x_{n_v,k}(t_1) & \cdots & x_{n_v,k}(t_{n_t}) \end{bmatrix}, \tag{18}$$

with $x_{i,k}(t_j) = H(v_i, u, t_j, \varphi_k)$.

Another key advantage of this formulation is that the candidate functions φ_k can be directly interpreted as candidate aggregation kernels. This enables a more refined function library that can incorporate physical or expert knowledge. For example, we can restrict the kernel library to functions fulfilling the symmetry property $\varphi_k(v_a, v_b, u) = \varphi_k(v_b, v_a, u)$ to enforce physical plausibility. Additionally, we require:

$$\varphi_k(v_a, v_b, u) \geq 0, \quad \text{if } v_a, v_b, u \geq 0. \tag{19}$$

It is also possible to add mechanistic kernels, which may describe certain aggregation phenomena.

3.2. Scaling and weighting strategy

To enable systematic scaling and weighting of the sparse kernel identification algorithm, we use the weighted model residual

$$\left\| \boldsymbol{W}_1 \left(\boldsymbol{Y} - \sum_{k=1}^{K} \boldsymbol{X}_k \xi_k \right) \boldsymbol{W}_2 \right\|_F, \tag{20}$$

with the diagonal weighting matrices $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ and the Frobenius norm $\|\cdot\|_F$. With this weighting, the ℓ_0-regularized sparse identification problem is written as:

$$\hat{\boldsymbol{\xi}} = \arg\min_{\boldsymbol{\xi}} \|\boldsymbol{Y} - \boldsymbol{\Theta}\boldsymbol{\xi}\|_{\boldsymbol{W}}^2 + \sigma \|\boldsymbol{\xi}\|_0, \tag{21}$$

with $\boldsymbol{Y} = \mathrm{vec}(\boldsymbol{Y})$, $\boldsymbol{\Theta} = \left[\theta_1, \ldots, \theta_K\right]$, $\theta_k = \mathrm{vec}(\boldsymbol{X}_k)$, $\boldsymbol{W} = \left(\boldsymbol{W}_2 \otimes \boldsymbol{W}_1\right)^2$. The regularization parameter σ promotes the sparsity of the parameter vector.

The weighting matrix $\boldsymbol{W}_1$ allows weighing of the model error along the internal coordinate, and the weighting matrix $\boldsymbol{W}_2$ allows the weighing of the model error along the temporal dimension. The weighing is necessary because the entries of the snapshot matrix $\boldsymbol{N}$ have different orders of magnitude. Typically, large particle sizes correspond with low particle numbers, even if those particles contain a large fraction of the mass of the particle population. Similarly, time points at which the average particle size is low correspond to large particle numbers since the total particle mass remains constant. Thus, without normalization, the model will be biased toward fitting small particle sizes only. Since the goal is to approximate $\boldsymbol{N}$ equally well for all particle sizes and all time points, we choose the weighting based on the largest elements in the rows and columns of the snapshot matrix:

$$\boldsymbol{W}_1 = \mathrm{diag}\left(\max_j(N_1(t_j)), \ldots, \max_j(N_{n_v}(t_j)) \right)^{-1}, \tag{22}$$

$$\boldsymbol{W}_2 = \mathrm{diag}\left(\max_i(N_i(t_1)), \ldots, \max_i(N_i(t_{n_t})) \right)^{-1}. \tag{23}$$

It can also be sensible to use other statistics than row and column maxima for normalization, e.g., if the model fit in small or large particle size ranges is particularly important. However, empirically we found that ℓ_∞-based normalization works well also for approximating the mass-based PSD.

Furthermore, normalization of the candidate kernels is essential for the identification algorithm. This is because small numeric values of the kernel φ_k cause large numeric values of the parameter ξ_k. This would make the kernel φ_k more likely to be selected by the algorithm than kernels with large numeric values. Therefore, we normalize the candidate kernels to have the same integral in the considered particle size range:

$$\frac{\int_{v_1}^{v_{n_v}} \int_{v_1}^{v_{n_v}} \varphi_k(v, v', u) \mathrm{d}v \mathrm{d}v'}{\int_{v_1}^{v_{n_v}} \int_{v_1}^{v_{n_v}} 1 \mathrm{d}v \mathrm{d}v'} = 1. \tag{24}$$

In addition, the control input $u(t)$ is normalized to the interval $[0, 1]$.

