Journal Pre-proof

Control oriented modeling of twin-screw granulation in the ConsiGma™-25 production plant

Selma Celikovic, Johannes Poms, Johannes Khinast, Martin Horn, Jakob Rehrl

PII: S0378-5173(23)00458-1
DOI: https://doi.org/10.1016/j.ijpharm.2023.123038
Reference: IJP 123038

To appear in: *International Journal of Pharmaceutics*

Received date : 2 February 2023
Revised date : 18 April 2023
Accepted date : 5 May 2023

Please cite this article as: S. Celikovic, J. Poms, J. Khinast et al., Control oriented modeling of twin-screw granulation in the ConsiGma™-25 production plant. *International Journal of Pharmaceutics* (2023), doi: https://doi.org/10.1016/j.ijpharm.2023.123038.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Journal Pre-proof

Control Oriented Modeling of Twin-Screw Granulation in the ConsiGma™-25 Production Plant

Selma Celikovic[a,b], Johannes Poms[a], Johannes Khinast[a,c], Martin Horn[b], Jakob Rehrl[a,*]

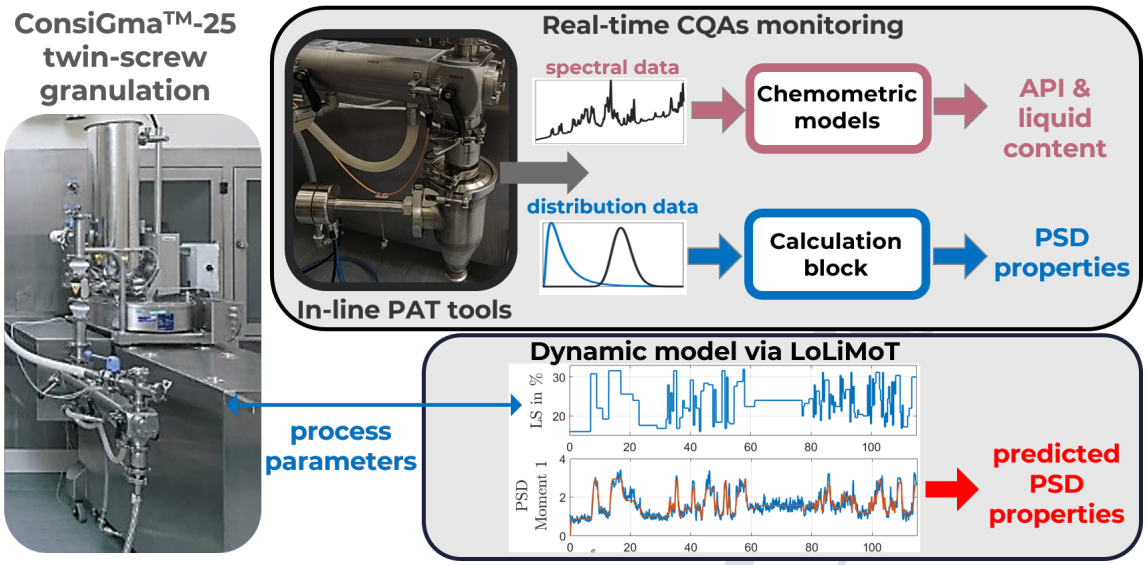
[a] *Research Center Pharmaceutical Engineering GmbH, Inffeldgasse 13/2, 8010 Graz, Austria*
[b] *Institute of Automation and Control, Graz University of Technology, Inffeldgasse 21b, 8010 Graz, Austria*
[c] *Institute for Particle and Process Engineering, Graz University of Technology, Inffeldgasse 13/III, 8010 Graz, Austria*

Abstract

ConsiGma™-25 is a continuous production plant integrating a twin-screw granulation, fluid bed drying, granule conditioning, and a tableting unit. The particle size distribution (PSD), active pharmaceutical ingredient (API) content, and liquid content of wet granules after twin-screw granulation affect the quality of intermediate and final products. This paper proposes methods for real-time monitoring of these quantities and control-oriented modeling of the granulator.
The PSD of wet granules is monitored via an in-line process analytical technology (PAT) probe based on the spatial velocimetry principle. The algorithm for signal processing and evaluation of PSD characteristics is developed and applied to the acquired PSD data. A dynamic process model predicting PSD characteristics from granulation parameters is trained via the local linear model tree (LoLiMoT) approach. The experimental data required for the model training are collected via systematically designed excitation runs. Finally, the performance of the identified model is examined and verified by means of a new set of validation runs. Furthermore, an in-line PAT probe based on Raman spectroscopy is developed and integrated after the granulator. The API- and liquid content of produced wet granules are evaluated from the spectral data by means of chemometric modeling, and chemometric models are validated on a separate set of experimental data. The solutions proposed in this research can be used as a reliable (and necessary) basis for the development of advanced quality-by-design control concepts (e.g., PSD process control). Such concepts would ultimately improve the ConsiGma™-25 process performance in terms of robustness against disturbances and quality of intermediate and final products.

Keywords: ConsiGma™-25, Twin Screw Granulation, Continuous Manufacturing, Data Driven Process Modeling, Local Linear Model Tree (LoLiMoT), Control Oriented Modeling, Process Analytical Technology (PAT)

*Corresponding author: jakob.rehrl@rcpe.at



1. Introduction

The pharmaceutical industry nowadays is in transition from standard batch-based to continuous manufacturing. The continuous manufacturing of pharmaceuticals comes with a range of benefits, such as increased production flexibility, better quality control, reduced energy consumption, and lower environmental footprint via waste reduction (Lee et al., 2015).

The ConsiGma^{TM}-25 is a well-known continuous manufacturing plant integrating a twin-screw granulator (TSG), a fluid-bed-dryer (FBD), a granule-conditioning unit (GCU), and a tablet press (TP). The common ConsiGma^{TM}-25 operation mode, i.e., the operation with empirically determined, constant process parameters and with only limited real-time monitoring of intermediate/final critical quality attributes (CQA), does not fully exploit the benefits of continuous manufacturing. Potential disturbances (equipment faults, material variability, or operator mistakes) lead to quality degradation of intermediate/final products. Thus, the process performance could be improved by means of the pharmaceutical quality-by-design (QbD) approach (ICH, 2009, 2005, 2008, 2012), specifically via the implementation of quality and process control concepts (Yu et al., 2014). There are some essential requirements for the development of model based control concepts: real-time monitoring of CQAs via process analytical technology (PAT) tools, and the process models linking the critical process parameters (CPP) to CQAs. The performance of the ConsiGma^{TM}-25 twin-screw wet granulation unit is significantly affecting the final product quality: Wet granulation is a size enlargement process that improves the flowability properties of raw material, reduces the risk of segregation, and improves content uniformity (Seem et al., 2015). The particle size distribution (PSD) of wet granules is considered an intermediate CQA affecting the final product quality (e.g., tablet dissolution (Markl et al., 2020; Zaborenko et al., 2019)), and the performance of the subsequent unit operations (e.g., filter clogging in the FBD due to formation of

fines). The active pharmaceutical ingredient (API) content of the wet granules is another intermediate CQA. A long-term out-of-specification (OOS) API content of wet granules due to material segregation or equipment faults would be reflected in the final product, i.e., tablet quality. Therefore, the paper at hand focuses on the solutions required for the real-time monitoring and control of PSD and API content (i.e., PAT equipment and process models).

An approach for the real-time monitoring and control of the granule size after the twin-screw granulation is proposed in (Nicolaï, 2019). This study raises the issue of the nonlinear behavior of the granulation process claiming that the granule size controller should consider this system property when choosing the correct control action. Linear controllers, e.g., PID controllers, therefore cannot fully exploit the potential of the granulation process. However, if the process model capturing the nonlinear behavior of the investigated system would be available, more sophisticated, nonlinear model-based control concepts could be developed. Although several modeling approaches for twin-screw granulation can be found in the literature (Barrasso et al., 2014, 2013; Barrasso and Ramachandran, 2015), they are typically not directly applicable to the design of model-based control strategies. Advanced modeling approaches, such as gPROMS (gProms, 2023), offer a possibility for in-silico process investigation via digital twin and flowsheet modeling (Wang et al., 2022; Metta et al., 2019). Although very promising for sensitivity analysis, these models are not yet suitable to capture the dynamic behavior of twin screw granulation. If such a feature is available, the modeling procedure could be transferred from the ConsiGma^{TM}-25 line to the simulation environment. However, due to computational demand, it is not possible to directly apply these models for model predictive control design. In contrast, (Shirazian et al., 2017) proposes a modeling approach based on an artificial neural network (ANN) algorithm for a static prediction of PSD characteristics for different operating conditions based on experimental data. Although the promising results presented in this work indicate the potential of the data-driven ANN approach (in terms of accuracy and computation time), they do not consider further important prerequisites, such as the dynamic behavior of the system or the real-time measurement of critical quantities.

The paper at hand closes the existing gaps by providing a systematic framework for developing a dynamic process model for granule size prediction that is suitable for model predictive control (MPC) concepts. It can be used as a step-by-step guide through the development of a PAT strategy, definition of model structure (model inputs and outputs), systematic design of excitation runs (collecting experimental data that accurately reflect the dynamic behavior of the system), signal processing, and model training. In order to monitor the PSD of wet granules in real-time, an in-line PAT probe based on the spatial filter velocimetry principle is mounted at the TSG outlet. The relevant PSD characteristics are extracted from acquired distribution data and the respective CPPs are identified. Furthermore, a process model describing the relation between the granulation CPPs and PSD characteristics is developed by means of a local-linear model tree (LoLiMoT) algorithm (Nelles, 1997). For control design purposes, LoLiMoT is a powerful data-driven alternative to physically motivated modeling approaches suggested in

Journal Pre-proof

(Barrasso et al., 2014, 2013; Barrasso and Ramachandran, 2015). Although this model identification approach is tailored to a specific pharmaceutical model formulation, the proposed method for the design of excitation runs is generally applicable and can be quickly adapted to different formulations. As shown in (Rehrl et al., 2019), process models based on the LoLiMoT approach allow a straightforward development of advanced model-based process control concepts, e.g., MPC. A nonlinear MPC integrating the process model identified in this work would not encounter the nonlinearity issue raised in (Nicolaï, 2019). To capture the remaining CQAs of the wet granulation, this paper proposes an in-line PAT solution based on Raman spectroscopy, as well as the appropriate chemometric modeling approach for extracting API content information from acquired spectral data. In addition to the API-, the liquid content of wet granules is extracted from the spectral data by means of chemometric modeling. This could be an alternative to the approach for the real-time monitoring of liquid content after the granulation via NIR proposed in (Nicolaï et al., 2018).
The performance of all the solutions developed in this study is confirmed via separate sets of validation experiments. Finally, based on these solutions, typical applications focusing on the quality/process control for the ConsiGma^{TM}-25 production plant are proposed.

2. Materials and Methods

2.1. Materials

Wet granules were produced from a powder pre-blend and deionized water. The pre-blend material consisted of Methyl 4-hydroxybenzoate (Sigma Aldrich, USA), also known as Methylparaben, as the API surrogate material, and three excipients, i.e., VIVAPHARM HPMC (Demacsa, Mexico), Avicel PH101 (DuPont, Ireland), and Granulac 200 (Meggle, Germany). Table 1 shows the nominal pre-blend composition.

Table 1: Formulation of pre-blend material.