3.3. Physically constrained kernel identification

Solving (21) by a combinatorial search of all possible model structures is computationally expensive. However, an approximate solution can be found, for example, with the sequential thresholded least-squares (STLS) algorithm [10]. In the context of partial differential equations (PDEs), where the columns of $\boldsymbol{\Theta}$ are usually highly correlated, the STRidge [15] is the typical choice. The algorithm iteratively performs ridge regression while setting parameters that fall under a certain threshold to zero.

Using any of the standard algorithms for problem (21) might lead to negative parameters $\hat{\xi}_k$. Because of the restriction (19), negative parameters imply a negative aggregation rate. Negative aggregation rates are not physically plausible and will most likely lead to unstable models.

Taking this into account, we proposed an adapted version of the STRidge algorithm to solve (21). The proposed algorithm only produces sparse parameter vectors $\hat{\boldsymbol{\xi}}$ that have non-negative elements. Instead of solving the unconstrained ridge regression, our modified STRidge algorithm solves a constrained quadratic program (QP). This way, the identified PBE is physically plausible and has improved stability properties. The modified STRidge algorithm is given below:

Algorithm 1: STRidge with positive coefficients

Input: $\boldsymbol{\Theta}$, $\boldsymbol{Y}$, λ, ε, iters
Output: $\hat{\boldsymbol{\xi}}$

$\hat{\boldsymbol{\xi}} = \arg\min_{\boldsymbol{\xi}} \|\boldsymbol{Y} - \boldsymbol{\Theta}\boldsymbol{\xi}\|_{\boldsymbol{W}}^2 + \lambda\|\boldsymbol{\xi}\|_2^2$, s.t. $\boldsymbol{\xi} \geq 0$ ▷ Solve QP

bigcoeffs $= \{j : \xi_j \geq \varepsilon\}$ ▷ Select large coefficients

$\hat{\boldsymbol{\xi}}[\sim \text{bigcoeffs}] = 0$ ▷ Apply hard threshold

$\hat{\boldsymbol{\xi}}[\text{bigcoeffs}] = \text{STRidge}(\boldsymbol{\Theta}\left[:, \text{bigcoeffs}\right], \boldsymbol{Y}, \lambda, \varepsilon, \text{iters}-1)$ ▷ Recursive call with fewer coefficients

return $\hat{\boldsymbol{\xi}}$

3.4. Extension to actuated particle systems

In many real-world applications, the aggregation is subject to a set of process parameters influencing the process. For example, in the presented case, the aggregation rate is dependent on the L/S ratio u. Hence, we augment the kernel library with functions of the system input, which is known as SINDy with control [11]. This makes the presented algorithm well-suited for control of particulate systems.

According to Sastry [16], the aggregation kernel can be written as the product of two functions $f_i(u)$ and $g_j(v_a, v_b)$:

$$\varphi_k(v_a, v_b, u) = f_i(u) \cdot g_j(v_a, v_b). \tag{25}$$

We note that this factorization applies to most, if not all, kernels found in the literature, e.g., all kernels listed by X. Liu and D. Litster [17] for granulation can be written in this way.

In the context of SINDy with control, structure (25) allows the systematic construction of the candidate library. For this, the user defines a set of functions of the input $f_i(u)$ and a set of size-dependent aggregation kernels $g_j(v_a, v_b)$. The candidate library is then constructed by multiplying all combinations of functions $f_i(u)$ and $g_j(v_a, v_b)$. If there are any literature kernels that cannot be factorized as in (25), these could additionally be added to the kernel library.

4. Experimental validation

4.1. Experimental setup

The pharmaceutical wet-granulation line QbCon®1 (L.B. Bohle Maschinen+Verfahren GmbH, Ennigerloh, Germany) implements the granulation process as described above. In contrast to the TSG depicted in Fig. 1, the TSG features two kneading zones, which are treated as one in terms of the PBM. After the granulation step, the machine features a continuous vibrated fluidized bed dryer. Except for the L/S ratio u, the process operated with constant process parameters during experimentation. The tested L/S ratios cover the entire operating range of the process to include the underlying nonlinearities. During realistic operation, the L/S ratio would be varied in a smaller region. Table 1 summarizes the process parameters used in the experiments. The raw material used for granulation is a blend of 72 % Pharmatose®300M (DFE Pharma GmbH & Co. KG, Goch, Germany), 15.3 % Vivapur®102 (JRS PHARMA GmbH & Co. KG, Rosenberg, Germany), 10 % Ibuprofen (Janssen Pharmaceutica N.V., Beerse, Belgium), and 2.7 % Plasdone™K-29/32 (Ashland Industries Deutschland GmbH, Düsseldorf, Germany). The used granulation liquid is demineralized water.