Nr.	Raw Materials	Quantity (wt. %)
1	Methyl 4-hydroxybenzoate API surrogate	4.12
2	VIVAPHARM HPMC	5.15
3	Avicel PH101	21.41
4	Granulac 200	69.32

2.2. Granule size modeling

2.2.1. Technical process description

The loss-in-weight feeder K-TRON KT 20 (Coperion K-Tron, Switzerland, (Coperion, 2023)) was placed at the inlet of the TSG of the ConsiGma^{TM}-25 plant (GEA, Belgium, (GEA, 2023)). It was filled with the pre-blend material which was fed at a

Journal Pre-proof

nominal solid feed rate (SFR) of 15 kg/h. A liquid tank stored deionized water, which was supplied to the TSG by means of a mass flow controlled peristaltic pump at a liquid feed rate (LFR) of 60 g/min, resulting in the nominal liquid-to-solid ratio (LS) of 24 %. The TSG with a screw diameter of 20 mm and a length-to-diameter ratio of 20:1 was operated with the nominal screw speed (SS) of 700 rpm. The chosen screw configuration was: 1K/6 / 2x1T / 2x1,5T / 4x2T / 6K/4 60° / 1x1,5T / 6K/4 60° / 1x1,5T / 2K/6 60°, with K representing kneading- and T representing transport elements (e.g., 2K/6 60° stands for two kneading elements shifted by 60°, with a length of 1/6 of their diameter). The nominal barrel temperature (BT) of 30° C was controlled by a PID controller. The pre-blend material was conveyed among the granulator screws, the liquid was distributed, and wet granules were produced. Figure 1 shows a schematic diagram of the investigated granulation unit.

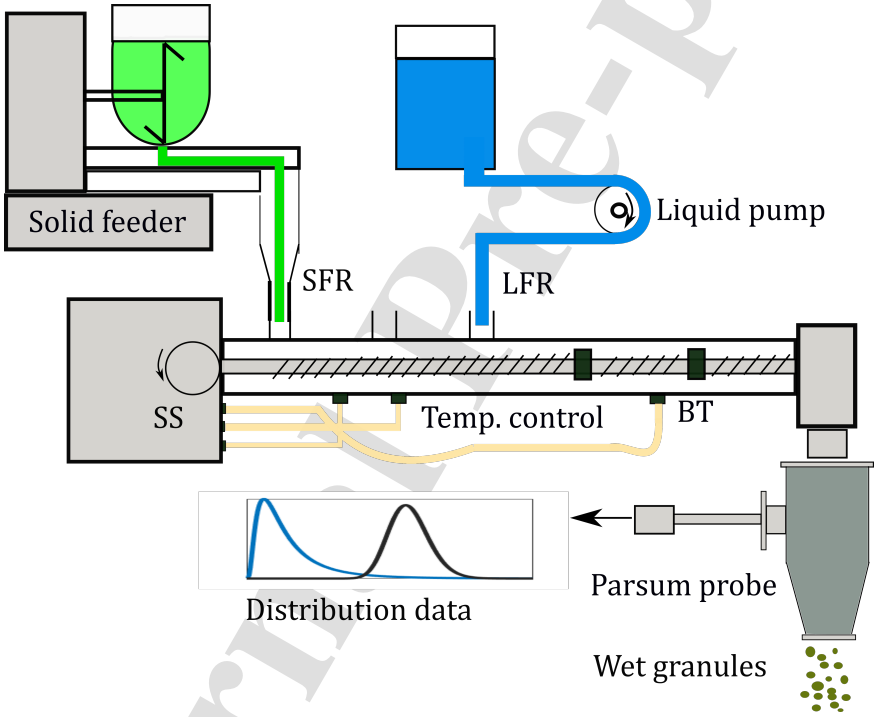


Figure 1: Schematic representation of the ConsiGma™-25 granulation unit with the Parsum probe monitoring the size of wet granules. Legend: Solid feed rate (SFR), liquid feed rate (LFR), screw speed (SS), barrel temperature (BT).

2.2.2. PAT strategy

The size of the wet granules was captured in-line by means of an IPP 80-P inline particle probe (Parsum GmbH, Germany, (Parsum, 2023)) mounted at the TSG outlet. The probe uses the spatial filter velocimetry principle. In brief, in the measurement volume a laser diode emits a beam sensed by an array of optical fibres connected to individual photo-detectors. Particles passing through the volume therefore cast a shadow on the detectors, generating a signal on the whole array whose frequency is proportional

Journal Pre-proof

to the velocity. The time of flight is determined by the time a single photo-detector is blocked by the shadow of a moving particle. From these two signals, the chord length distribution and subsequently the PSD is calculated by the instrument software IPP V9.00 (Dieter et al., 2011; Silva et al., 2013).
The IPP 80-P probe was mounted directly in the product stream at the granulator outlet with the help of a custom built mechanical interface, depicted in Figure 2. The design consists of a Ø100 mm cylindrical tube intersected with a Ø80 mm tube sideways, with welded tri-clamp flanges, and a conical shape toward the bottom. The stainless steel interface can be opened for inspection and is easy to clean. The D12 disperser unit with an inlet opening of Ø7.5 mm that includes a teflon coated cap for less adhesion of the wet granules, was attached to the probe. An air unit constantly supplies the disperser and cleans the optical windows with an air flow set to 6 and 30 l/min, respectively. Additional purge air pulses are used every 8 s to remove granules if they got stuck or block the disperser inlet. An additional valve is mounted on top of the TSG for pressure compensation, such that the internal ConsiGma^{TM}-25 pressure sensor does not trigger a false alarm due to the pulses.
The instrument software was set to a ring buffer of 5000 particles for fast response, using the whole measurement range from 50 to 6000 μm and an acquisition rate of 1s. The software outputs are the number-density and volume-density distribution of the PSD, as well as the velocity distribution, aspect ratio, and particle rate (the number of particles per second passing the probe). In this work a volumetric density distribution $q_3(x)$, given for a size array x of $n_{dist} = 36$ size fractions, was used. All process data were centrally stored and made available on the SIMATIC SIPAT 5.1.1.0 platform (Siemens AG, Germany, (Siemens, 2023)). Methods and collectors were defined in SIPAT to interface with the Parsum OPC-DA server (Parsum GmbH, Germany, (Parsum, 2023)) and the iFix OPC-DA (GE Digital, USA, (iFix, 2023)) for the ConsiGma^{TM}-25 SCADA/HMI system.

2.2.3. Particle size distribution (PSD) characteristics to be modeled

In order to use the PSD information captured via the Parsum probe for modeling or control purposes, the relevant scalar characteristics should be extracted from measured distribution data. These PSD characteristics should reflect the physical properties of the produced granules and also be available in real-time (important for application in a feedback control concept). Following that idea, the four statistic moments (M_1, M_2, M_3 and M_4) are defined as PSD characteristics (Stieß, 2008; Ramsey et al., 2002). The connotation of individual moments is explained by an example with three arbitrary distributions depicted in Figure 3.

PSD-Moment 1. The first moment (mean/expected value) of a distribution is defined as

$$M_1 = \int_{-\infty}^{\infty} q_3(x) x \, dx. \tag{1}$$

An increase in the average size of produced granules is reflected by an increase of M_1.

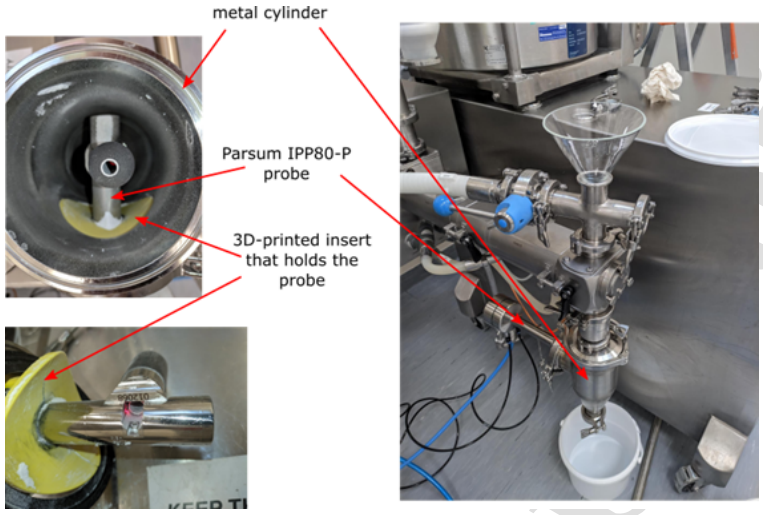


Figure 2: Mechanical interface integrating the Parsum IPP 80-P inline particle measuring probe mounted at the granulator outlet (CAPRI, 2023a).

PSD-Moment 2. The second central moment (variance) of a distribution is defined as

$$M_2 = \int_{-\infty}^{\infty} q_3(x)(x - M_1)^2 \, dx. \tag{2}$$

M_2 captures the distribution broadness, i.e., the higher M_2 gets, the higher the possibility to find the granules further away from the mean value M_1.

PSD-Moment 3. The third standardized moment (skewness) of a distribution is defined as

$$M_3 = \int_{-\infty}^{\infty} q_3(x) \left(\frac{x - M_1}{\sqrt{M_2}} \right)^3 dx. \tag{3}$$

M_3 provides information on the distribution shape, i.e, the relative size of the distribution tails. A negative value of M_3 implies a longer distribution tail on the left, and a positive value of M_3 implies a longer tail on the right side of the observed distribution.

PSD-Moment 4. The fourth standardized moment (kurtosis) of a distribution is defined as

$$M_4 = \int_{-\infty}^{\infty} q_3(x) \left(\frac{x - M_1}{\sqrt{M_2}} \right)^4 dx. \tag{4}$$

Similar to M_3, M_4 also provides information on the distribution shape. It is a measure of the overall tail extremity, without giving information on which side the effect is more pronounced.
In addition to the statistic moments, the following deviation measures are defined. Again, the meaning of individual measures is illustrated with the help of three distribution examples in Figure 3.

Journal Pre-proof

Reference Deviation. The first deviation measure represents the deviation from an arbitrary reference distribution q_{3ref} and is calculated as

$$e_{ref} = (M_1 - M_{1ref}) \sum_{i=1}^{n_{dist}} |q_{3,i} - q_{3ref,i}|. \tag{5}$$

This measure is a suitable choice for implementing a PSD control loop: A distribution that yields the best final product properties and that results in a robust process performance can be selected as the reference distribution. The control loop would then aim at driving e_{ref} to zero. For the model development purpose, an arbitrary Gaussian distribution with a mean value of 1555 μm and a standard deviation of 900 μm is chosen as a reference distribution (for this reference distribution, e_{ref} approximately equals zero at nominal process parameters).

Normal Deviation. The second deviation measure is calculated as

$$e_{normal} = \sum_{i=1}^{n_{dist}} \left(\frac{q_{3,i} - q_{3normal,i}(M_1, M_2)}{q_{3normal,i}(M_1, M_2)} \right)^2 \tag{6}$$

and it represents the deviation from the Gaussian distribution with the equivalent first and second characteristic moments. This quantity is relevant for the development of the signal processing concept. However, it is not relevant for the modeling as it cannot be directly correlated to the specific physical properties of the granules.

PSD signal processing. The cleaning of the Parsum PAT probe (for more details please refer to Section 2.2.2) sometimes results in corrupted PSD measurements. The corrupted PSD data provide no useful information and could compromise the process modeling. The developed PSD signal processing algorithm detects and replaces the corrupted PSD data in the following way: The current value of normal deviation e_{normal} is compared to the mean value over a time window $n_{filter} = 20$, and the respective difference is calculated. If the calculated difference exceeds the threshold of 2.5 standard deviations over the same time window, the measured PSD is detected as corrupted. The corrupted PSD is replaced with the mean value over the same time window. Figure 4 depicts the performance of this concept. Furthermore, in order to improve the signal-to-noise ratio (SNR), all introduced PSD characteristics are low-pass filtered via the following difference equation

$$y_{filt,k} = 0.15 y_k + 0.85 y_{filt,k-1} \tag{7}$$

with y_k representing the evaluated and $y_{filt,k}$ representing the filtered value of the signal at the k-th time instant. The filter coefficients are chosen as a compromise between the achieved SNR and the time delay brought into the system by filtering.