The RTD of the granulation unit was measured using the tracer method. We used the ExtruVis system (MeltPrep, Graz, Austria) for this purpose. The tracer experiment was repeated three times, and the RTD model fitting uses the mean of the measured RTD curves. When fitting the time-delay T_d of the RTD model, we only considered multiples of the sample time T_s.

In the experimental setup, a Parsum®probe IPP 70-S (Parsum GmbH, Chemnitz, Germany) measures the particle size in-line at the outlet of the dryer [18]. The particle ring buffer of the Parsum®probe was configured to hold 3000 particles, the internal sample rate was set to 500 kHz, and the PSD was logged with a sample time $T_s = 1$ s. The Parsum®probe measures the particle diameter d on a logarithmic grid on the interval $d \in [50\,\mu\text{m}, 5000\,\mu\text{m}]$ with $n_v = 20$ classes. With the simplifying assumption of spherical particles, we calculate the particle volume according to $v = \pi d^3/6$.

Table 1
Process parameters used in the experimental setup.

Name	Symbol	Value	Unit
L/S ratio	u	[15, 30]	%
Screw speed	ω_{screw}	100	min^{-1}
Dry-solid mass flow rate	$\dot{m}_{\text{s,in}}$	1	kg h^{-1}
Drying air temperature	T_{air}	65	°C
Drying air volume flow	Q_{air}	25	$\text{N m}^3\,\text{h}^{-1}$
Dryer vibration	a_{vib}	8	m s^{-2}

The particle size measurement $\tilde{N}_i(t)$ contains no information about the absolute number of particles leaving the process. However, under stationary flow conditions, the mass of the outflowing particles remains constant. Thus, it is natural to normalize the particle size measurement to have constant volume:

$$N_i(t) = \frac{\tilde{N}_i(t)}{\sum_{i=1}^{n_v} \tilde{N}_i(t) \cdot v_i}. \tag{26}$$

Finally, the particle size measurement is filtered along the temporal direction with a 5th-order Savitzky–Golay filter with a window length of $T_{\text{filt}} = 61$ s. Data smoothing prevents noise amplification when taking the derivative of the measurement signal. Savitzky–Golay filters are thus often used in the SINDy approach [19]. We also tested other filter window lengths: $T_{\text{filt}} = 31$ s and $T_{\text{filt}} = 121$ s produced the same model structure with only minor variations in the model parameters, showing that the model identification is not sensitive to the noise suppression.

Due to the position of the Parsum®probe, the dryer causes a transport delay in the particle size measurement. For simplicity, particle transport in the dryer is approximated by plug flow. For the selected process parameters, we experimentally determined the residence time to be $T_{\text{dryer}} = 28$ s and compensated the particle size measurement accordingly.

Granulation was performed on the experimental setup in two experiments to generate training and validation data, respectively.

1. **Training data**: The L/S ratio u varies as an amplitude-modulated pseudorandom binary sequence (APRBS) between 15 % and 30 % with a minimum hold time of 5 min. The duration of the first experiment is 60 min.
2. **Test data**: The five operating points $u \in \{25.5\,\%, 16.5\,\%, 22.5\,\%, 28.5\,\%, 19.5\,\%\}$ are each held for 15 min.

We will evaluate the model performance based on the mass-based PSD $q_3(d)$ for its high relevance in many engineering applications. Where practical, we will only consider the 10th, 50th, 90th percentiles of this distribution, which will be denoted by $d_{(10)}$, $d_{(50)}$, and $d_{(90)}$, respectively. As a performance measure, we will use the mean absolute error (MAE).

4.2. Algorithmic setup

The aggregation kernel library is constructed as described in Section 3.4. Table 2 summarizes the functions $f_i(u)$ and $g_j(v_a, v_b)$. The function library consists of $K = 8 \cdot 11 = 88$ candidate kernels. Table 2 contains both canonical kernels and kernels previously proposed in the literature to describe the granulation process. The functions $f_i(u)$ have been chosen monotonically increasing, which encodes the physical knowledge that higher L/S ratios lead to higher particle sizes.

The tuning parameters of Algorithm 1 are the threshold ε and the ridge regression parameter λ. Depending on the choice of the threshold

Table 2
Overview of the functions used to construct the aggregation kernel library.