2.2.4. Definition of model structure

The aim of modeling is to develop a simple, yet comprehensive description of the system of interest. The process parameters that have an influence on the product quality

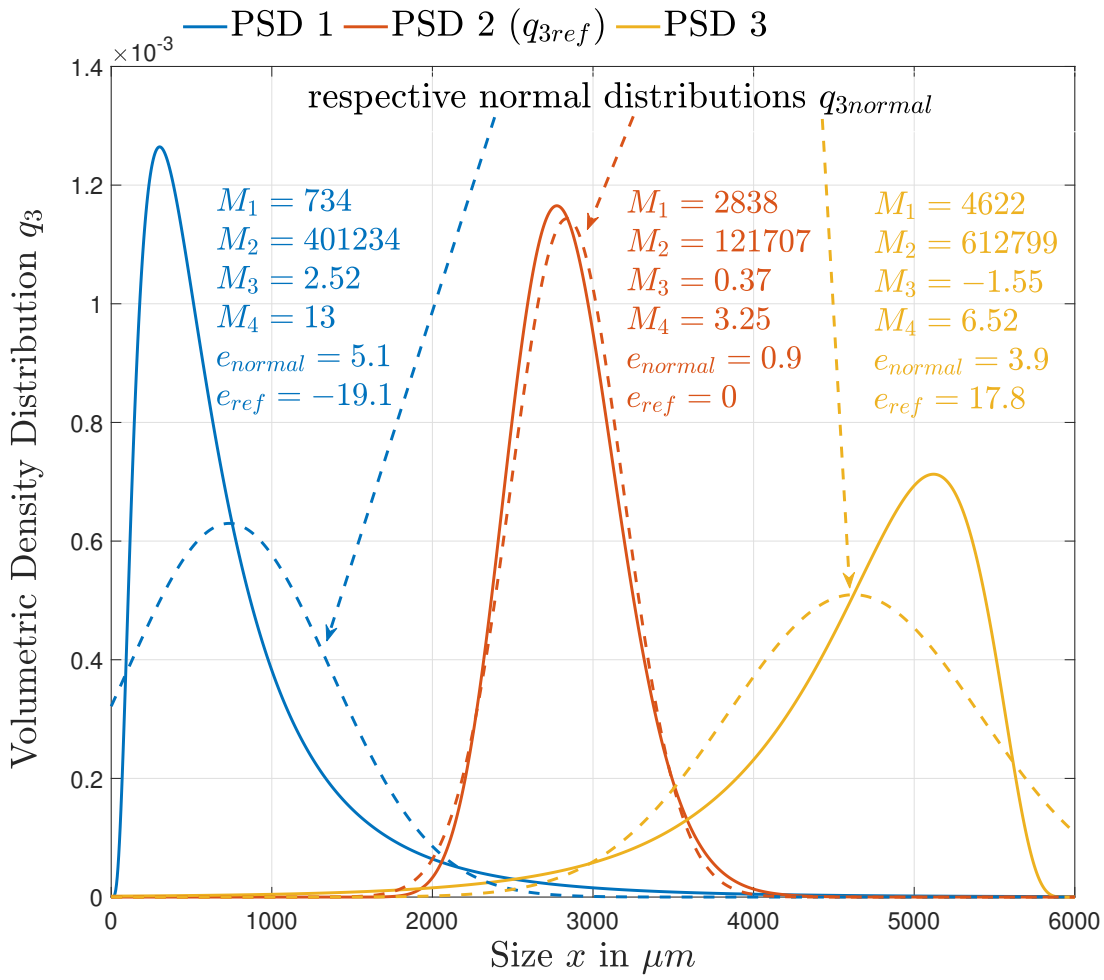


Figure 3: PSD characteristics illustrated by three arbitrary distributions: A distribution shift from left (PSD 1) to the right (PSD 3) correlates with an increase of the respective first moment M_1, i.e., with an increase in the average size of the produced granules. The low M_2 value of the narrow second distribution indicates the low chance of finding the granules further away from the calculated M_1. The first distribution with the long tail on the right implies an overall higher amount of agglomerates than fines (reflected in the positive sign of M_3). The opposite can be stated for the third distribution characterized by the negative M_3 value. The lowest tendency toward granule size outliers of the second distribution is correlated with the lowest M_4 value (compared to PSD 1 and PSD 3), and vice versa. The second distribution is similar to a Gaussian and exhibits a low e_{normal} value. The first and third distributions show higher, yet similar deviation, again reflected in the respective e_{normal} values. To illustrate the meaning of the reference deviation, PSD 2 is chosen as the reference distribution. Positive/negative e_{ref} values imply the production of larger/smaller granules than for the reference distribution. The total extent of the deviation is reflected in the absolute value of the reference deviation.

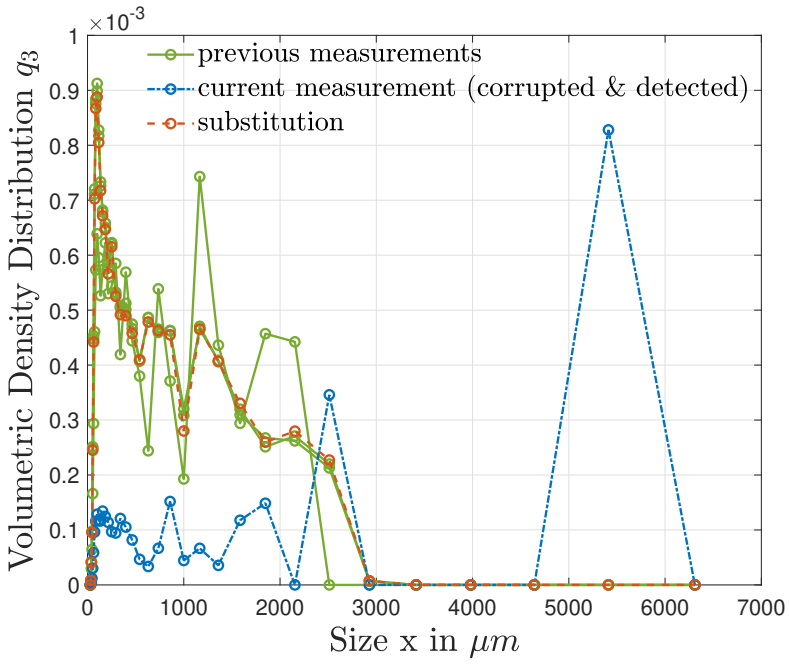


Figure 4: The developed PSD signal processing concept detects and replaces a corrupted measurement.

attributes are determined and considered as model inputs. In control-oriented modeling, the model inputs can be perceived as process actuators (i.e., can be externally manipulated in real-time). Real-time monitored quantities that reflect the product quality attributes are considered as model outputs.

In order to select the model structure, a systematic investigation of the influence of granulation process parameters on the granule size was performed. Solid feed rate, liquid-to-solid (LS) ratio, screw speed, and barrel temperature were considered to be potential model inputs, and the PSD characteristics introduced in Section 2.2.3 were considered to be potential model outputs. A design of experiments (DoE) was performed using the Plackett Burman screening design in Modde DoE software (Sartorius, Germany, (Sartorius, 2023a)). This DoE is shown in Table 2 and it involved eleven runs with simultaneous changes in granulation process parameters. The duration of individual runs was adjusted during the experiments, such that both model inputs and outputs reach a steady state. Obtained results indicate a strong relationship between the granule size (more specifically its first, second, fourth moment, and the reference deviation) and granulation LS. Influence of other process parameters was not confirmed. Therefore, these are excluded from the modeling procedure. The proposed structure of the granule size process model is depicted in Figure 5. Due to reproducibility issues, the third characteristic moment is excluded from the model structure.

Journal Pre-proof

Journal Pre-proof

Table 2: Executed DoE for capturing the effect of granulation process parameters on the size of wet granules.

Run Nr. [/]	LS [%]	SFR [$\frac{kg}{h}$]	BT $°C$	SS $[rpm]$	M_1 $[mm]$	M_2 $[mm^2]$	M_3 [/]	M_4 [/]	e_{ref} [/]
1	24	15	30	700	2.2	1.7	-0.07	1.7	0.07
2	18	10	35	900	1.0	0.6	0.22	1.7	-0.21
3	18	20	35	900	1.2	0.6	0.04	1.78	-0.11
4	30	10	35	500	3.0	2.4	0.13	2.2	0.22
5	30	20	35	500	2.9	2.0	-0.2	2.1	0.19
6	24	15	30	700	2.0	1.5	-0.1	1.8	0.06
7	18	20	25	500	1.3	0.7	-0.06	1.7	-0.10
8	18	10	25	500	1.2	0.6	-0.07	1.7	-0.14
9	30	20	25	900	3.0	1.6	0.176	2.4	0.21
10	30	10	25	900	3.0	2.3	0.1	2.2	0.21
11	24	15	30	700	2.0	1.3	-0.18	1.9	0.06

Acronyms:
Granulation process parameters: Liquid-to-solid ratio (LS), solid feed rate (SFR), barrel temperature (BT), screw speed (SS).
Particle size distribution characteristics: Moment 1 (M_1), moment 2 (M_2), moment 3 (M_3), moment 4 (M_4), reference deviation (e_{ref}).

2.2.5. Local linear model tree (LoLiMoT) approach

The process modeling is performed by means of the LoLiMoT approach (Nelles, 1997). LoLiMoT is an algorithm for the data-driven identification of nonlinear systems by means of weighted local linear models (LLM).

General description. The standard model structure depicted in Figure 6a involves p inputs $u_1, \ldots, u_p$, and one output $\hat{y}$. Each of M LLMs outputs is associated with the respective validity function $\Phi_1, \ldots, \Phi_M$. LoLiMoT is a two-step approach: The first step involves LLM parameter identification ($w_{i0}, \ldots, w_{ip}, i \in 1, \ldots, M$), with individual LLM outputs defined as

$$y_i = w_{i0} + w_{i1}u_1 + \ldots + w_{ip}u_p. \tag{8}$$

The second step involves the input range partitioning in order to determine the tree structure of the M LLMs. The final model output is calculated as a weighted sum of individual LLMs

$$\hat{y} = \sum_{i=1}^{M} \hat{y}_i \Phi_i(u). \tag{9}$$

Figure 6b illustrates the application of the LoLiMoT algorithm for the approximation of a nonlinear static function.