	$f_i(u)$	$g_j(v_a, v_b)$
1	1	1 [17]
2	$\sqrt{u}$	$v_a^{1/3} + v_b^{1/3}$
3	u	$v_a + v_b$
4	u^2	$v_a^{1/3} \cdot v_b^{1/3}$
5	u^3	$v_a \cdot v_b$
6	u^4	$v_a^{-1/3} + v_b^{-1/3}$
7	$\left(1 + \tanh(\frac{u-0.5}{0.1})\right)/2$	$v_a^{-1} + v_b^{-1}$
8	$\left(1 + \tanh(\frac{u-0.5}{0.3})\right)/2$	$v_a^{-1/3} \cdot v_b^{-1/3}$
9		$v_a^{-1} \cdot v_b^{-1}$
10		$\left(v_a^{2/3} + v_b^{2/3}\right) \cdot \left(v_a^{-1} + v_b^{-1}\right)$ [16]
11		$\left(v_a^{-1/3} + v_b^{-1/3}\right) \cdot \left(v_a^{1/3} + v_b^{1/3}\right)$ [20]

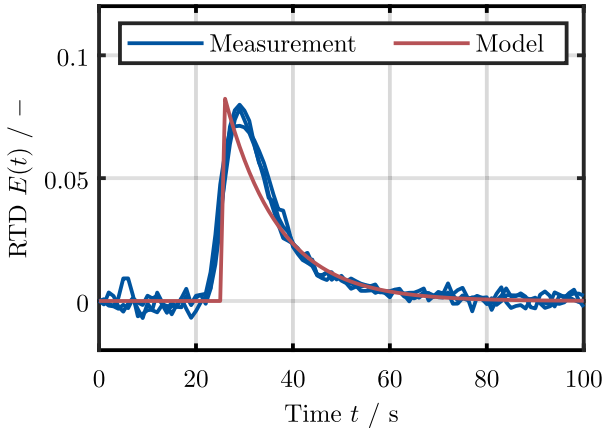


Fig. 3. Residence time distribution measurements compared to the fitted model.

ε, the algorithm chooses different numbers of active terms. In the presented scenario, having more than one active term only decreased the model's residual $\|\boldsymbol{Y} - \boldsymbol{\Theta}\boldsymbol{\xi}\|_{\boldsymbol{W}}^2$ on the training data by about 1 %. This result indicates that one aggregation kernel suffices to describe the system adequately. The choice $\varepsilon = 0.1$ was made because it produces one active aggregation kernel. With only one active term in the parameter vector $\boldsymbol{\xi}$, the risk of overfitting due to correlated columns in $\boldsymbol{\Theta}$ vanishes. We thus chose $\lambda = 0$. Note that in this case, the STRidge algorithm coincides with the STLS algorithm.

5. Results

The results of fitting the RTD model given by (5) to tracer experiment data are shown in Fig. 3. The identified model parameters are $T_d = 25\,\text{s}$ and $T = 11.1\,\text{s}$. The MAE of the model is $3.1 \cdot 10^{-3}$.

Based on the training data, the sparse kernel identification algorithm finds

$$\beta\left(v_a, v_b, u\right) = 0.0597 \cdot u^3 \cdot \left(v_a^{-\frac{1}{3}} + v_b^{-\frac{1}{3}}\right) \cdot \left(v_a^{\frac{1}{3}} + v_b^{\frac{1}{3}}\right). \tag{27}$$

Fig. 4 shows the simulated PBM with the identified aggregation kernel compared to the temporally filtered test data. The MAEs of the 10th, 50th, and 90th percentiles are 93.9 µm (15 %), 134.8 µm (10 %), and 780.2 µm (27 %). For comparison, the identification results for other normalization metrics than the ℓ_∞-based normalization are given in Table 3. This shows that the normalization given by (22) and (23) produces the best results with regard to the mass-based PSD.

Fig. 5 shows the mass-based PSD for the two operating points $u = 16.5\,\%$ and $u = 28.5\,\%$. The depicted PSD is the averaged measurement over the duration of the respective operating point.

Table 3
Identification results for different normalization schemes.

Norm. metric	MAE/µm		
	$d_{(10)}$	$d_{(50)}$	$d_{(90)}$
ℓ_∞	93.9	134.8	780.2
ℓ_2	644.7	1960.9	1324.0
Median	340.3	379.0	1026.6

Table 4
Non-sparse regression results for literature kernels.

Aggr. kernel	$\|\boldsymbol{Y} - \boldsymbol{\Theta}\boldsymbol{\xi}\|_{\boldsymbol{W}}^2$	MAE/µm		
		$d_{(10)}$	$d_{(50)}$	$d_{(90)}$
X. Liu and D. Litster [17]	59.8	94.3	135.2	772.0
Sastry [16]	50.1	93.9	134.8	780.2
Friedlander and Wang [20]	90.6	272.8	357.3	1013.6

To show robustness of the identification process, we divide the training data into five folds, each containing a contiguous 12 min segment. Repeating the identification with the kth fold removed from the training data yields the same kernel $g_{11}(v_a, v_b)$ as before in all five cases. The same input-dependent function $f_5(u) = u^3$ was chosen in four of the five times, while one time the function $f_6(u) = u^4$ was chosen instead.

Finally, non-sparse least-squares regression is performed for the literature aggregation kernels reported in Table 2. For better comparison, based on the sparse regression results, we only consider the input-dependent function $f_5(u) = u^3$ for the non-sparse regression.

6. Discussion

Fig. 3 demonstrates that the RTD predicted by the model is in good agreement with the experimental data. This result gives experimental justification for assumptions A3 and A4. The rising edge of the modeled RTD is infinitely sharp, whereas the measured PSD exhibits a smoother increase. This deviation can be attributed to the neglect of axial dispersion in the screw zone (A3). Nevertheless, A3 is necessary to derive the simplified PBM given by (4). Unfortunately, since the particle size cannot be measured inside the TSG, a validation of assumptions A3 and A4 beyond the RTD is not possible.

Ignoring dispersion and aggregation in the transport zone will likely lead to an overestimation of the aggregation rate constant, since the effective residence time is increased. In terms of the input–output behavior, however, these effects cancel out. Consequently, there are no restrictions on the applicability of the model in control applications. To verify this experimentally, we repeated the identification with an assumed $T = 13.3\,\text{s}$ (20 % increase). The algorithm identified the same model structure with the aggregation rate constant reduced by 8.3 %. The MAEs of the percentiles $d_{(10)}$, $d_{(50)}$, and $d_{(90)}$ were not effected by the change. This exemplifies that the identified aggregation rate constant should be interpreted as an effective value, which absorbs the modeling assumptions A2-A4. The model structure, on the other hand, does not seem to be sensitive to the RTD parameters.

The PSD measurements are inherently noisy due to the stochastic nature of the process and the measurement principle. The relatively low signal-to-noise ratio makes it difficult to quantify the prediction performance of the model. The Parsum®probe computes the PSD on the basis of 3000 particles in its ring buffer. Increasing the buffer size reduces noise but introduces a delay because the buffer must be fully replenished before the signal reflects new process conditions. With an average measurement rate of 268 particles/s, the ring buffer gets fully replaced after approximately 11 s. Additional smoothing is achieved through a Savitzky–Golay filter. The window size of the filter is several times longer than the replacement time of the Parsum probe, to mitigate the influence of particularly rare, i.e., large particles. These particles may not be adequately represented in the particle buffer,