Design of excitation run. Data-driven modeling implies a direct relationship between the experimental data used for the model identification and the model quality. The

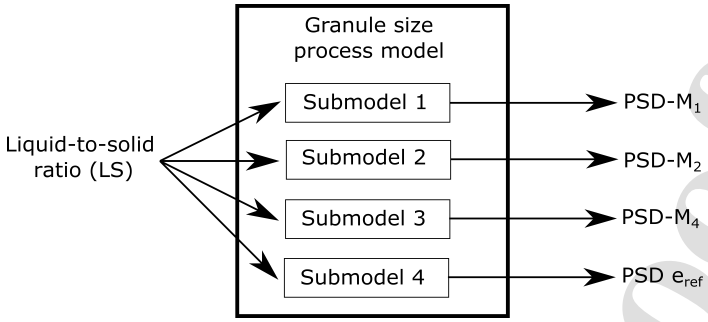
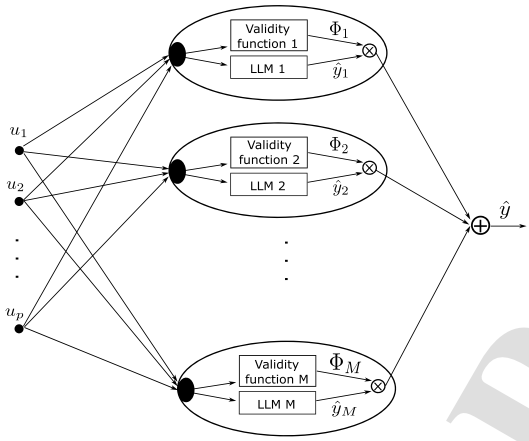
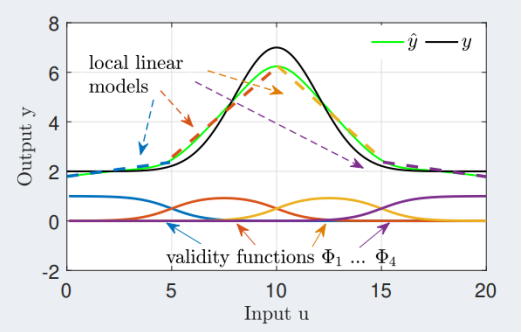


Figure 5: Proposed structure of the granule size process model.



(a) Standard LoLiMoT model structure.



(b) The nonlinear static function y is approximated via four LLMs. LLM parameters, i.e., the offset and the slope of straight lines, and the respective validity functions are determined. The final model output $\hat{y}$ is calculated as a weighted sum of individual LLMs.

Figure 6: LoLiMoT approach.

choice of appropriate excitation signals for modeling experiments is crucial to obtain an accurate model. Therefore, a combination of two excitation runs was performed. The first excitation run involved amplitude modulated pseudo random binary signal (APRBS) LS variations. APRBS, characterized by random level amplitude (A) and level duration (L) is a standard choice for the LoLiMoT excitation signal. The results of the first excitation run were utilized to design the second excitation run. Again, the LS sequence was chosen in an APRBS-like fashion, but A and L were optimized to achieve the maximal average distance between the new and existing data points. The following paragraph can be used as a step-by-step guide for the design of the excitation run (for the visual representation of the introduced steps please refer to Section 3.1).

Step 1. An APRBS LS sequence was designed using the System Identification Toolbox in MATLAB (MATLAB, 2023c) and applied to the real production plant. The PSD of produced wet granules was captured via the Parsum probe, and the first moment was evaluated according to Equation 1. The first part of the experiment (approximately thirty minutes) involved slow transitions between the random amplitude levels offering an insight into the steady-state behavior of the system. The second part of the experiment

Journal Pre-proof

involved fast transitions between the random amplitude levels capturing the system behavior in high dynamic regions.

Step 2. The designed LS- and measured M_1 sequences were provided as the identification data for the LoLiMoT algorithm. Preliminary LoLiMoT training was performed and the neuro-fuzzy-model(NFM) (Mishra, 2020) describing the relation between the LS and M_1 was obtained. This NFM was used to predict the first characteristic moment $M_{1,LOL}$. Also, the collected data was used to examine the coverage of input-output (IO) space. Typically, there are a few non-covered areas remaining in the IO space (for more details please refer to Section 3.1). To improve the model performance in these areas, the second excitation run was designed.

Step 3. In order to avoid repeating the already available experimental data points, the new APRBS-like LS sequence was designed in an optimal manner. The work (Heinz and Nelles, 2018) introduces an approach for the iterative optimization of LoLiMoT excitation signals. Level amplitude A and level duration L are considered optimization variables in each iteration. The optimization objective is defined by means of the following cost function

$$J = \frac{1}{L} \sum_{k=N+1}^{N+L} d_{NN}(\mathbf{X}_{old}, \mathbf{x}_k). \tag{10}$$

The function d_{NN} calculates the smallest distance between any existing data point stored in $\mathbf{X}_{old}$ and the newly created one $\mathbf{x}_k$, i.e., maximizing J is equivalent to maximizing the distance between the existing and new data points. The proposed algorithm involves the following steps:

- Initialization: The data matrix $\mathbf{X}_{old}$ is filled with the available data set of size N. In the first iteration, $\mathbf{X}_{old}$ contains the identification data collected in the Step 1.
- Optimization problem is defined as

$$\begin{aligned} \max_{A,L} \quad & \frac{1}{L} \sum_{k=N+1}^{N+L} d_{NN}(\mathbf{X}_{old}, \mathbf{x}_k), \\ \text{s.t.} \quad & 17\,\% \leq A \leq 32\,\% \\ & 20\,s \leq L \leq 240\,s \\ & L \in \mathbb{Z} \end{aligned} \tag{11}$$

  with $\mathbf{x}_k$ holding the to-be-optimized $LS_{new} = [A_{new} \dots A_{new}]$ signal of the length L_{new}, and the respective first moment $M_{1,new}$ predicted via the preliminary NFM obtained in step 2. The optimization problem is solved by means of a genetic algorithm using the Global Optimization Toolbox in MATLAB (MATLAB, 2023a). The genetic algorithm is chosen due to its capability to consider integer constraints (i.e., $L \in \mathbb{Z}$). The optimal A_{new} and L_{new} are obtained, the new LS sequence is appended to the existing one, and the matrix $\mathbf{X}_{old}$ is accordingly extended. Note:

Journal Pre-proof

For more details on the introduced approach (e.g., $d_{NN}, \mathbf{X}_{old}, \mathbf{x}_k$ definition) please refer to (Heinz and Nelles, 2018; Universität Siegen, 2023).

- The optimization procedure is repeated until the specified experiment duration is reached.

Step 4. The optimal LS sequence obtained in the Step 3 was applied to the ConsiGma$^{\text{TM}}$-25, the resulting PSD variations were measured (the respective M_1 evaluated), and the final IO space coverage was examined.

Journal Pre-proof

Journal Pre-proof

2.3. Granule API- and liquid content

2.3.1. Technical process description

For the creation of the chemometric model, granules of different levels of API content need to be presented to the Raman probe. In order to generate the intentional variations of the API concentration, a second loss-in-weight feeder was installed at the granulator inlet. The first and second feeder were filled with the pre-blend materials with the API concentration of 8 % and 0 %, respectively. The quantity of individual excipients in the pre-blends was adjusted to keep the same ratio as for the pre-blend depicted in Table 1. The two feeders supplied the pre-blend powders with the nominal SFR of $7.725\,kg/h$, and $7.275\,kg/h$, resulting in the nominal API concentration of 4.12 % at the granulator inlet. The remaining granulation parameters stayed the same as introduced in Section 2.2.1. A schematic representation of the ConsiGma^{TM}-25 granulation unit depicted in Figure 1 remains similar with two adaptations: The setup is extended with an equivalent second loss-in-weight feeder placed at the granulator inlet, and the Parsum measurement probe providing size distribution data is replaced by a Raman probe providing spectral data at the granulator outlet (please refer to Section 2.2.1 for more details).

2.3.2. PAT Strategy

The API content of the produced wet granules was captured by means of a sampling device and the Raman method. The Raman method is the study of in-elastically scattered light (i.e., photons scattered as different wave number) from a monochromatic light source in the context of vibrational spectroscopy. A fingerprint of specific bands is observed representing the different molecular bonds of the material and its attributes like polymorphic form. This makes the method interesting to be utilized as a PAT tool, as the peak heights correlate with concentration. Compared to near-infrared spectroscopy (NIRS), only a modest influence of PSD is present (Paudel et al., 2015). However, the in-line implementation is more complex than NIRS, as the measurement volume has to be shielded from external light sources and the sampling device has to be built from materials that do not show reflective or auto-fluorescence behavior. Therefore, a sampling device was designed that is suitable for wet granules and dry powder handling, and allows hosting a wide-angle illumination Raman probe. Specifically, the Rxn2^{TM} Hybrid Raman process spectrometer with attached PhAT (Pharmaceutical Area Testing) probe (Kaiser Optical Systems, USA (Kaiser Optical Systems, 2023)) was used in the setup. Because of its rapid prototyping capability and flexible design options, 3D-printing was used to create the needed parts. They were made of matte black PLA NX2 (Extrudr, Austria (Extrudr, 2023b)) printed on a i3 MK3S+ (Prusa, Czech Republic (Prusa, 2023)). The design consists of two rotating compartments with a volume of approximately $8\,ml$ each, in an hourglass shape with a saphire glass (Edmund optics, Germany (Edmund Optics, 2023)) on the other side, with printed cleaning seals of Flex TPU (Extrdr, Austria (Extrudr, 2023a)), and purging air ducts. The device is shown in Figure 7. The complete construction fits inside the $Ø80\,mm$ Tri-Clamp flange adapter as used with the Parsum probe (see Section 2.2.2). Open source hardware and software solutions were used when possible. A closed loop NEMA17 stepper motor (Bigtreetech,

Journal Pre-proof

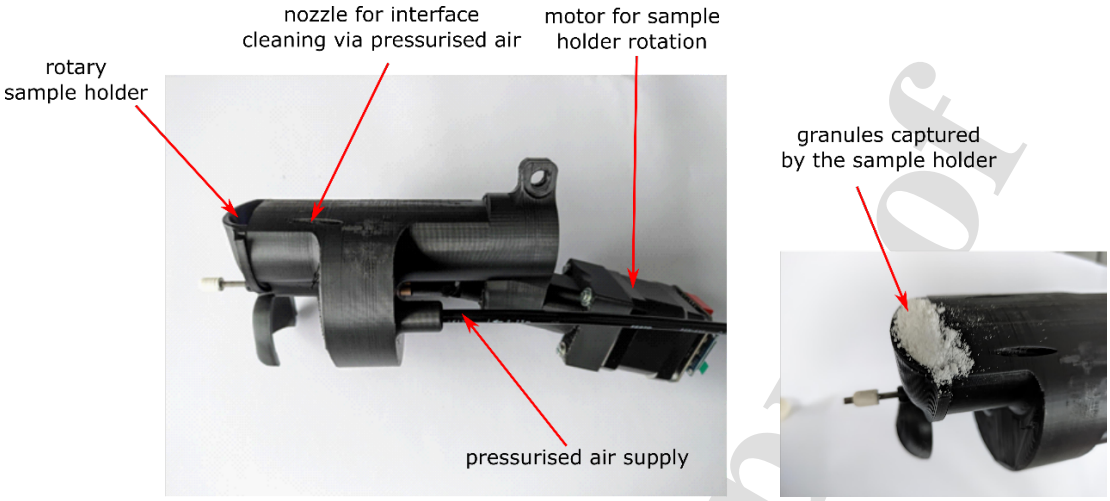


Figure 7: Sampling device used to present the granules to the Raman probe (CAPRI, 2023a).

China) with GRBL1.1 stepper motor controller board rotates the cups. Air pulses are generated with a VPPE-3-1-1 pressure regulator valve (Festo, Germany (Festo, 2023)) and $4-20\,mA$ T-Click interface board (MikroE, Serbia (MIKROE, 2023)). The communication and synchronization between Raman and sampling device is implemented over OPC-UA protocol.
The OPC-UA server, using the open source python-opcua library (FreeOpcUa, 2023), runs on the Raman station and waits for a new recorded spectrum from the iC Raman 4.1.915 software (Kaiser Optical Systems), that are pushed to subscribed clients and starts the next sampling cycle. The sampling device control box houses a Raspberry Pi 4 Model B (Raspberry Pi, 2023) that executes two Python scripts (using the opcua-asyncio library): The OPC-UA client for sample device control and the chemometric model prediction engine. After the signal for the new sample cycle is received, the purging air pulse is generated, and the step motor is instructed to rotate the cups by 180 degrees. Then the cups are filled with fresh granules and a $15\,s$ Raman spectrum acquisition takes place (until the measurement cycle starts over again). The complete cycle time of $20\,s$ is chosen as a compromise between signal-to-noise ratio and measurement frequency.