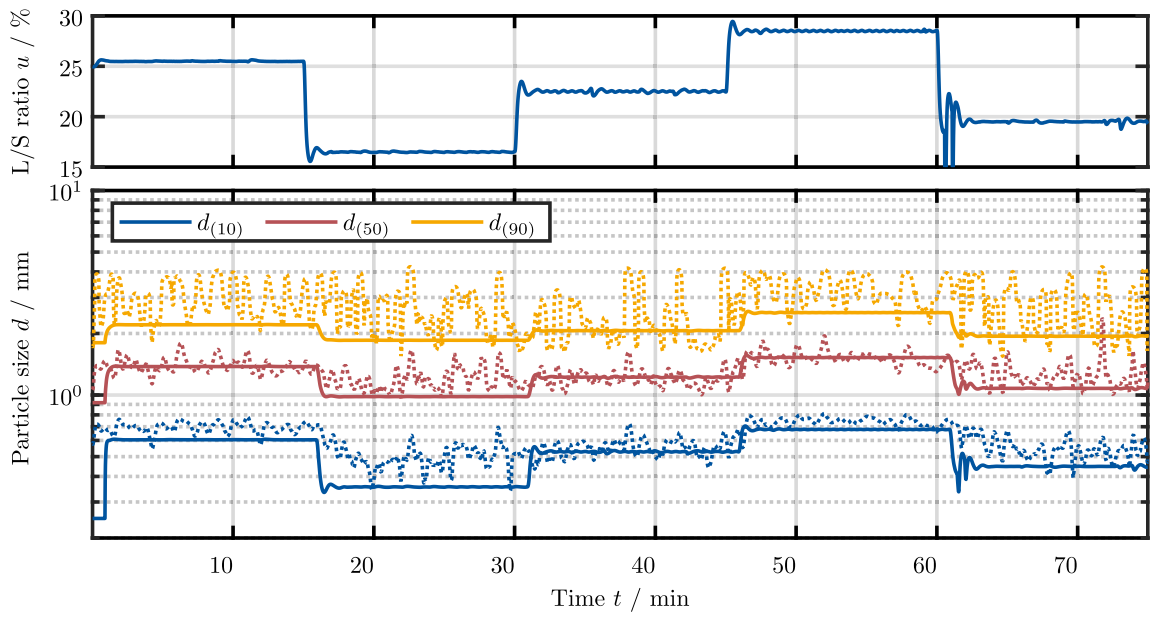
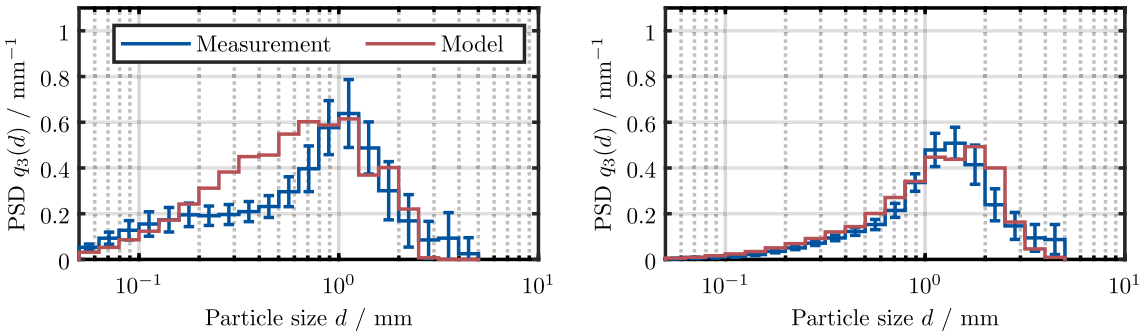


Fig. 4. Model prediction (solid) and filtered measurement (dotted) of the percentiles $d_{(10)}$, $d_{(50)}$, and $d_{(50)}$ of the mass-based particle size distribution.



(a) $u = 16.5\,\%$. (b) $u = 28.5\,\%$.

Fig. 5. Average measured and predicted mass-based particle size distribution.

which makes additional smoothing necessary. Despite this, the signal $d_{(90)}$ in Fig. 4 still exhibits significant noise. This demonstrates the effect of large particles, which, although less frequent, contribute a disproportionately large share of the total mass. Consequently, the upper tail of the mass-based PSD is particularly susceptible to low-frequency, high-amplitude fluctuations.

Despite the measurement noise – also present in the training data – the kernel identification algorithm succeeds in capturing both the overall shape of the PSD (Fig. 5) and its response to changes in the L/S ratio u (Fig. 4). A comparison between Figs. 5(a) and 5(b) shows how higher L/S ratios reduce the fraction of small particles. The model exhibits the largest deviation from the measurement for low L/S ratios in the medium particle size range, as seen in Fig. 5(a). This leads to a slight underestimation of $d_{(50)}$. One reason for the PSD shape mismatch in the low L/S ratio range might be the omission of particle breakage according to assumption A2. Ismail et al. [21] reports that for low L/S ratios, particle breakage plays a more pronounced role than at higher L/S ratios.

The identified aggregation kernel (27) reveals a strongly nonlinear relationship between the L/S ratio and the coalescence rate. For low L/S ratios, the aggregation rate is low and insensitive to further changes, whereas for high L/S ratios, it increases sharply. This behavior is also clearly seen in the measurement in Fig. 4 and it aligns well with the theoretical analysis of wet granulation reported by Liu et al. [22].

The algorithm identified Brownian coagulation [20] as the most suitable mechanism to describe the data. This suggests that random particle motion is the dominant driver of coalescence under the investigated conditions. However, since the true physical mechanism is not known, this interpretation must be treated with caution. In fact, the non-sparse identification results in Table 4 reveal that the constant aggregation kernel from [17] is similarly well suited to describe the data. In the presented wet-granulation case, which is already widely covered in literature, the proposed algorithm does not discover a new aggregation mechanism but instead finds the best-suited aggregation kernel from literature. The main benefit the algorithm provides in this case is the circumvention of a computationally expensive exhaustive search of all model structures.