2.3.3. Chemometric modeling

The Raman spectral data collected every $20\,s$ (cycle time of the sampling device) and the respective API concentration at the granulator inlet C_{in} (computed from the pre-blend concentrations and the feeder mass flow rates) as

$$C_{in} = \frac{SFR1 \cdot 8\%}{SFR1 + SFR2} \tag{12}$$

were used for the chemometric model development. The baseline of the captured Raman spectra was corrected using Whittaker's asymmetric least square (ALS) fitting

Journal Pre-proof

algorithm (Eilers, 2003). Also, the standard-normal-variate (SNV) processing of the selected part of the baseline-corrected spectra (i.e., subtraction of the mean value and normalization by the standard deviation) was performed. For initial chemometric model development SIMCA multivariate data analysis software (Sartorius Stedim Data Analytics AB, Sweden, (Sartorius, 2023b)) was used. A partial least squares (PLS) regression model with three PLS components was trained on the spectral range from $445\,cm^{-1}$ to $1725\,cm^{-1}$, giving a root mean square error in cross validation of 0.20 wt% API. Model coefficients were exported and used in a custom chemometric prediction engine implemented in a Python script: An OPC-UA client runs using the opcua-asyncio library, the spectra pre-processing steps and PLS calculations are performed, and the predicted values are made available via an OPC-UA server.
Furthermore, the Raman spectral data were used to estimate the liquid content of the granules at the granulator outlet (LS_{out}). In this case, the Raman spectral data and the respective LS_{in} at the TSG inlet calculated as

$$LS_{in} = \frac{LFR}{SFR1 + SFR2} \tag{13}$$

were provided as the identification data for the development of the second chemometric model. Again, the baseline correction and SNV processing of the Raman spectra were performed. MATLAB with the Statistics and Machine Learning Toolbox (MATLAB, 2023b) was used to train a PLS model with six components on the spectral range from $100\,cm^{-1}$ to $1600\,cm^{-1}$. For a detailed description of DoE used for model training, please refer to Section 3. Furthermore, the Raman analysis of pre-blend raw materials was performed using the Raman PhAT probe with an exposure time of $15\,s$ (the same exposure time, as for the inline measurement). The obtained raw spectra are shown in Figure A.13 (please refer to Appendix A).

3. Results and discussion

3.1. Granule size model

Experimental data for model identification. The experimental data required for model training were collected via systematically designed excitation runs. As introduced in Section 2.2.5, two system excitation experiments were performed, the first excitation run with a typical LoLiMoT input sequence (APRBS), and the second excitation run with an optimal input sequence. Figure 8 outlines the steps taken for the design of two excitation runs (please refer to Section 2.2.5 for details). The introduced four-step approach involves the following: In Step 1 the APRBS LS sequence was designed and applied to the ConsiGma^{TM}-25 (first excitation run), the induced PSD variations were captured via the PAT solution introduced in Section 2.2.2, and the first characteristic moment of the PSD was evaluated. Step 2 involves preliminary LoLiMoT training (a model predicting M_1 from LS trained on the identification data collected in Step 1) and examination of IO space coverage. In Step 3, the LS sequence was designed in an optimal manner to cover the blank regions of IO space (empty or poorly covered regions

Journal Pre-proof

in the first excitation run). For that purpose, the identification data from Step 1 and the preliminary model trained in Step 2 were used by the optimization algorithm. Step 4 basically reiterates the first two steps with the LS sequence designed in Step 3, i.e., the optimal LS sequence was applied to the ConsiGma^{TM}-25 (second excitation run), the first characteristic moment of PSD was evaluated, and the final IO space coverage was examined. The experimental data obtained in Step 4 successfully extended the preliminary IO space (Step 2), i.e., the initially empty ranges were filled with new data points. Therefore, it can be stated that the collected experimental data precisely reflect the system behavior on the operating range of interest and can be used for model identification.
Note: Although the obtained experimental data are only valid for modeling of the investigated pharmaceutical formulation, the proposed method for the design of excitation runs is generally applicable and can be re-executed for different formulations in a timely manner.

Model identification. The PSD data collected in the two excitation runs were merged, the PSD signal processing was performed, and the PSD characteristics were evaluated (according to Equations 1, 2, 4, and 6 introduced in Section 2.2.3). The combined sequences (LS, M_1, M_2, M_4, and e_{ref}) were provided as the identification data set for the LoLiMoT training. The training was performed in MATLAB using the LMN-Toolbox LMN-Tool (2023); Hartmann et al. (2012) and repeated for each of the proposed model outputs, i.e., a submodel per PSD characteristics was created (please see Figure 5). As a result, the final granule size process model involves four neuro-fuzzy-models. For the M_1 model, the time-delayed LS and size sequences were arranged as

$$\mathbf{u}^T = [u_1, \ldots, u_p] = [LS_{k-1} \ldots LS_{k-n_{LOL}}, M_{1,k-1} \ldots M_{1,k-n_{LOL}}], \tag{14}$$

and used as the inputs for the LoLiMoT structure proposed in Figure 6a. The time delay $n_{LOL} = 3$ was chosen as a compromise between the model accuracy and complexity. The analog approach was followed for the remaining PSD characteristics. The LoLiMoT algorithm suggested to use the structures with 15, 12, 14, and 12 LLMs for M_1, M_2, M_4, and e_{ref} models, respectively. These results (together with the obtained experimental data) confirm the statements regarding the non-linear behavior of granule size introduced in (Nicolaï, 2019), and justify the choice of modeling algorithm (in contrast to LoLiMoT, linear models like, e.g., transfer functions would not be sufficient to accurately reflect the system behavior in the complete operating range).

Evaluation of model performance. The model performance was investigated both on the training and validation data set. The modeling results depicted in Figures 9 and 10 indicate a good agreement between the measured and predicted PSD characteristics and confirm the quality of the identified model. The validation experimental data set was not chosen randomly but originates from the preliminary control concept experiments. In these experiments, the reference value for M_1 is changed, and the LS is accordingly adjusted, i.e., this data set involves gradual changes of LS over the complete operating range, as well as the dynamic short-time deviations from the nominal point. Therefore,

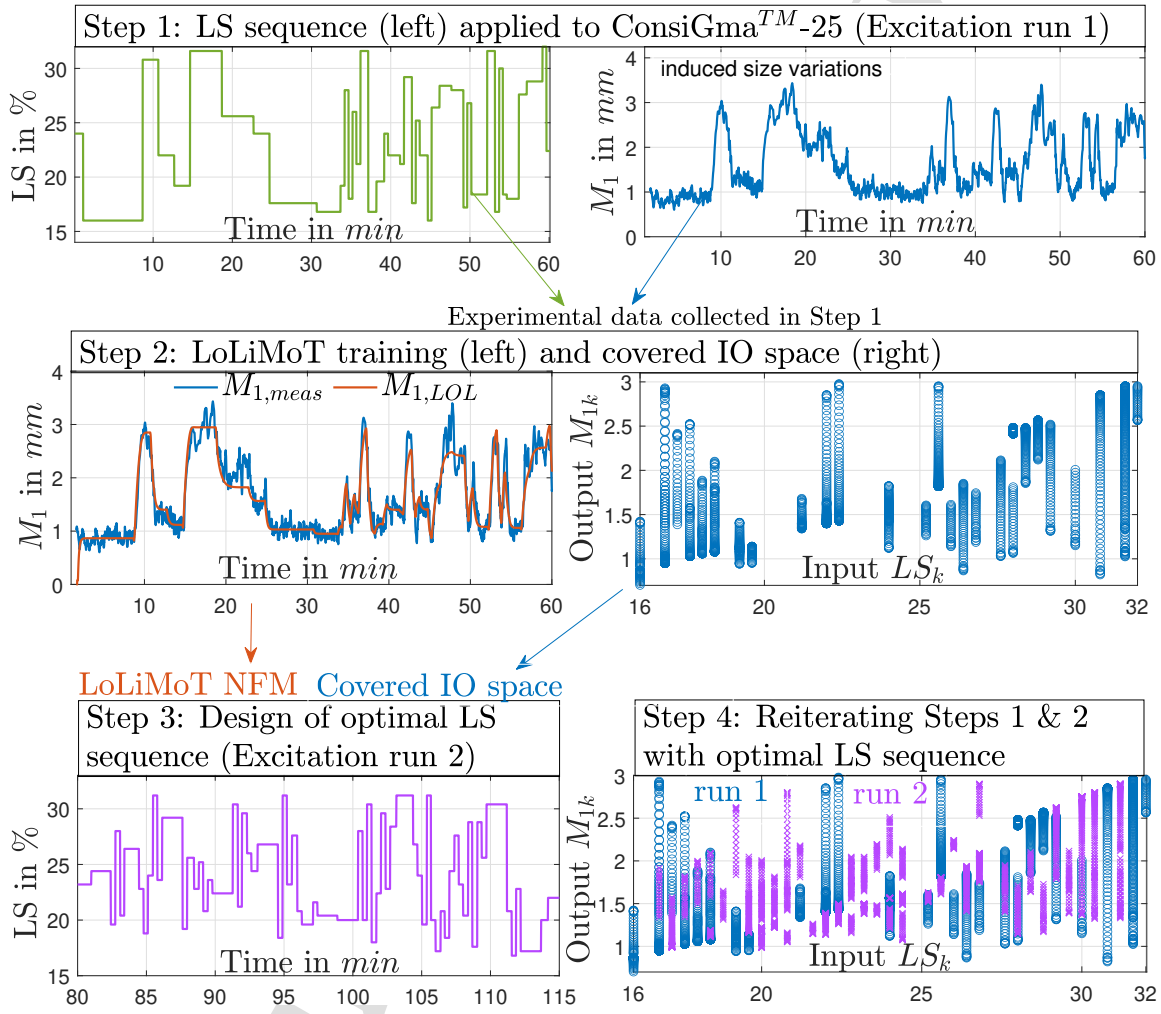


Figure 8: Step-by-step design of the LoLiMoT excitation run.

Journal Pre-proof

the good model performance is even more significant, as this corresponds to a realistic application example. A discrepancy between the measurement and the model prediction at the end of the validation experiment (please see the violet section in Figure 10) originates from the difference between the LS set-point (SP) used as model input and its respective actual value (AV). A fairly similar event would also occur in the case of raw material variability (e.g., PB with different PSD) or equipment faults (e.g., inaccurate feeding) where the model prediction would not match the measured PSD characteristics. This indicates that the difference between the measurement and model prediction could be used for the development of a fault detection algorithm. Furthermore, the SP and AV of LS could be compared in real-time, and the potential deviations could be used as correction terms, additionally improving the model performance.