Another important limitation of the proposed approach is that not all aggregation kernels found in the literature can be directly used within the kernel library. Specifically, only kernels that are linearly parametrized can be used within the SINDy framework. For example, Kapur [23] proposes a kernel, where the exponent of the kernel is an adjustable parameter. In this case, one might need to use a two-step approach, by first identifying the unknown exponents and then adding the kernel to the library.

7. Outlook

Due to the novelty of the proposed approach, the scope of this work was deliberately restricted to the variation of a single process parameter and the consideration of pure aggregation. Extending the framework to multiple input variables is conceptually straightforward but increases both the number of operating points required for training data generation and the size of the candidate function library. Fortunately, the SINDy algorithm scales efficiently with an increasing number of candidate functions. Moreover, the methodological advances introduced in this paper can be readily adapted to enable the identification of breakage kernels, which were neglected in the present study, but are expected to further enhance the prediction.

Finally, the model provides a promising foundation for model predictive control. Leveraging the identified PBM within a control framework could enable real-time adaptation of operating conditions and improved product quality in continuous granulation processes.

CRediT authorship contribution statement

Stefan R. Tölle: Writing – original draft, Visualization, Software, Methodology, Investigation, Conceptualization. **Lorenz Dörschel:** Writing – original draft, Methodology. **Stefan Klinken-Uth:** Writing – review & editing, Investigation, Conceptualization. **Alana Delvos:** Writing – review & editing, Investigation, Data curation. **Jörg Breitkreutz:** Supervision, Project administration. **Heike Vallery:** Writing – review & editing, Supervision, Project administration. **Sebastian Stemmler:** Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Institut of Pharmaceutics and Biopharmaceutics reports equipment, drugs, or supplies was provided by LB Bohle Machines and Processes GmbH. Institut of Pharmaceutics and Biopharmaceutics reports equipment, drugs, or supplies was provided by DFE Pharma. Institut of Pharmaceutics and Biopharmaceutics reports equipment, drugs, or supplies was provided by JRS Pharma LP. Institut of Pharmaceutics and Biopharmaceutics reports equipment, drugs, or supplies was provided by Janssen Pharmaceutica NV. Institut of Pharmaceutics and Biopharmaceutics reports equipment, drugs, or supplies was provided by Ashland Industries Deutschland GmbH. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The measurement data and software needed to reproduce the above findings are freely available in the accompanying data publication https://doi.org/10.18154/RWTH-2026-01122.

References

[1] U. Nandi, V. Trivedi, S.A. Ross, D. Douroumis, Advances in twin-screw granulation processing, Pharmaceutics 13 (5) (2021) http://dx.doi.org/10.3390/pharmaceutics13050624.

[2] T. Glaser, C.F. Sanders, Fu.Y. Wang, I.T. Cameron, J.D. Litster, J.M.-H. Poon, R. Ramachandran, C.D. Immanuel, F.J. Doyle, Model predictive control of continuous drum granulation, J. Process Control 19 (4) (2009) 615–622, http://dx.doi.org/10.1016/j.jprocont.2008.09.001.

[3] T. Shaqarin, A. Al-Rawajfeh, M. Hajaya, N. Alshabatat, B. Noack, Model-based robust H∞ control of a granulation process using Smith predictor with reference updating, J. Process. Control. 77 (2019) 38–47, http://dx.doi.org/10.1016/j.jprocont.2019.03.003.

[4] D. Ramkrishna, M.R. Singh, Population balance modeling: Current status and future prospects, Annu. Rev. Chem. Biomol. Eng. 5 (1) (2014) 123–146, http://dx.doi.org/10.1146/annurev-chembioeng-060713-040241.

[5] A. Chaudhury, H. Wu, M. Khan, R. Ramachandran, A mechanistic population balance model for granulation processes: Effect of process and formulation parameters, Chem. Eng. Sci. 107 (2014) 76–92, http://dx.doi.org/10.1016/j.ces.2013.11.031.

[6] R. Ramachandran, P.I. Barton, Effective parameter estimation within a multi-dimensional population balance model framework, Chem. Eng. Sci. 65 (16) (2010) 4884–4893, http://dx.doi.org/10.1016/j.ces.2010.05.039.

[7] X. Chen, L.G. Wang, F. Meng, Z.-H. Luo, Physics-informed deep learning for modelling particle aggregation and breakage processes, Chem. Eng. J. 426 (2021) 131220, http://dx.doi.org/10.1016/j.cej.2021.131220.

[8] J. Chakraborty, J. Kumar, M. Singh, A. Mahoney, D. Ramkrishna, Inverse problems in population balances. Determination of aggregation kernel by weighted residuals, Ind. Eng. Chem. Res. 54 (42) (2015) 10530–10538, http://dx.doi.org/10.1021/acs.iecr.5b01368.

[9] S.I.X. Tiong, F. Ahamed, Y.K. Ho, Data-driven discovery of population balance equations for the particulate sciences [Preprint], 2025, http://dx.doi.org/10.48550/ARXIV.2502.09010.

[10] S.L. Brunton, J.L. Proctor, J.N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proc. Natl. Acad. Sci. USA 113 (15) (2016) 3932–3937, http://dx.doi.org/10.1073/pnas.1517384113.

[11] E. Kaiser, J.N. Kutz, S.L. Brunton, Sparse identification of nonlinear dynamics for model predictive control in the low-data limit, Proc. R. Soc. A 474 (2219) (2018) 20180335, http://dx.doi.org/10.1098/rspa.2018.0335.

[12] T. Chan Seem, N.A. Rowson, I. Gabbott, M. De Matas, G.K. Reynolds, A. Ingram, Asymmetric distribution in twin screw granulation, Eur. J. Pharm. Biopharm. 106 (2016) 50–58, http://dx.doi.org/10.1016/j.ejpb.2016.01.013.

[13] D. Ramkrishna, Population Balances: Theory and Applications to Particulate Systems in Engineering, Academic Press, Amsterdam, 2000.

[14] S. Kumar, D. Ramkrishna, On the solution of population balance equations by discretization – I. A fixed pivot technique, Chem. Eng. Sci. 51 (8) (1996) 1311–1332, http://dx.doi.org/10.1016/0009-2509(96)88489-2.

[15] S.H. Rudy, S.L. Brunton, J.L. Proctor, J.N. Kutz, Data-driven discovery of partial differential equations, Sci. Adv. 3 (4) (2017) e1602614, http://dx.doi.org/10.1126/sciadv.1602614.

[16] K.V. Sastry, Similarity size distribution of agglomerates during their growth by coalescence in granulation or green pelletization, Int. J. Miner. Process. 2 (2) (1975) 187–203, http://dx.doi.org/10.1016/0301-7516(75)90021-6.

[17] L. X. Liu, J. D. Litster, Population balance modelling of granulation with a physically based coalescence kernel, Chem. Eng. Sci. 57 (12) (2002) 2183–2191, http://dx.doi.org/10.1016/S0009-2509(02)00110-0.

[18] D. Petrak, S. Dietrich, G. Eckardt, M. Köhler, In-line particle sizing for real-time process control by fibre-optical spatial filtering technique (SFT), Adv. Powder Technol. 22 (2) (2011) 203–208, http://dx.doi.org/10.1016/j.apt.2010.11.002.

[19] B.M. de Silva, K. Champion, M. Quade, J.-C. Loiseau, J.N. Kutz, S.L. Brunton, PySINDy: A Python package for the sparse identification of nonlinear dynamics from data, 2020, http://dx.doi.org/10.48550/ARXIV.2004.08424.

[20] S. Friedlander, C. Wang, The self-preserving particle size distribution for coagulation by brownian motion, J. Colloid Interface Sci. 22 (2) (1966) 126–132, http://dx.doi.org/10.1016/0021-9797(66)90073-7.

[21] H.Y. Ismail, M. Singh, A.B. Albadarin, G.M. Walker, Complete two dimensional population balance modelling of wet granulation in twin screw, Int. J. Pharm. 591 (2020) 120018, http://dx.doi.org/10.1016/j.ijpharm.2020.120018.

[22] L.X. Liu, J.D. Litster, S.M. Iveson, B.J. Ennis, Coalescence of deformable granules in wet granulation processes, AIChE J. 46 (3) (2000) 529–539, http://dx.doi.org/10.1002/aic.690460312.

[23] P. Kapur, Kinetics of granulation by non-random coalescence mechanism, Chem. Eng. Sci. 27 (10) (1972) 1863–1869, http://dx.doi.org/10.1016/0009-2509(72)85048-6.