Journal Pre-proof

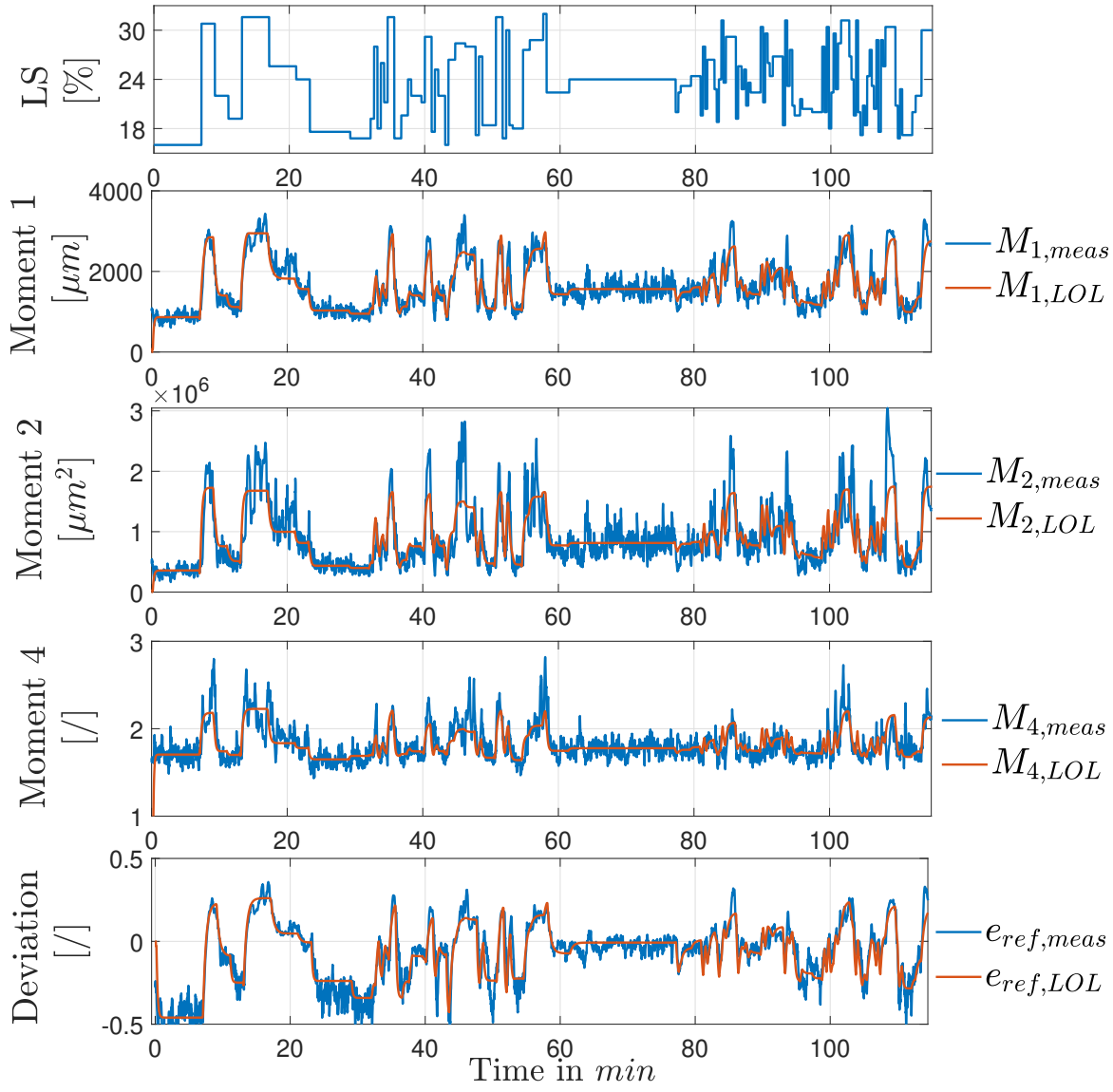


Figure 9: The granule size model trained via the LoLiMoT approach predicts PSD characteristics accurately on the training data set (CAPRI, 2023b).

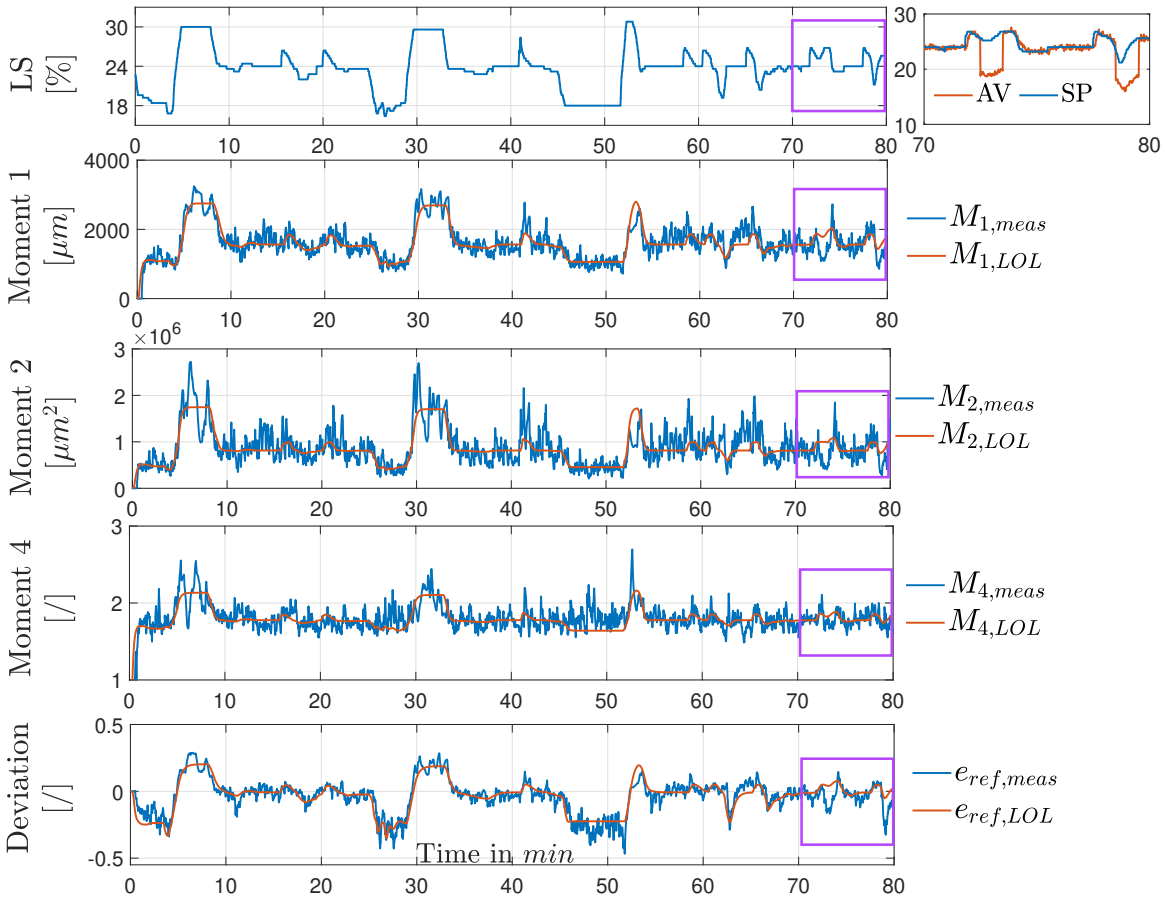


Figure 10: The granule size model trained via the LoLiMoT approach predicts PSD characteristics accurately on the validation data set (CAPRI, 2023b).

Journal Pre-proof

3.2. Granule API- and moisture content model identification and validation

The experimental data required for the chemometric model development were collected by means of the DoE data depicted in Table 3. The variations of inlet API concentration in a range of 1.33 % to 7 % were realized via SFR variations. Additionally, in order to assure the model robustness by different amounts of granulation liquid, LS variations in a range of 18 % to 30 % were introduced to the system via LFR variations. Thus, the Raman spectral data at the boundary (1.33 % and 7 %) and central (4.12%) DoE points were captured at different LS levels. Each run was executed for approximately six minutes in order to achieve the steady-state operation. The chemometric models predicting the outlet API concentration and outlet LS were developed according to the procedure described in Section 2.3.3. The model calibration results are depicted in Figure B.14 (please refer to Appendix B).
The performance of the developed chemometric models was investigated for the training- and the validation data set (please refer to Figures 11 and 12). In both cases, a satisfactory agreement between the set inlet- and estimated outlet API concentration is confirmed. The same can be stated for the results obtained with the second chemometric model estimating the liquid content at the TSG outlet from Raman spectral data. The 20 s time delay between the inlet- and via models estimated outlet quantities is not a sign of the model weakness, but a consequence of the sampling time of the sampling device (please refer to Section 2.3.2 for more details).

Table 3: Development of chemometric models DoE.

Run Nr. [/]	SFR1 $[\frac{kg}{h}]$	SFR2 $[\frac{kg}{h}]$	LFR $[\frac{g}{min}]$	LS_{in} [%]	C_{in} [%]
1	2.5	12.5	60	24	1.33
2	2.5	12.5	45	18	1.33
3	5.6	9.4	75	30	3.0
4	7.73	7.27	75	30	4.12
5	7.73	7.27	60	24	4.12
6	7.73	7.27	45	18	4.12
7	9.38	5.62	75	30	5.0
8	13.13	1.87	75	30	7.0
9	13.13	1.87	60	24	7.0
10	13.13	1.87	45	18	7.0

Acronyms:
Solid feed rate of API feeder (SFR1), solid feed rate of the excipient feeder (SFR2), liquid feed rate (LFR), liquid-to-solid ratio at the granulator inlet (LS_{in}), API concentration at the granulator inlet (C_{in}).

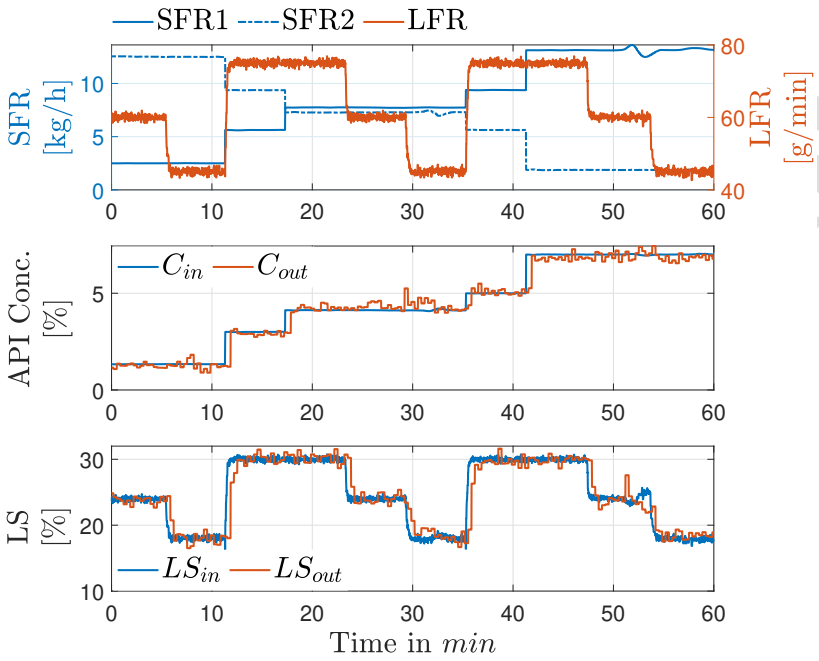


Figure 11: Chemometric model predicting API- and liquid content of wet granules from Raman spectral data for the training data set.

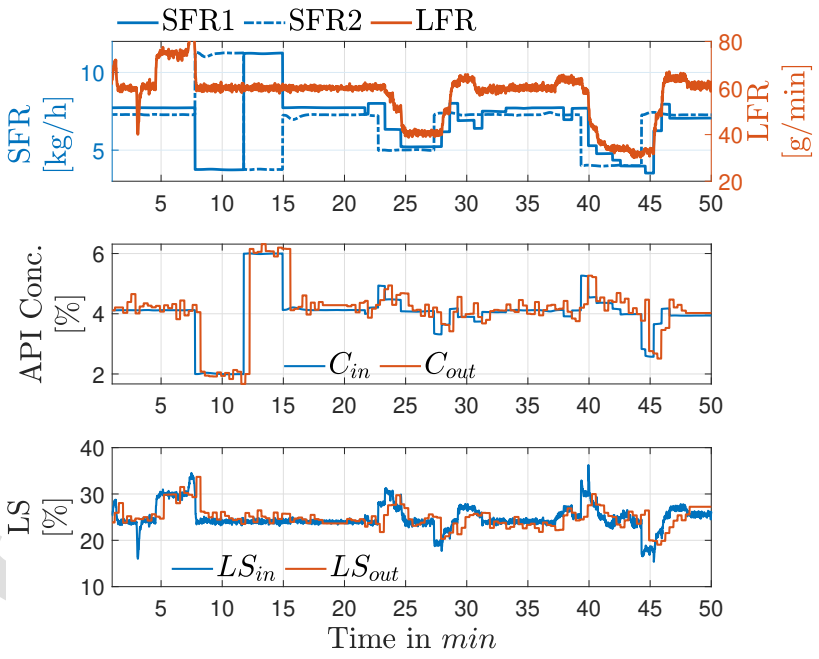


Figure 12: Chemometric model predicting API- and liquid content of wet granules from Raman spectral data for the validation data set.

Journal Pre-proof

4. Conclusion and outlook

The solutions proposed in this work allow the development of advanced control concepts for granule size, API-, and liquid content of wet granules in ConsiGma^{TM}-25. The validation results of the dynamic model for granule size indicate very good conformity between the measurement and model estimation and thereby confirm the model quality. The same can be stated for the results of validation experiments obtained for the chemometric models predicting the API- and liquid content from the Raman spectral data. As such, these solutions can be used as a reliable (and necessary) basis for the development of the following use cases.

Fault detection and digital assistant. The granule size process model can run in parallel with the process. The difference between the predicted (i.e., $M_{1,LOL}$, $M_{2,LOL}$, $M_{4,LOL}$, $e_{ref,LOL}$) and via PAT probe measured PSD characteristics (i.e., M_1, M_2, M_4, e_{ref}) can be used as trigger signals for the fault detection algorithm. This algorithm can be developed to detect and distinguish between different process disturbances, such as equipment or material faults. Furthermore, the fault detection algorithm can support the operator of the manufacturing line via an appropriate digital assistant concept. The digital assistant concept will generate valuable suggestions to the operator, e.g., to check the line for potential faults and eliminate them.

Soft-sensor for granule size. The identified granule size process model can be used as a soft-sensor, acting as a potential replacement for the Parsum probe. The soft-sensor application will be exceptionally valuable for the ConsiGma^{TM}-25 constellations where the mounting of the Parsum PAT probe is not feasible. The equipment setup with the Raman probe placed at the TSG outlet would be an example of such a constellation: this configuration does not allow the additional installation of the Parsum probe due to space limitations. In this case, the soft-sensor application will allow the simultaneous monitoring of all intermediate CQAs, i.e., the measurement of API- and liquid content via Raman, and the prediction of the wet granules size via the LoLiMoT approach.

Quality control concept. Intermediate CQAs, i.e., granule size, API-, and liquid content can be monitored in real-time either via installed PAT equipment or via a soft-sensor. This information can be used to discard non-conforming material by means of an advanced discharge control concept.

Process control concept. The granule size dynamic process model can be used for the development of nonlinear MPC in a straightforward manner. The model quality significantly impacts the MPC performance: on the one hand, the model is used during MPC parameter tuning via simulation studies (acting as a replacement for the real system), and on the other hand, the MPC prediction algorithm that is part of the MPC concept uses this process model as a core component. In such an application, the MPC algorithm adjusts the granulation process parameters (e.g., LS) in order to keep the granule size (e.g., M_1) close to the reference. Furthermore, the API- and liquid content predicted via chemometric models can be utilized as controlled variables in an appropriate feedback

Journal Pre-proof

control concept. Similarly to the granule size MPC, the granulation process parameters can be adjusted in order to keep API- and liquid content close to the reference values.

The introduced use cases would ensure an increased product quality and allow the mitigation of process disturbances, ultimately improving the performance of the ConsiGma^{TM}-25 production plant.

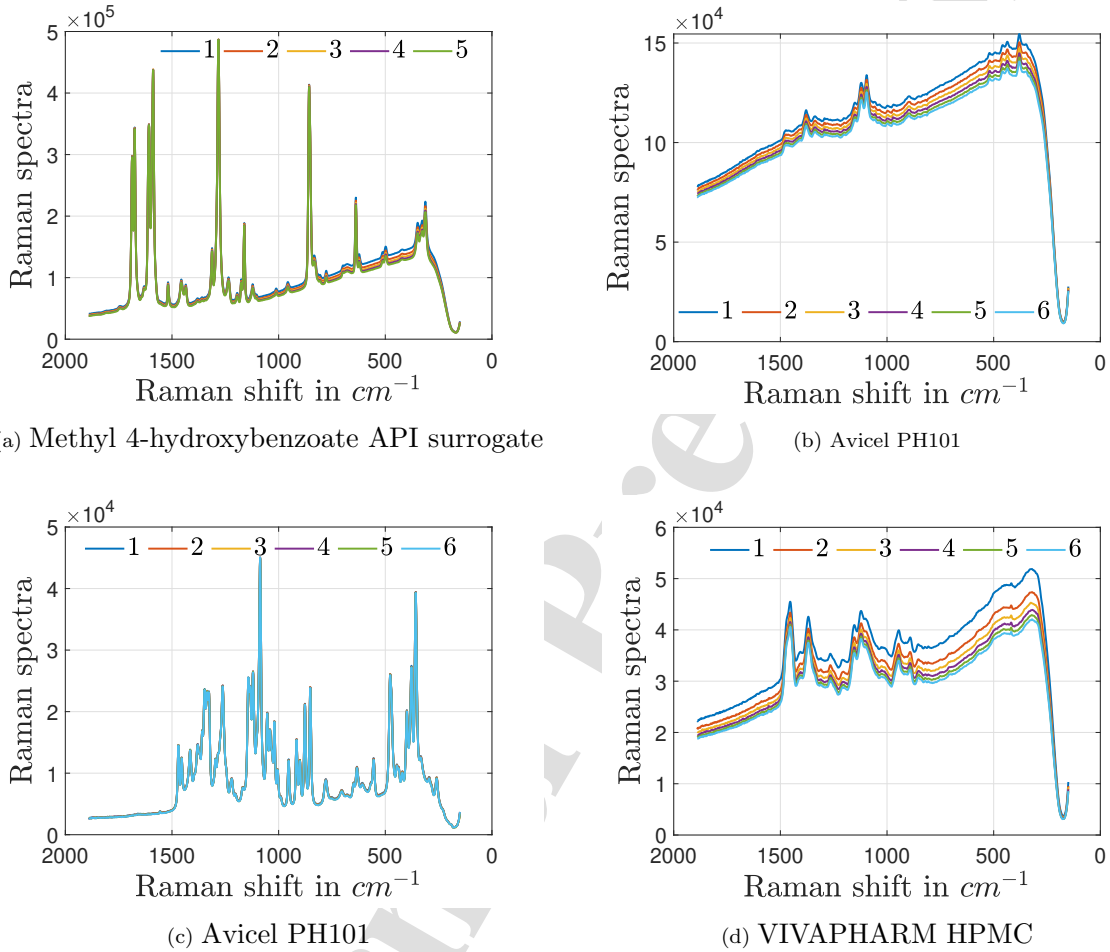
5. Acknowledgments

The COMET Center Research Center Pharmaceutical Engineering (RCPE) is funded within the framework of COMET - Competence Centers for Excellent Technologies by BMK, BMDW, Land Steiermark and SFG. The COMET program is managed by the FFG.
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870062.

Journal Pre-proof

Appendices

Appendix A. Raman analysis of raw pre-blend materials.



(a) Methyl 4-hydroxybenzoate API surrogate

(b) Avicel PH101

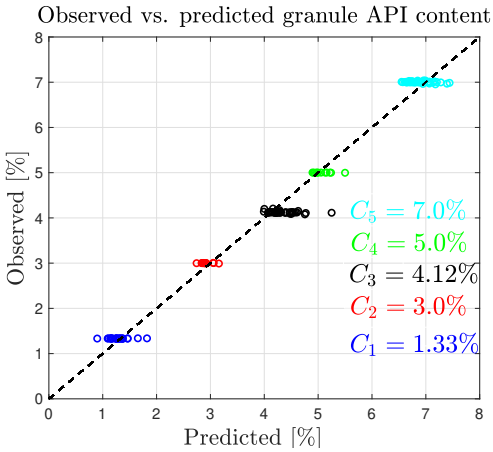
(c) Avicel PH101

(d) VIVAPHARM HPMC

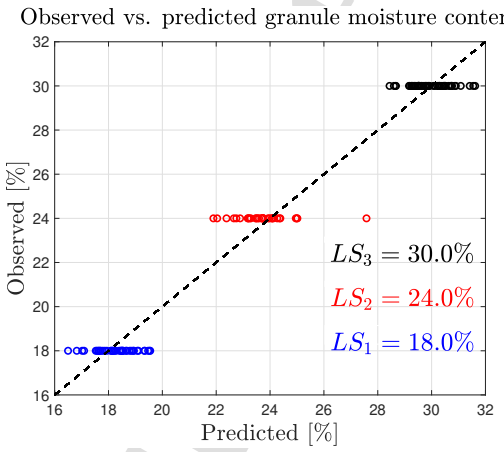
Figure A.13: Raman analysis of raw materials in the pre-blend.

Journal Pre-proof

Appendix B. Raman calibration



(a) Calibration of the chemometric model for API content.



(b) Calibration of the chemometric model for liquid content.

Figure B.14

Journal Pre-proof

Journal Pre-proof

Bibliography

Barrasso, D., Ramachandran, R., 2015. Qualitative assessment of a multi-scale, compartmental pbm-dem model of a continuous twin-screw wet granulation process. Journal of Pharmaceutical Innovation 11. doi:10.1007/s12247-015-9240-7.

Barrasso, D., Tamrakar, A., Ramachandran, R., 2014. A reduced order pbm–ann model of a multi-scale pbm–dem description of a wet granulation process. Chemical Engineering Science 119, 319–329. URL: https://www.sciencedirect.com/science/article/pii/S0009250914004230, doi:https://doi.org/10.1016/j.ces.2014.08.005.

Barrasso, D., Walia, S., Ramachandran, R., 2013. Multi-component population balance modeling of continuous granulation processes: A parametric study and comparison with experimental trends. Powder Technology 241, 85–97. URL: https://www.sciencedirect.com/science/article/pii/S0032591013001666, doi:https://doi.org/10.1016/j.powtec.2013.03.001.

CAPRI, 2023a. Project deliverable will be available in .pdf format under number D3.2. https://www.capri-project.com/project-deliverables, accessed on 26/01/2023.

CAPRI, 2023b. Project deliverable will be available in .pdf format under number D4.1. https://www.capri-project.com/project-deliverables, accessed on 26/01/2023.

Coperion, 2023. K-TRON Feeders. https://www.coperion.com/en/products-services/process-equipment/feeders/twin-screw-feeders, accessed on 26/01/2023.

Dieter, P., Stefan, D., Günter, E., Michael, K., 2011. In-line particle sizing for real-time process control by fibre-optical spatial filtering technique (sft). Advanced Powder Technology 22, 203–208. URL: https://www.sciencedirect.com/science/article/pii/S0921883110002165, doi:https://doi.org/10.1016/j.apt.2010.11.002. special issue of the 6th World Congress on Particle Technology.

Edmund Optics, 2023. Sapphire Window. https://www.edmundoptics.de/p/15mm-diameter-uncoated-sapphire-window/4991/, accessed on 27/01/2023.

Eilers, P.H.C., 2003. A Perfect Smoother. Analytical Chemistry 75, 3631–3636. URL: https://doi.org/10.1021/ac034173t, doi:10.1021/ac034173t.

Extrudr, 2023a. Flex TPU. https://www.extrudr.com/en/products/catalogue/tpu-medium-schwarz_2240/, accessed on 27/01/2023.

Extrudr, 2023b. PLA NX2. https://www.extrudr.com/en/products/catalogue/pla-nx2-schwarz_1966/, accessed on 27/01/2023.

Journal Pre-proof

Festo, 2023. VPPE-3-1-1 pressure regulator valve. https://www.festo.com/de/en/a/567540/, accessed on 27/01/2023.

FreeOpcUa, 2023. FreeOpcUa:python-opcua. https://github.com/FreeOpcUa/python-opcua, accessed on 18/04/2023.

GEA, 2023. ConsiGma^{TM}-25 manufacturing line. https://www.gea.com/de/products/tablet-presses/continuous-tableting-lines/consigma-continuous-tablet-line.jsp, accessed on 26/01/2023.

gProms, 2023. gProms. https:/https://www.psenterprise.com/products/gproms, accessed on 17/04/2023.

Hartmann, B., Ebert, T., Fischer, T., J.Belz, Kampmann, G., Nelles, O., 2012. " LM-Ntool - Toolbox zum automatischen Trainieren lokaler Modellnetze". 22. Workshop Computational Intelligence, Dortmund .

Heinz, T.O., Nelles, O., 2018. Excitation signal design for nonlinear dynamic systems with multiple inputs – a data distribution approach. at - Automatisierungstechnik 66, 714–724. URL: https://doi.org/10.1515/auto-2018-0027, doi:doi:10.1515/auto-2018-0027.

ICH, 2005. ICH Harmonised Tripartite Guideline, Quality Risk Management Q9. https://database.ich.org/sites/default/files/Q9_Guideline.pdf, accessed on 27/01/2023.

ICH, 2008. ICH Harmonised Tripartite Guideline, Pharmaceutical Quality SystemQ10. https://database.ich.org/sites/default/files/Q10Guideline.pdf, accessed on 27/01/2023.

ICH, 2009. ICH Harmonised Tripartite Guideline, Pharmaceutical Development Q8(R2). https://database.ich.org/sites/default/files/Q8_R2_Guideline.pdf, accessed on 27/01/2023.

ICH, 2012. ICH Harmonised Tripartite Guideline, Development and Manufacture of Drug Substances Q11. https://database.ich.org/sites/default/files/Q11Guideline.pdf, accessed on 27/01/2023.

iFix, 2023. iFix OPC. https://www.ge.com/digital/documentation/ifix/version61/Subsystems/DBB/content/dbb_using_the_opc_client_io_driver.htm, accessed on 26/01/2023.

Kaiser Optical Systems, 2023. Rxn2TM Hybrid Raman process spectrometer. https://www.us.endress.com/en/field-instruments-overview/optical-analysis-product-overview/raman-rxn2-analyzer?t.tabId=product-overview, accessed on 27/01/2023.

Journal Pre-proof

Lee, S., O'Connor, T., Yang, X., Cruz, C., Chatterjee, S., Madurawe, R., Moore, C., Yu, L., Woodcock, J., 2015. Modernizing pharmaceutical manufacturing: from batch to continuous production. Journal of Pharmaceutical Innovation 10. doi:10.1007/s12247-015-9215-8.

LMN-Tool, 2023. Matlab-toolbox for local model networks, universität Siegen. https://www.mb.uni-siegen.de/mrt/lmn-tool/?lang=de, accessed on 26/01/2023.

Markl, D., Warman, M., Dumarey, M., Bergman, E.L., Folestad, S., Shi, Z., Manley, L.F., Goodwin, D.J., Zeitler, J.A., 2020. Review of real-time release testing of pharmaceutical tablets: State-of-the art, challenges and future perspective. International Journal of Pharmaceutics 582, 119353. URL: https://www.sciencedirect.com/science/article/pii/S0378517320303379, doi:https://doi.org/10.1016/j.ijpharm.2020.119353.

MATLAB, 2023a. Global Optimization Toolbox. https://de.mathworks.com/products/global-optimization.html, accessed on 26/01/2023.

MATLAB, 2023b. Statistics and Machine Learning Toolbox. https://www.mathworks.com/products/statistics.html, accessed on 26/01/2023.

MATLAB, 2023c. System Identification Toolbox. https://www.mathworks.com/products/sysid.html, accessed on 27/01/2023.

Metta, N., Ghijs, M., Schäfer, E., Kumar, A., Cappuyns, P., Assche, I., Singh, R., Ramachandran, R., De Beer, T., Ierapetritou, M., Nopens, I., 2019. Dynamic flowsheet model development and sensitivity analysis of a continuous pharmaceutical tablet manufacturing process using the wet granulation route. Processes 7, 234. doi:10.3390/pr7040234.

Mishra, S., 2020. Neuro-Fuzzy Models and Applications. doi:10.4018/978-1-5225-5793-7.ch004.

Nelles, O., 1997. LOLIMOT - Lokale, lineare Modelle zur Identifikation nichtlinearer, dynamischer Systeme. at - Automatisierungstechnik 45, 163–174. URL: https://doi.org/10.1524/auto.1997.45.4.163, doi:10.1524/auto.1997.45.4.163.

Nicolaï, N., De Leersnyder, F., Copot, D., Stock, M., Ionescu, C.M., Gernaey, K.V., Nopens, I., De Beer, T., 2018. Liquid-to-solid ratio control as an advanced process control solution for continuous twin-screw wet granulation. AIChE Journal 64, 2500–2514. URL: http://doi.wiley.com/10.1002/aic.16161, doi:10.1002/aic.16161.

Nicolaï, N., 2019. Supervisory process monitoring, identification and control for continuous pharmaceutical wet granulation. Ph.D. thesis. Ghent University - Faculty of Bioscience Engineering.

Journal Pre-proof

Parsum, 2023. IPP 80-P inline particle-measuring probe. https://www.parsum.de/en/particle-probes/ipp-80-p-pharma-probe/parsum-messsonde-ipp80p-3/, accessed on 26/01/2023.

Paudel, A., Raijada, D., Rantanen, J., 2015. Raman spectroscopy in pharmaceutical product design. Advanced Drug Delivery Reviews 89, 3–20. URL: https://www.sciencedirect.com/science/article/pii/S0169409X15000599, doi:https://doi.org/10.1016/j.addr.2015.04.003. pharmaceutical applications of Raman spectroscopy – from diagnosis to therapeutics.

Prusa, 2023. i3 MK3S+. https://www.prusa3d.com/, accessed on 27/01/2023.

Ramsey, J., Newton, H., Harvill, J., 2002. The Elements of Statistics: With Applications to Economics and the Social Sciences. Duxbury/Thomson Learning. URL: https://books.google.at/books?id=q1clOAAACAAJ.

Raspberry Pi, 2023. Raspberry Pi 4 Model B. https://www.raspberrypi.com/products/raspberry-pi-4-model-b/, accessed on 27/01/2023.

Rehrl, J., Kirchengast, M., Celikovic, S., Sacher, S., Kruisz, J., Khinast, J., Horn, M., 2019. Improving pellet quality in a pharmaceutical hot melt extrusion process via pid control and lolimot-based mpc. Journal of pharmaceutical innovation doi:10.1007/s12247-019-09417-0.

Sartorius, 2023a. Modde DoE. https://www.sartorius.com/en/products/process-analytical-technology/data-analytics-software/doe-software/modde, accessed on 26/01/2023.

Sartorius, 2023b. SIMCA, Multivariate Data Analysis. https://www.sartorius.com/en/products/process-analytical-technology/data-analytics-software/mvda-software/simca, accessed on 26/01/2023.

Seem, T.C., Rowson, N.A., Ingram, A., Huang, Z., Yu, S., de Matas, M., Gabbott, I., Reynolds, G.K., 2015. Twin screw granulation — a literature review. Powder Technology 276, 89–102. URL: https://www.sciencedirect.com/science/article/pii/S0032591015001023, doi:https://doi.org/10.1016/j.powtec.2015.01.075.

Shirazian, S., Kuhs, M., Darwish, S., Croker, D., Walker, G.M., 2017. Artificial neural network modelling of continuous wet granulation using a twin-screw extruder. International Journal of Pharmaceutics 521, 102–109. URL: https://www.sciencedirect.com/science/article/pii/S037851731730090X, doi:https://doi.org/10.1016/j.ijpharm.2017.02.009.

Siemens, 2023. SIMATIC-SIPAT. https://support.industry.siemens.com/cs/document/109758269/simatic-sipat-v5-1-is-available?dti=0&lc=en-AR, accessed on 26/01/2023.

Journal Pre-proof

Silva, A.F., Burggraeve, A., Denon, Q., Van der Meeren, P., Sandler, N., Van Den Kerkhof, T., Hellings, M., Vervaet, C., Remon, J.P., Lopes, J.A., De Beer, T., 2013. Particle sizing measurements in pharmaceutical applications: Comparison of in-process methods versus off-line methods. European Journal of Pharmaceutics and Biopharmaceutics 85, 1006–1018. URL: https://www.sciencedirect.com/science/article/pii/S0939641113001215, doi:https://doi.org/10.1016/j.ejpb.2013.03.032.

Stieß, M., 2008. Mechanische Verfahrenstechnik - Partikeltechnologie 1. Springer, Berlin, Heidelberg.

Universität Siegen, 2023. Design of Excitation Signals for Identification. https://www.mb.uni-siegen.de/mrt/research/kapitel7/chapter7.html?lang=de, accessed on 27/01/2023.

MIKROE, 2023. Click boards. https://www.mikroe.com/click-boards, accessed on 27/01/2023.

Wang, L.G., Omar, C., Litster, J., Slade, D., Li, J., Salman, A., Bellinghausen, S., Barrasso, D., Mitchell, N., 2022. Model driven design for integrated twin screw granulator and fluid bed dryer via flowsheet modelling. International Journal of Pharmaceutics 628, 122186. URL: https://www.sciencedirect.com/science/article/pii/S0378517322007402, doi:https://doi.org/10.1016/j.ijpharm.2022.122186.

Yu, L., Amidon, G., Khan", M., 2014. Understanding pharmaceutical quality by design. The AAPS Journal 16, 771–783. doi:10.1208/s12248-014-9598-3.

Zaborenko, N., Shi, Z., Corredor, C., Smith-Goettler, B., Zhang, L., Hermans, A., Neu, C., Alam, M.A., Cohen, M., Lu, X., Xiong, L., Zacour, B., 2019. First-principles and empirical approaches to predicting in vitro dissolution for pharmaceutical formulation and process development and for product release testing. The AAPS Journal 21. doi:10.1208/s12248-019-0297-y.

Journal Pre-proof

Journal Pre-proof

Selma Celikovic: Conceptualization, Methodology, Investigation, Writing - Original Draft, Visualization

Johannes Poms: Methodology, Investigation, Writing - Original Draft, Visualization

Johannes Khinast: Conceptualization, Writing - Review & Editing, Supervision

Martin Horn: Conceptualization, Writing - Review & Editing, Supervision

Jakob Rehrl: Conceptualization, Writing - Review & Editing, Supervision

Journal Pre-proof

Journal Pre-proof

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof