
Article:
AlAlaween, W., Khorsheed, B., Mahfouf, M. orcid.org/0000-0002-7349-5396 et al. (2 more authors) (2020) An interpretable fuzzy logic based data-driven model for the twin screw granulation process. Powder Technology, 364. pp. 135-144. ISSN 0032-5910

https://doi.org/10.1016/j.powtec.2020.01.052

An Interpretable Fuzzy Logic Based Data-Driven Model for the Twin Screw Granulation Process

Wafa' H. AlAlaween[a], Bilal Khorsheed[b], Mahdi Mahfouf[c], Gavin K. Reynolds[d] and Agba D. Salman[b*]

[a]*Department of Industrial Engineering, The University of Jordan, Jordan*
[b]*Department of Chemical and Biological Engineering, The University of Sheffield, UK*
[c]*Department of Automatic Control and Systems Engineering, The University of Sheffield, UK*
[d]*Pharmaceutical Technology & Development, AstraZeneca, UK*
(E-mails: w.alaween@ju.edu.jo; bkhorsheed1@sheffield.ac.uk; m.mahfouf@sheffield.ac.uk; gavin.reynolds@astrazeneca.com; a.d.salman@sheffield.ac.uk[*])

Abstract

In this research, a new framework based on fuzzy logic is proposed to model the twin screw granulation (TSG) process. First, various fuzzy logic systems (FLSs) having different structures are developed to define various rule bases. The extracted fuzzy rules are assessed and reduced accordingly into a single rule base by utilizing the singular value decomposition-QR factorization (SVD-QR) approach. The resulted reduced FLS is, then, implemented to describe the TSG process mathematically and linguistically via simple to understand IF-THEN rules. The linguistic output provides an accessible framework to increase the understanding of this complex process within an industrial context. Validated on laboratory-scale experiments, it is shown that the newly proposed model successfully predicts the granule size and enhances the understanding of the TSG process. Furthermore, the proposed framework outperforms the standard FLS and the Artificial Neural Network (ANN), with an overall improvement of approximately 16% and 29% in R^2, respectively.

Keywords: Fuzzy logic system; SVD-QR approach; Twin screw granulation.

1. Introduction

In the pharmaceutical industry, it is of interest to develop a continuous tablets production line, this being due to its potential advantages in terms of cost, time, controllability and scalability [1]. In order to develop such a continuous production line, all the unit operations that include, but not limited to, mixing, granulation, drying and tabletting, should operate in a continuous mode. Granulation, by which fine particles are aggregated in order to obtain specific properties (e.g. flowability and homogeneity), is considered to be one of the key processes in the tablet production line [2]. In general, various granulation processes (e.g. high shear, roller compactor and fluidized bed) can be utilized to aggregate fine particles [3]. Among all the granulation processes, the twin screw granulation (TSG), as a continuous wet granulation process, has recently received a great deal of interest, particularly, in the pharmaceutical industry. This can be attributed to the short residence time and the flexible design [4-6].

Since the granulation process is considered to be a key stage process in the pharmaceutical industry as well as in the other industries, several research papers and books have focused on understanding and modelling such a process [2, 4, 6]. Some of these papers focused on the effect of the three main types of the granulation input parameters, namely, process, equipment and formulation parameters, on the granule properties [5, 7]. In addition, a regime map was developed to describe the three mechanisms of the granulation process, namely; wetting and nucleation, growth and consolidation, and breakage and attrition [8]. Several studies have utilized such a comprehensive understanding to develop and implement various modelling paradigms to predict the granule properties and the granulation behaviour [2, 9, 10]. These modelling paradigms are normally either data-driven or physically-based models (i.e. semi-mechanistic models). For instance, one and/or multi-dimensional population balance models (PBMs), as semi-mechanistic ones, have been developed to characterize the

evolution with time of the produced granules in terms of granule size, porosity and binder content [4, 9]. For this purpose, a two-dimensional PBM was implemented for the TSG process in order to characterize the evolution of the granule size and the liquid distribution [4]. The PBMs have also been successfully implemented to provide a deeper insight into such a process at the micro level. However, the number of the granule properties that can be investigated in such models is limited, with up to three granule properties have been examined in most of the published studies [2]. Furthermore, the interactions among its three main mechanisms have not been taken into account. Considering all the interactions among these mechanisms may result in a complex model that cannot be easily solved [11]. Therefore, and in order to investigate a greater number of properties and to consider the effect of equipment dynamics, such models have been combined with various modelling paradigms such as the Monte Carlo and the discrete element methods [9, 12].

With the recent huge advances in computing power and the new algorithms that have hitherto been presented in the related literature [12-13], data based modelling paradigms have been extensively implemented in many areas (e.g. manufacturing, marine technology, chemical engineering, powder technology, health care, etc.) [12, 15-16]. For example, Artificial Neural Networks (ANNs) and Adaptive Network Based Fuzzy Inference System (ANFIS), as data based models, have been implemented as surrogate modelling approaches for online optimization and control of various processes (e.g. polymerization, groundwater remediation, casting and flood forecasting processes) that have only hitherto been investigated via other computationally elicited models, such as the kinetic model [13-15]. The use of data-driven paradigms for modelling complex and uncertain environments is obviously not new since such approaches have been used to model the various granulation processes by mapping the investigated input parameters to the properties of the granules. For instance, regression models have been implemented to predict the main granule properties and to find the optimal input

parameters [17-18]. However, regression models are, in fact, incapable of representing the sophisticated input/output relationships and the complex interactions among the process inputs [2]. Therefore, ANNs, as data-driven models, with different configurations (i.e. different number of hidden layers, hidden neurons and activation functions) have been utilized as a surrogate model to predict the granule properties such as the granule size, which is usually characterized by its three descriptors (i.e. D_{10}, D_{50} and D_{90}) [19-20]. The ANN was also investigated in the related literature to scale-up the granulation processes [21]. However, data based models, in particular ANNs, are considered to be strictly powerful interpolators, thus, they may not perform as expected beyond the investigated range [22-23]. Other data based models, such as Fuzzy Logic Systems (FLSs), have also been developed to model the granulation process [6]. For instance, a modelling framework that integrated an interval type-2 FLS and a Gaussian mixture model was presented to represent the TSG process [6]. In addition, fuzzy logic was also adopted as a surrogate paradigm to control the granule growth with a high accuracy in the fluidized bed granulation process [24].

In general, the physical and data based approaches have their own limitations and strengths and, thus, electing the appropriate approach will undoubtedly depend on many factors including, but not limited to, (i) the nature of the process under investigation, (ii) the availability of the associated physical equations, (iii) the availability of the process data and its distribution in the space considered, and (iv) the degree of accuracy required. Therefore, modelling frameworks that integrate the two types of models have been utilized to represent the various granulation processes, in order to circumvent the limitations of implementing each type separately [9]. For instance, a multi-scale three-dimensional PBM was integrated with the partial least square regression to model the fluidized bed granulation process, where the latter was utilized to determine the unknown kernels required for the PBM by mapping the investigated process inputs to these kernels. Once the model was developed, an online

optimization strategy was adopted to minimize the difference between the predicated and the actual/experimental granule properties [25]. Furthermore, a model that integrated a computational fluid dynamics (CFD) model, PBM and radial basis function (RBF) model, which was utilized to estimate the empirical parameters (e.g. kernels) required to implement the PBM, was proposed to represent the high shear granulation process [9].

In this research work that is an extension of a previous research that was done by the same authors [6], a new predictive modelling framework that is based on fuzzy logic systems (FLSs) is proposed to model the TSG process. First of all, different FLSs having different structures are developed and optimized to extract meaningful and comprehensive information, which are presented as linguistic rules, from a limited and sparse data set. Since a large number of rules may lead to a computationally expensive and biased model with redundant rules, the extracted fuzzy rules are then analysed for their relative contributions and, consequently, reduced using the singular value decomposition-QR (SVD-QR) factorization method. Finally, the reduced FLS that consists of the optimal number of rules is utilized to predict the granules size distribution and to provide a simple understanding of the TSG process. The rest of this research paper is organized as follows: the experimental work conducted using the twin screw granulator is briefly described in Section 2. The FLS and the related theoretical background of the framework are presented in Section 3. In Section 4, the modelling results are presented and discussed. Finally, Section 5 concludes the paper.

2. Experimental Work

In this research work, two pharmaceutically relevant excipients, namely, Microcrystalline cellulose (MCC, Avicel PH-101, D_{50}=50µm, FMC Corporation, Ireland) and spray-dried Mannitol (Pearlitol® 100 SD, D_{50}=100 µm, Roquette, France) were granulated.

Such excipients were chosen because of their different interaction with water. For instance, Mannitol is a water-soluble powder, while MCC is not. However, it absorbs water. Therefore, these materials are expected to have different rates of 'wetting and nucleation', 'growth and consolidation' and 'breakage' inside the granulation barrel.

The powders were granulated using a co-rotating twin screw granulator (16mm Prism Eurolab TSG, Thermo Fisher Scientific, Karlsruhe, Germany). The granulator has a length to diameter (L/D) ratio of 25:1 (continuous, open end or die less system). Deionised water, as a liquid binder, and the two powders were fed to the granulator using a peristaltic pump (Watson Marlow, Cornwall, UK) and a gravimetric twin screw feeder (K-PH-CL-24-KT20, K-Tron Soder, Niederlenz, Switzerland), respectively. The screw configuration usually consists of conveying elements, kneading elements or a combination of both of them [6]. In this research work, three screw configurations were used: (i) a configuration that consists of conveying elements only; (ii) a configuration that consists of conveying elements and one zone of kneading elements (16 kneading discs in total); (iii) a configuration that consists of conveying elements and two zones of kneading elements (32 kneading discs in total). A schematic diagram of the co-rotating twin screw granulator and the three configurations are shown in Figure 1. After completing the granulation process, the granules were left to dry at room temperature. The Retsch Camsizer (Retsch Technology GmbH, Germany), as a dynamic image analyser, was then utilized to measure the granules size presented as volume fraction for each experiment, where the size of the granules is photo-optically recorded and measured.

Figure 1. Schematic diagram for the twin screw granulator and the screw configurations: Long pitch conveying element (LPCE, Length=2Diameter), Short pitch conveying element (SPCE, Length=Diameter) and Kneading disc at 60° pitch (K 60° staggering angle, Length=Diameter/4).

In addition to the material type and the screw configuration, three input variables, namely, screw speed, liquid to solid (L/S) ratio and powder feed rate, were investigated in this research paper. These variables with their levels are listed in Table 1. Although, there exist several input parameters that may have an influence on the granulation process, the ones listed in Table 1 are the most crucial parameters for the TSG process [6]. It is worth emphasising at this stage that the levels of the granulation inputs for each material are different, this due to the different properties and behaviour of the two powders examined in this research work. A statistical linear correlation analysis (i.e. Pearson correlation coefficient) was implemented between the examined input parameters and the granules size that was represented using the main three diameters (i.e. D_{10}, D_{50} and D_{90}). In general, Pearson correlation coefficient is a value in the range of -1 to 1 that usually describes the strength and direction of a linear relationship between two variables. Correlation values among the investigated parameters and the three diameters were obtained and these are shown in Table 2. Some of the parameters have different correlation coefficients between the two types of materials. For instance, the relationships between the powder feed rate and the three diameters for Mannitol are stronger than the same relationships for MCC. The analysis of variance, as a parametric statistical model that is commonly used to compare data sets, showed that the material type and the screw configuration have significant effects on the three diameters, where the P-values (i.e. the observed significance levels) are less than 0.05.

Figure 2 shows a sample of the data distributions (i.e. data density that represents the number of experiments) in the space investigated using two variables at a time. Developing a data-driven model for the TSG process using such a data set can represent a ‘tricky’ exercise because of the nonlinear behaviour, uncertainties and the sparse and limited amount of data presented in Figure 2.

Table 1. The input parameters of the TSG process.

Table 2. The correlation coefficients.

Figure 2. Data density for the screw speed with the L/S ratio.

3. Modelling Framework Based on Fuzzy Logic

Recently, artificial (computational) intelligence has changed the way researchers in both industry and academia think, and has also allowed for utilizing computer systems in several areas. In the pharmaceutical industry, this means a shift from the traditional modelling and control paradigms that can be too complex to derive or simply does not exist, towards data based paradigms that depend on mimicking the human way of thinking and analysing observed/collected data [2]. Various data based models have hitherto been successfully utilized and applied in many research areas [13-14, 16, 26]. These paradigms include, but not limited to, regression modelling approaches and ANNs. Some of these paradigms (e.g. linear regression models) are, in fact, incapable of embedding the sophisticated nonlinear relationships for complex processes. In addition, some of them (e.g. ANNs) are considered as black-box models due to lack of interpretability [22]. Therefore, fuzzy logic has found its way into many industrial and academic applications, in order to develop a simple transparent/interpretable model [27]. Furthermore, such a system can efficiently characterise uncertainties intrinsically.

Figure 3. The FLS structure.

In general, FLS framework, which usually consists of four elements, is depicted in Figure 3. As shown in the figure, the FLS usually maps a crisp input space (**x**) to a crisp output space (**y**). First, the fuzzifier maps crisp inputs to fuzzy input sets with their membership functions. In this research paper, the membership function is presented in the form of the

Gaussian function. The reason behind such a choice can be attributed to the characteristics of such a function (i.e. continuity and smoothness) that allow the FLS to be utilized as a *universal approximator*. The Gaussian membership function ($\mu_j^i(x_j)$) can be expressed as follows [6]:

$$\mu_j^i(x_j) = \exp\left[-\frac{1}{2}\left(\frac{x_j - m^i}{\sigma^i}\right)\right] \tag{1}$$

where the parameters m^i and σ^i stand for the mean and the standard deviation of the i^{th} set, respectively, and the x_j is the j^{th} input. The fuzzification step is usually performed to consider the scenarios of uncertainties; by uncertainty in the granulation process, in particular, one means not only measurement uncertainties but also uncertainties that result from the heterogeneous distributions of the granules porosity. In this research work, the singleton fuzzification is utilized to simplify the model without any loss of generality [28].

The linguistic representation of the process under investigation is presented by rules. Such rules can be either extracted from a data set or provided by process experts (hand-crafted). It is worth emphasizing at this stage that, in this research paper, the rule base is extracted from a set of laboratory-scale experiments that were conducted using different operating settings and material types. The fuzzy inputs are integrated with the rules by the fuzzy inference engine, which is considered to be the heart of the FLS. Fuzzy basis functions ($\phi_i(\mathbf{x})$) are commonly used to mathematically represent such an integration. This function can be written as follows [27]:

$$\phi_i(\mathbf{x}) = \frac{\prod_{j=1}^{n} \mu_j^i(x_j)}{\sum_{i=1}^{R}\prod_{j=1}^{n} \mu_j^i(x_j)} \tag{2}$$

where **x** is the input vector, R is the number of rules and the other parameters are as defined previously in Equation (1). For a FLS that has n inputs ($x \in \mathbf{x}$) and one output ($y \in \mathbf{y}$), the i^{th} rule can be linguistically expressed as the follows [27]:

Rule^i: IF x_1 is A_1^i … and x_n is A_n^i, **THEN** y is B^i.

where A_j^i represents the j^{th} antecedent membership function of the i^{th} rule, and B^i represents the consequent of the same rule. In this research paper, the Mamdani FLS is considered such that B^i is represented by a membership function. In the Mamdani system, membership functions, which are usually used to express words such as low, high, etc., are represented by fuzzy sets that express the subjective information one may have about the process. These rules give FLSs their interpretability. The output of the inference engine is fuzzy output sets. Since a crisp output is always required for engineering applications, a defuzzification step is finally performed to map the fuzzy output into a crisp value.

All data based models, including FLSs, significantly depend on the process data. Thus, having a good data set, in terms of the number of data points and the distribution of the data in the space under investigation, is very important, in order to obtain a good predictive performance. In the case of the granulation processes, the modelling challenges may emanate from (i) a sparse and limited amount of data, (ii) the highly nonlinear input/output relationships that can result from the interaction among the granulation three mechanisms, and (iii) the strong need to provide a simple understanding of the input/output relationships. Therefore, in this research work, a modelling framework based on a FLS is presented. The proposed modelling framework consists of a number of FLSs that have various structures in terms of the number of rules and the parameters of the fuzzy sets. Therefore, the idea behind the development of such a framework is not only to map the process inputs to the outputs (i.e. granule properties),

but also to extract all the possible informative rules that may not all be easy to extract by a single FLS, this being due to the complementary role that the various FLS structures can play in modelling and capturing the possible patterns of the process under investigation. However, such FLSs may include the similar rules (i.e. highly overlapping rules) in their rule bases, therefore, these highly overlapping rules need to be assessed and reduced to one rule.

Figure 4 illustrates the main steps of the proposed algorithm. First, *M* different FLSs having different structures are developed using the process data in order to obtain a number of rules that can cover all the space areas of the investigated parameters, in particular, those areas where the data are not sufficient. Such a number of rules can be "huge" and, consequently, may result in a computationally expensive model. In addition, some of the rules may be redundant, this being due to the high overlapping between these rules, and, thus, they may lead to a biased model. Therefore, the extracted rules are assessed and, accordingly, reduced to construct a reduced FLS. Various approaches have been presented in the related literature to develop a reduced FLS model by extracting the most important rules from the extracted rule bases [27]. These approaches include, but are not limited to, an orthogonal least-squares method, eigenvalue decomposition and the SVD-QR method [28]. Since the latter has been widely used in different applications, it is utilized, in this research work, to assess the extracted rules. Such an approach can be utilized to assess the rules by several steps starting with identifying the so-called fuzzy basis function matrix (Φ). The latter can be mathematically represented as follows [27]:

$$\Phi = \begin{bmatrix} \phi_1\left(\mathbf{x}^{(1)}\right) & \cdots & \phi_P\left(\mathbf{x}^{(1)}\right) \\ \vdots & \ddots & \vdots \\ \phi_1\left(\mathbf{x}^{(N)}\right) & \cdots & \phi_P\left(\mathbf{x}^{(N)}\right) \end{bmatrix} \tag{3}$$

where $\phi_i\left(\mathbf{x}^{(k)}\right)$ represents the i^{th} fuzzy basis function that was defined previously in Equation (2), and the superscript number is utilized to identify the input vector of the k^{th} data set (i.e. experiment). Once the fuzzy basis function matrix is defined, the singular value decomposition (SVD) matrix, the so-called matrix factorization, of such a matrix can calculated. Such a step is usually followed by estimating the numerical rank of the fuzzy basis functions matrix via examining the calculated singular values. The QR algorithm is, then, utilized to rank-order the basis functions that correspond to the ordered singular values that are larger than a predefined threshold. The size of the resulted matrix is smaller than the original matrix presented in Equation (3). Finally, the denominators of the basis functions defined in Equation (2) need to be re-normalized followed by determining the system parameters [28]. The SVD-QR algorithm can be utilized to identify the dominant and subdominant spans and, thus, it allows one to identify which fuzzy basis functions would contribute the most to the system and how many of them are required. For more in-depth analysis related to such an approach, readers are referred to [27-28]. It is worth emphasising at this stage that such an approach keeps the meaning of the linguistic information that may be expressed as fuzzy rules. In addition, available expert knowledge should be incorporated when available and required.

Figure 4. Flow chart of the presented algorithm (*M* is the maximum number of the FLSs counted by m, I is the maximum number of rules counted by i, and SD stands for steepest descent algorithm).

4. Results and Discussion

The algorithm relating to the presented model was implemented to predict the granules size and to represent the TSG process in a linguistic way that can be easily understood. In order to elicit such a multi-input single-output (MISO) model, different FLS structures were developed in order to obtain a number of rule bases. For each FLS, the data collected (i.e. 26

experiments that were carried-out under different operating conditions and material types) were randomly divided into two sets: training (70%) and testing (30%). The former set is utilized to identify the input/output relationships by extracting informative rules and the latter set is used to assess each model's generalization capabilities. Partitioning the available data into such two sets can significantly affect the performance of the FLS, in particular when the available data are limited. By data partitioning one means both the number of data points in both sets (i.e. training and testing data sets) and the data distribution in the space investigated. Various partitioning approaches, including the well-known k-fold cross-validation with different partitioning ratios, were utilized in this research paper. It was found that both the 5-fold cross-validation and the random partitioning (i.e. training (70%) and testing (30%)) were the best approaches in order to develop a predictive FLS with a reliable performance. Since there was no significant difference in the predictive performance values between the two models that were developed using the aforementioned approaches (i.e. both the 5-fold cross-validation and the random partitioning), the random partitioning was selected as it was relatively faster. It is worth mentioning that the numbers of data points in the training and testing sets were the same for all FLSs that were developed but the data distributions in the space under investigation were different. For this reason, the FLSs can provide various rule bases that can better represent the TSG process. Since some of the TSG input variables are continuous (e.g. screw speed, L/S ratio and powder feed rate) and some of them are discrete (e.g. the material type and screw configuration), the model was developed in a way that the discrete variables were modelled as crisp ones.

For a pre-defined number of rules, the parameters for each model (e.g. mean and standard deviation) were initialized using hierarchical clustering algorithm [29]. These parameters were then optimized by utilizing the steepest descent (SD) algorithm [27]. Overfitting is a well–known phenomenon in machine learning. It occurs when a model learns

about the data too well, and as a result, the model fails to generalise to all the data patterns [30]. Since real experimental data always contain unavoidable noise (e.g. imprecise measurements or the insignificant effect of uncontrollable factors), the model structure needs to be developed in such a way that allows it to generalise as much as possible. For this, the SD optimization algorithm, was terminated when the maximum pre-defined number of epochs (i.e. one epoch refers to one training cycle that includes one forward and one backward pass) was reached or when the error was smaller than a threshold value. In the fuzzy logic system, the cardinality of the fuzzy sets and rules that are greater than the optimal, may also lead to overfitting [30]. In this research paper, various cardinalities for the fuzzy rules were assigned to the M different FLSs that were first developed, these being in the range of 1 to 18. This number (i.e. 18) is also the number of training experiments (i.e. each experiment is considered as one rule). The number of fuzzy rules that was finally selected was the one that ascertained a trade-off between a good generalization and training capabilities, in other words, it was the one that led to the minimum error residuals and to a small difference between the training and testing error values. It is worth emphasizing at this stage that the error was evaluated via the root mean square error (RMSE). The SVD-QR method was then utilized to assess the extracted rules that resulted from the various FLSs developed and to identify the ones that contribute the most to the system. Consequently, these rules were assessed and, accordingly, reduced. Reducing the number of these rules reduced the computational burden and avoided the likelihood of system redundancy.

Figure 5. The developed FLS for the volume fraction (%.µm^{-1}) for the size class (1188-1232µm): (a) Training (R^2=0.91), (b) Testing (R^2=0.91) (with 10% bands).

In addition to the results summarized in Table 3 for the D values and the overall size, the model performance for the volume fraction for the size class (1188-1232µm), as an example, for both the training and the testing data sets using 7 rules is shown in Figure 5, with a coefficient of determination (R^2) value (training, testing) = [0.91, 0.92]. Such a size class was

selected to show that the proposed model can predict the volume fraction in those classes where some of the experiments showed a bi-model behaviour. The RMSE values for the training and testing sets are 0.005 and 0.005, respectively. Similarly, Figure 5 shows that most of the predicted values lie within a 90% confidence interval. As an example, Figure 6 shows a sample of the fuzzy rule base. The rules presented in such a figure can be linguistically expressed as follows:

Rule 1: ***IF*** *the material is Mannitol and screw speed is small and screw configuration consists of conveying and 32 kneading elements and L/S ratio is small and powder feed rate is small,* ***THEN*** *the volume fraction is medium.*

Rule 2: ***IF*** *the material is MCC and screw speed is medium and screw configuration consists of conveying and 16 kneading elements and L/S ratio is high and powder feed rate is small,* ***THEN*** *the volume fraction is high.*

Rule 3: ***IF*** *the material is Mannitol and screw speed is high and screw configuration consists of conveying elements and L/S ratio is small and powder feed rate is small,* ***THEN*** *the volume fraction is high.*

Figure 6. The rule base of the reduced FLS for the volume fraction (%.μm^{-1}) for the size class (1188-1232μm).

It is worth mentioning also that the third rule presented in Figure 6 was extracted by only two FLSs, this being due to the limited data points available in such a region. Such a rule was kept and retained by the SVD-QR method. The reason behind this can be attributed to the significant contribution of such a rule to the reduced FLS model. To illustrate this further, eliminating such a rule from the rule base of the reduced FLS would have resulted in a lower predictive performance (approximately 11% lower). Figure 7 shows examples of response surfaces using two variables at a time. As can be observed in the response surfaces of such a

figure, the volume fraction is represented as highly non-linear functions of the screw speed and powder feed rate. In addition, it is noticeable that the volume fraction for the size class (1188-1232μm) has reached the saturation point. To illustrate, increasing the screw speed (more than 600rpm) does not have an effect on the volume fraction for this size class, which indicates that there is a balance between the "growth and consolidation" mechanism and the "breakage" one. One can also notice such a balance at different L/S ratios. The previous behaviour was observed for the volume fraction values of the majority of the size classes in the size range (900-2000μm). For the powder feed rate, a different behaviour can be observed; at small value of the powder feed rate, the volume fraction value has reached a saturation point when the screw speed is in the range of 400rpm to 600rpm, then it decreases and then increases as the screw speed increases (more than 600rpm). However, at higher values of powder feed rate, the volume fraction value increases as the screw speed increases, as shown in Figure 7.

Figure 7. The response surface for the volume fraction (%.μm^{-1}) for the size class (1188-1232μm).

In a similar manner, the volume fraction values for the other size classes were predicted using the presented model. It is worth mentioning that different numbers of rules (i.e. fuzzy basis functions) were used for the various size classes, such numbers were in the range of 3 to 9. The models that were developed to predict the volume fraction values for the large size classes (larger than 3500μm) had, in general, a lower number of rules when compared to smaller size classes, this being due to the fact that, for most of the experiments, the volume fraction values for large size classes are zeros (or very small), thus three rules were sufficient to achieve a good predictive performance. As examples, the predicted (o) and the experimental (*) size distributions for two experiments, which were carried-out using different materials and under different operating conditions, are shown in Figure 8. It is worth emphasizing that the volume fraction values for each size class presented in such a figure was predicted using a

MISO model. It is apparent that the predicted values for all size classes are close to the experimental ones. This proves the effectiveness and efficiency of the presented FLS framework in representing the TSG process efficiently.

Figure 8. The reduced FLS: the predicted (o) and the experimental (*) distributions for the size (a) using the MCC powder, screw speed=800rpm, screw configuration consists of conveying and 32 kneading elements, L/S ratio=1.25 and feed rate=1Kg/h; (b) using Mannitol powder, screw speed=200rpm, screw configuration consists of conveying elements, L/S ratio=0.3 and feed rate=1Kg/h.

Although the material type is considered as an input parameter, the two powders examined in this research work behaved differently. For instance, Figure 9 shows the surface responses for one of the granule size descriptors (i.e. D_{50}) with the L/S ratio and powder feed rate for the granules produced using the MCC and Mannitol powders. It is noticeable that the powder feed rate has insignificant effect on D_{50} for the granules produced from the Mannitol powder. However, it has a relatively considerable effect on D_{50} for the granules produced from the MCC powder. To illustrate, for the MCC powder, as the powder feed rate increases the D_{50} value increases at both high and low values of the L/S ratio, as shown in Figure 9 (a). It is also unexpectedly noticeable that the D_{50} value is in the range of 900μm to 1000μm (i.e. a medium range) when the L/S ratio is high (approximately 1.25), this can be attributed to the various screw configurations investigated in this research, as explained in Section 2, and not shown in this three-dimensional response surface. In addition, for the Mannitol powder, as the powder feed rate increases, the D_{50} value remains almost the same at a specific L/S ratio, as shown in Figure 9 (b). Unlike the MCC powder, it is obvious that increasing the L/S ratio in the range of 0.2 to 0.25 has no significant effect on the D_{50} value for Mannitol powder, however, increasing it in the range of 0.25 to 0.3 has considerably increased the D_{50} value. Likewise, one can clearly see from this figure that the function that describes the relationship between the

D_{50}, as an output, and the L/S ratio and the powder feed rate, as inputs, is highly nonlinear for the MCC powder, whereas such a function looks simpler for the Mannitol powder.

Figure 9. The response surface for D_{50} with the L/S ratio and feed rate for: (a) MCC and (b) Mannitol.

For comparison purposes, the common FLS framework was implemented to represent the TSG process using the available data set. For the volume fraction for the size class (1188-1232μm), the performance values for such a model measured via the RMSE (training, testing) and R^2 (training, testing) are [0.009, 0.008] and [0.75, 0.63], respectively, as shown in Figure 10. Obviously, the performance measures of the presented FLS framework are better than the ones obtained by the known FLS. For comparison purposes, the well-known ANN was also applied to model the TSG process using the same data. The performance of such a network and the FLS ones, both the proposed and the traditional FLSs, for the granule size, which is commonly represented by the three main diameters/descriptors, measured via the R^2 and RMSE values are presented in Table 3, also the performance of the neural network for the size class (1188-1232μm) is shown in Figure 11. In Table 3, one can notice that the RMSE values for the testing set for some ANNs are higher than the ones for the training set, this can be an indication of a training issue. Such an issue was investigated further and in more depth and it was found that this issue is due to the different values of the output modelled. From Figures 10 and 11, and Table 3, it is apparent that the structure presented in this research paper outperformed the standard FLS and ANN models, with an overall improvement of approximately 16% and 29% in R^2, respectively, in addition to being transparent and interpretable. Therefore, the FLS framework presented can be successfully used to linguistically represent the TSG process, as shown in Figure 6, and to predict the granules size accurately, as shown in Figures 5 and 8, and Table 3. The reason behind this can be attributed to the proposed model structure. To illustrate this, the presented model has a number of FLSs with different structures that can capture the

granulation input/output relationships because of the number of fuzzy basis functions included. Furthermore, the FLSs with various structures can usually play a complementary role in representing and modelling the process possible patterns. Although the SVD-QR algorithm is utilized in the modelling structure to reduce the number of fuzzy rules and the fuzzy basis functions that were extracted by the FLSs, the fuzzy rules and the basis functions are assessed carefully before eliminating any of them, thus, the negative effect on the predictive modelling performance of the SVD-QR algorithm can be negligible.

Figure 10. The traditional FLS for the volume fraction (%.µm-1) for the size class (1188-1232µm): (a) Training (R^2=0.75), (b) Testing (R^2=0.63) (with 10% bands).

Table 3. The performances of the models represented by RMSE and R^2.

Figure 11. The artificial neural network for the volume fraction (%.µm-1) for the size class (1188-1232µm): (a) Training, (b) Testing (with 10% bands).

In summary, a transparent but accurate data-driven predictive model was successfully developed to represent the TSG process mathematically and linguistically. The elicited model has the potential to lead to significant impact on the granulation process and the related industries, and various other equally challenging processes with limited amount of data. For instance, in addition to providing a simple linguistic understanding of the granulation process that can be utilized by users to control the process, the proposed model can accurately predict the whole size distribution of the granules produced by exploring all the areas in the space investigated, in particular, those areas where the amount of data is very conservative and/or is simply sparse. Therefore, such a model can be used on a relatively large-scale, if trained carefully, through the product development process. Such a development can positively affect those industries where the granulation process is one of the unit operations in their production lines by, for instance, minimizing the recycling ratio, improving the supply chain management and enhancing the company competitiveness.

4. Conclusions

Predicting the granule properties is a key step towards the control of the granulation process, in particular, in those industries where such a process can be identified as one of the key unit operations in their production lines. However, modelling and predicting the granulation process and the granule properties are by no means trivial exercises; this is due to the fact that (i) such a process is surrounded by uncertainties that can impact on the performance of some models (e.g. an artificial neural network), and (ii) the sparse and limited amount of available data. In this research work, a new interpretable predictive modelling framework that is based on fuzzy logic was proposed to model the twin screw granulation (TSG) process. Various Fuzzy Logic Systems (FLSs) having different structures were first developed to (i) extract linguistic and informative rules that cover all the areas in the investigated space, and (ii) capture the complex relationships between the granulation inputs and outputs. Since the number of the extracted rules from such structures can be large, they were assessed to identify the ones that contribute the most to the system and, consequently, reduced using the singular value decomposition-QR factorization (SVD-QR) algorithm. The main idea behind the SVD-QR algorithm is to identify the dominant and subdominant spans, thus, it allowed one to determine the contribution of the various fuzzy basis functions to the system. Finally, the reduced FLS was implemented to represent the TSG process, to provide a simple but effective understanding of such a process and to predict concomitantly the granule size distribution accurately. The experimental results showed that the new proposed model (i) predicted the size distribution successfully, (ii) provided a simple linguistic understanding of the TSG process, and (iii) outperformed the well-known FLS and the artificial neural network (ANN), with an overall improvement of approximately 16% and 29% in R^2, respectively.

References

[1] A. Rogers, A. Hashemi, M. Ierapetritou, Modeling of particulate processes for the continuous manufacture of solid-based pharmaceutical dosage forms, Processes, 1 (2013) 67-127.

[2] W.H. AlAlaween, M. Mahfouf, A.D. Salman, Predictive modelling of the granulation process using a systems-engineering approach, Powder Technology, 302 (2016) 265-274.

[3] H. Yu, J. Fu, L. Dang, Y. Cheong, H. Tan, H. Wei, Prediction of the particle size distribution parameters in a high shear granulation process using a key parameter definition combined artificial neural network model, Industrial and Engineering Chemistry Research, 54 (2015) 10825-10834.

[4] D. Barrasso, A. Hagrasy, J.D. Litster, R. Ramachandran, Multi-dimensional population balance model development and validation for a twin screw granulation process, Powder Technology, 270 (2015) 612-621.

[5] D. Djuric, P. Kleinebudde, Impact of screw elements on continuous granulation with a twin-screw extruder, Journal of Pharmaceutical Sciences, 97 (2008) 4934-4942.

[6] W.H. AlAlaween, B. Khorsheed, M. Mahfouf, I. Gabbott, G.K. Reynolds, A.D. Salman, Transparent Predictive Modelling of the Twin Screw Granulation Process using a Compensated Interval Type-2 Fuzzy System, European Journal of Pharmaceutics and Biopharmaceutics, 124 (2018) 138-146.

[7] M. Benali, V. Gerbaud, M. Hemati, Effect of operating conditions and physico–chemical properties on the wet granulation kinetics in high shear mixer, Powder Technology, 190 (2009) 160-169.

[8] S. Iveson and J. Litster, Growth regime map for liquid-bound granules, American Institution of Chemical Engineers Journal 44 (1998) 1510-1518.

[9] W.H. AlAlaween, M. Mahfouf, A.D. Salman, Integrating physics with data analytics for the hybrid modelling of the granulation process, American Institute of Chemical Engineers, 63 (2017) 4761-4773.

[10] A. Braumann, M. Goodson, M. Kraft, P. Mort, Modelling and validation of granulation with heterogeneous binder dispersion and chemical reaction, Chemical Engineering Science, 62 (2007) 4717-4728.

[11] S. Iveson, Limitations of one-dimensional population balance models of wet granulation processes, Powder Technology, 124 (2002) 219-229.

[12] D. Barrasso, A. Tamrakar, R. Ramachandran, A reduced order PBM-ANN model of a multi-scale PBM-DEM description of a wet granulation process, Chemical Engineering Science, 119 (2014) 319-329.

[13] S.S. Miriyala, P. Mittal, S. Majumdar, K. Mitra, Comparative study of surrogate approaches while optimizing computationally expensive reaction networks, Chemical Engineering Science, 140 (2016) 44-61.

[14] S.S. Miriyala, V.R. Subramanian, K. Mitra, TRANSFORM-ANN for online optimization of complex industrial processes: Casting process as case study, European Journal of Operational Research, 264 (2018) 294-309.

[15] H. Ghalkhani, S. Golian, B. Saghafian, A. Farokhnia, A. Shamseldin, Application of surrogate artificial intelligent models for real-time flood routing, Water and Environment Journal, 27 (2013) 535-548.

[16] M. Alshafiee, W.H. AlAlaween, D.Markl, M. Soundaranathan, A. Almajaan, K. Walton, L. Blunt, K. Asare-Addo, A predictive integrated framework based on the radial basis function for the modelling of the flow of pharmaceutical powders, International Journal of Pharmaceutics, in press.

[17] J. Westerhuis, P. Coenegracht, C. Lerk, Multivariate modelling of the tablet manufacturing process with wet granulation for tablet optimization and in-process control, International Journal of Pharmaceutics 156 (1997) 109-117.

[18] Y. Miyamoto, S. Ogawa, M. Miyajima, M. Matsui, H. Sato, K. Takayama, T. Nagai, An application of the computer optimization technique to wet granulation process involving explosive growth of particles, International Journal of Pharmaceutics 149 (1997) 25-36.

[19] S. Shirazian, M. Kuhs, S. Darwish, D. Croker, G.M. Walker, Artificial neural network modelling of continuous wet granulation using a twin-screw extruder, International Journal of Pharmaceutics, 521 (2017) 102-109.

[20] S.S. Behzadi, J. Klocker, H. Huttlin, P. Wolschann, H. Viernstein, Validation of fluid bed granulation utilizing artificial neural network, International Journal of Pharmaceutics, 291 (2005) 139-148.

[21] S. Watano, Y. Sato, K. Miyanami, Application of a neural network to granulation scale-up, Powder Technology 90 (1997) 153-159.

[22] C. Bishop, Pattern Recognition and Machine Learning, Springer, New York, 2006.

[23] C. Bishop, Neural Networks for Pattern Recognition, Oxford: Clarendon Press, New York, 1995.

[24] S. Watano, K. Miyanami, Control of granulation process by fuzzy logic, 18th International Conference of the North American Fuzzy Information Processing Society, New York, USA (1999).

[25] H. Liu, S. Galbraith, B. Cha, Z. Huang, S. Park, S. Yoon, Foundations of Computer Aided Process Operations/Chemical Process Control, Arizona (2017).

[26] H.J. Kim, M. Mahfouf, Y.Y. Yang, Modelling of hot strip rolling process using a hybrid neural network approach, Journal of Materials Processing Technology, 201 (2008) 101-105.

[27] J.M. Mendel, Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions, Prentice Hall, 2001.

[28] G.C. Mouzouris, J.M. Mendel, Non-Singleton Fuzzy Logic Systems: Theory and Applications, IEEE Transactions on Fuzzy Systems, 5 (1997) 56-71.

[29] Q. Zhang, M. Mahfouf, A hierarchical Mamdani-type fuzzy modelling approach with new training data selection and multi-objective optimisation mechanisms: A special application for the prediction of mechanical properties of alloy steels, Applied Soft Computing, 11 (2011) 2419-2443.

[30] P.A. Wałega, Overfitting problem in a virtual sensor obtained with W–M method, Procedia Computer Science 35 (2014) 54-62.

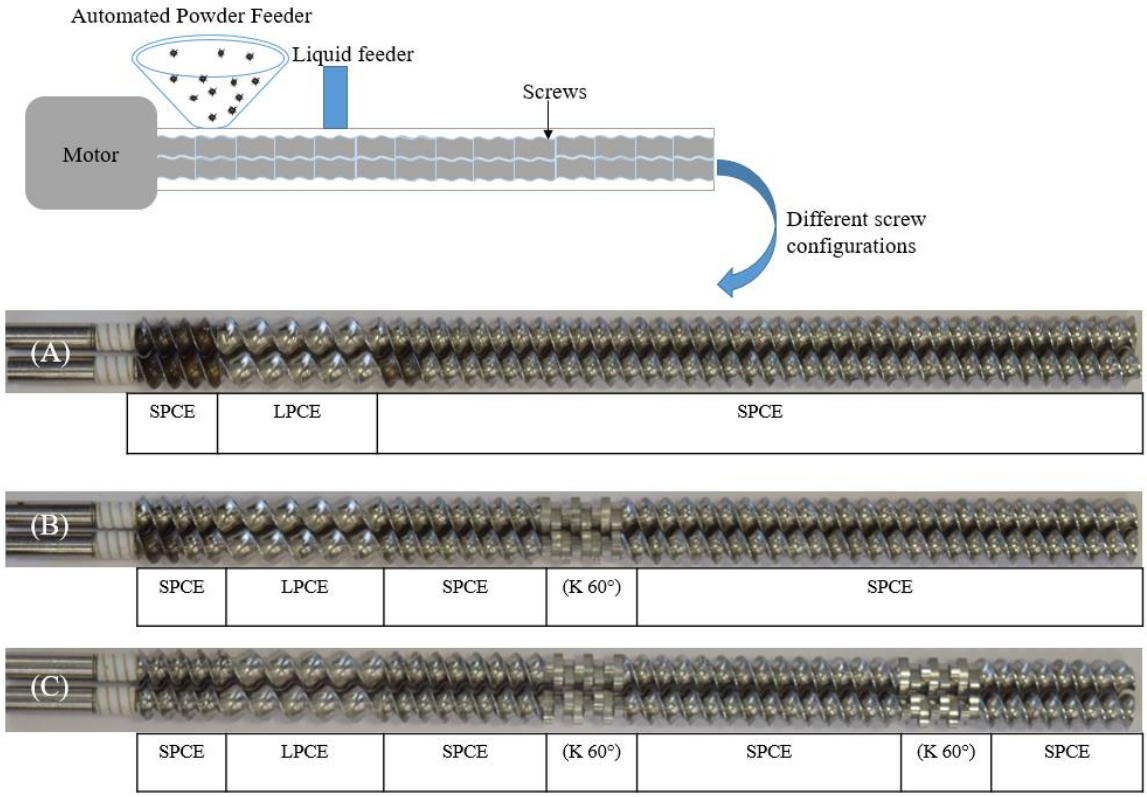


Figure 1. Schematic diagram for the twin screw granulator and the screw configurations: Long pitch conveying element (LPCE, Length=2Diameter), Short pitch conveying element (SPCE, Length=Diameter) and Kneading disc at 60° pitch (K 60° staggering angle, Length=Diameter/4).

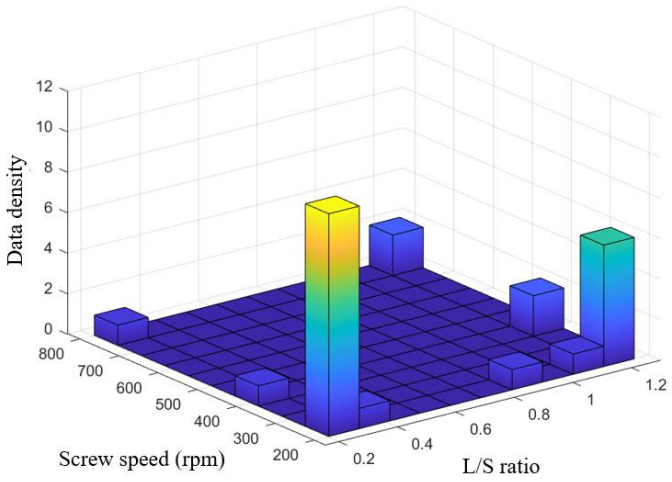


Figure 2. Data density for the screw speed with the L/S ratio.

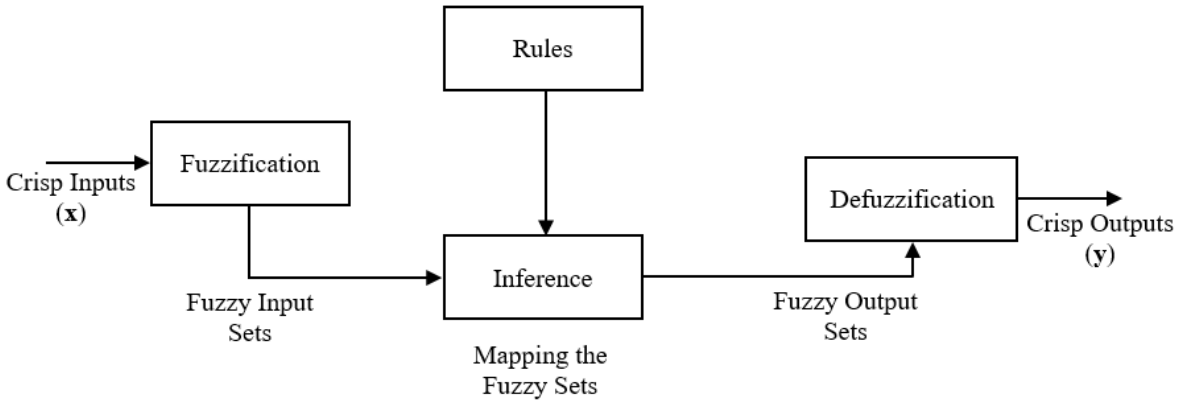


Figure 3. The FLS structure.

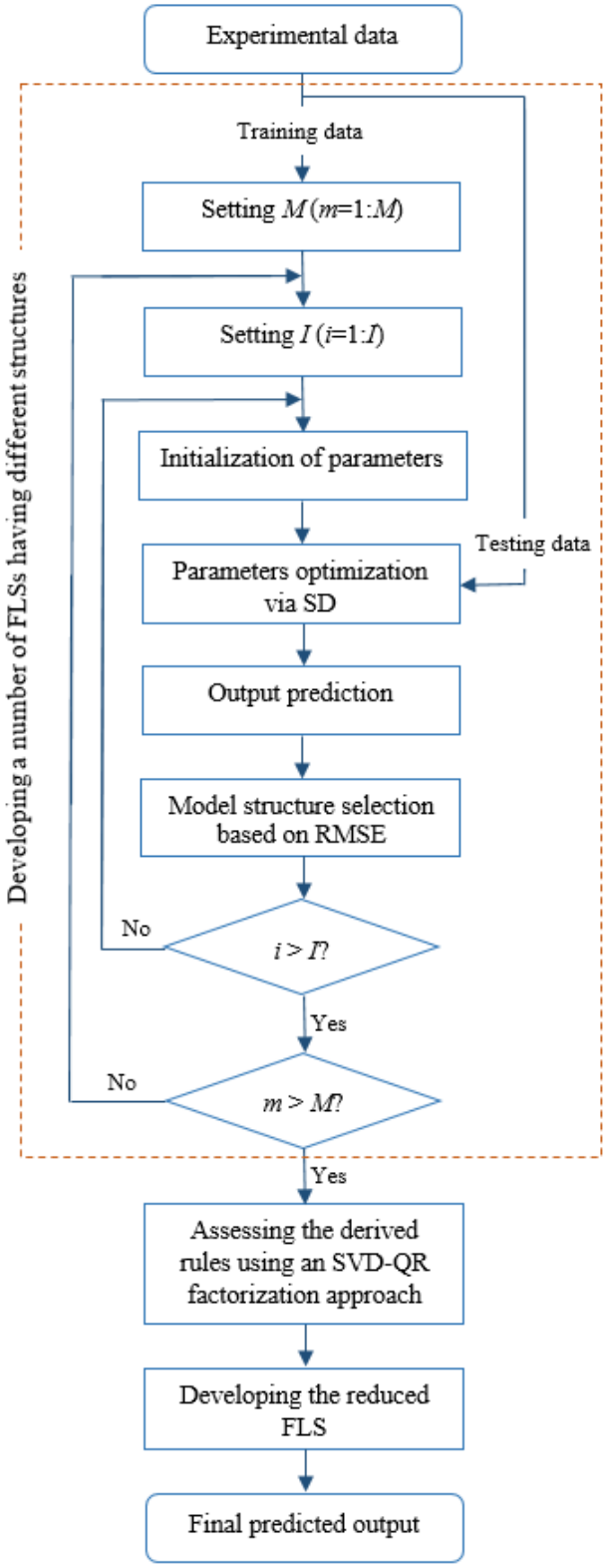

Experimental data
Training data
Setting M (m=1:M)
Setting I (i=1:I)
Initialization of parameters
Parameters optimization via SD
Testing data
Output prediction
Model structure selection based on RMSE
$i > I$?
No
Yes
$m > M$?
No
Yes
Developing a number of FLSs having different structures
Assessing the derived rules using an SVD-QR factorization approach
Developing the reduced FLS
Final predicted output

Figure 4. Flow chart of the presented algorithm (M is the maximum number of the FLSs counted by *m*, I is the maximum number of rules counted by *i*, and SD stands for steepest descent algorithm).

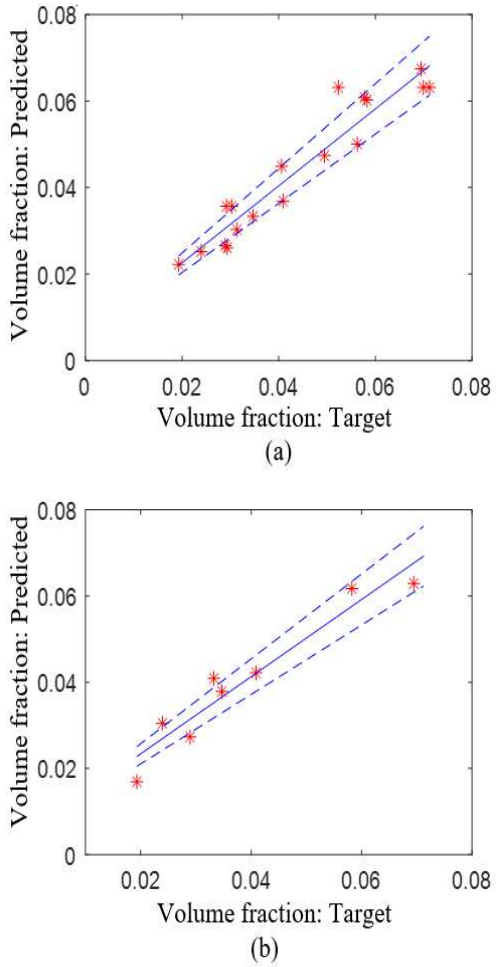
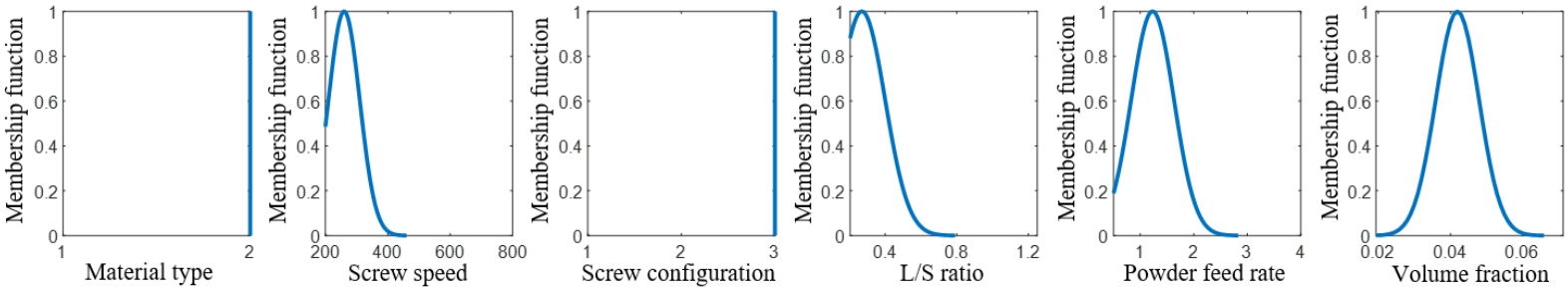
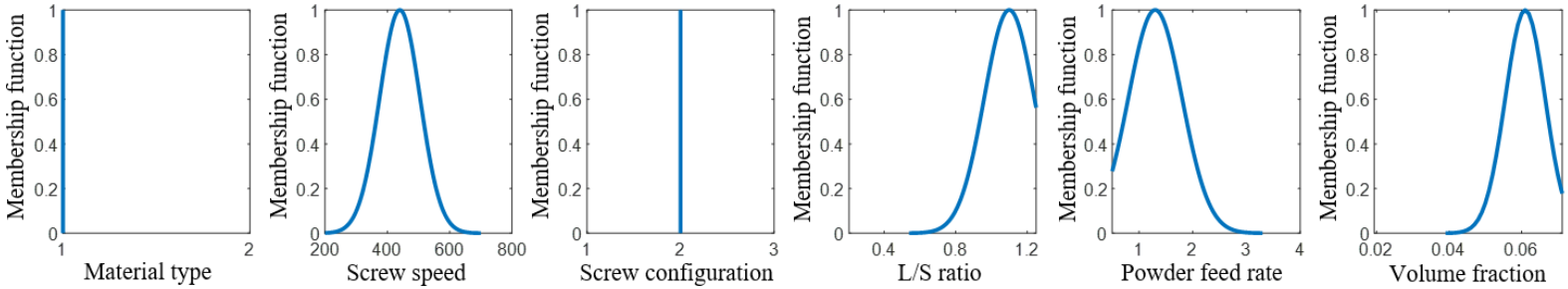


Figure 5. The developed FLS for the volume fraction (%.μm^{-1}) for the size class (1188-1232μm): (a) Training (R^2=0.91), (b) Testing (R^2=0.91) (with 10% bands).

Rule 1:



Rule 2:



Rule 3:

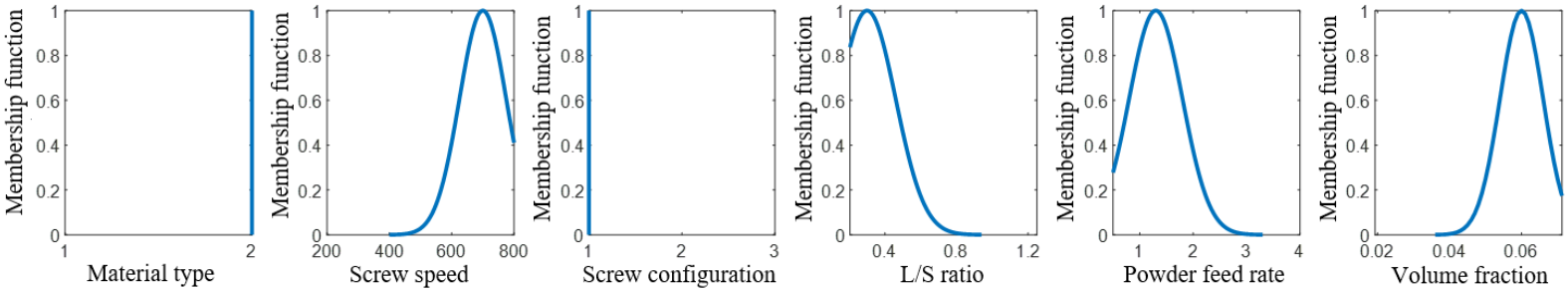


Figure 6. The rule base of the reduced FLS for the volume fraction (%.μm^{-1}) for the size class (1188-1232μm).

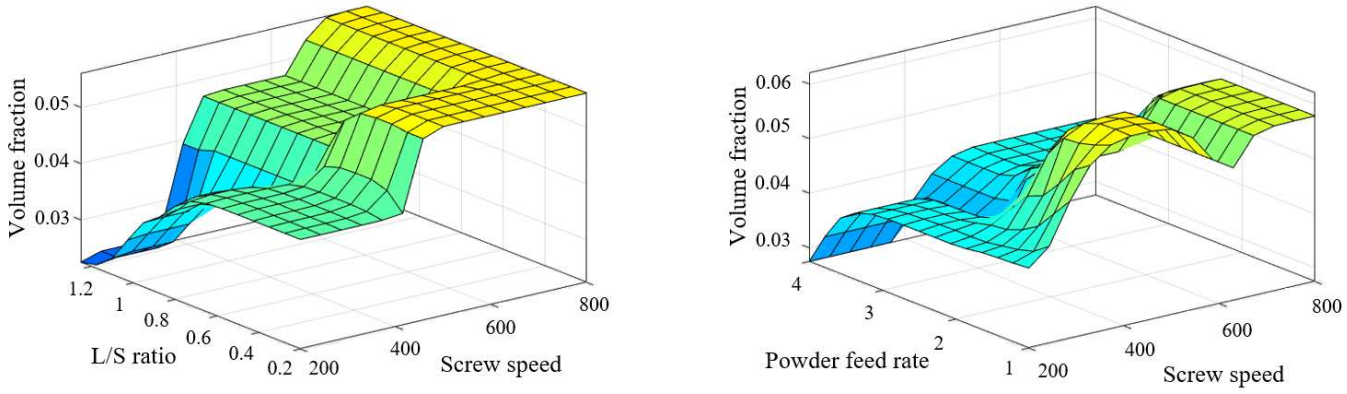


Figure 7. The response surface for the volume fraction (%.µm^{-1}) for the size class (1188-1232µm).

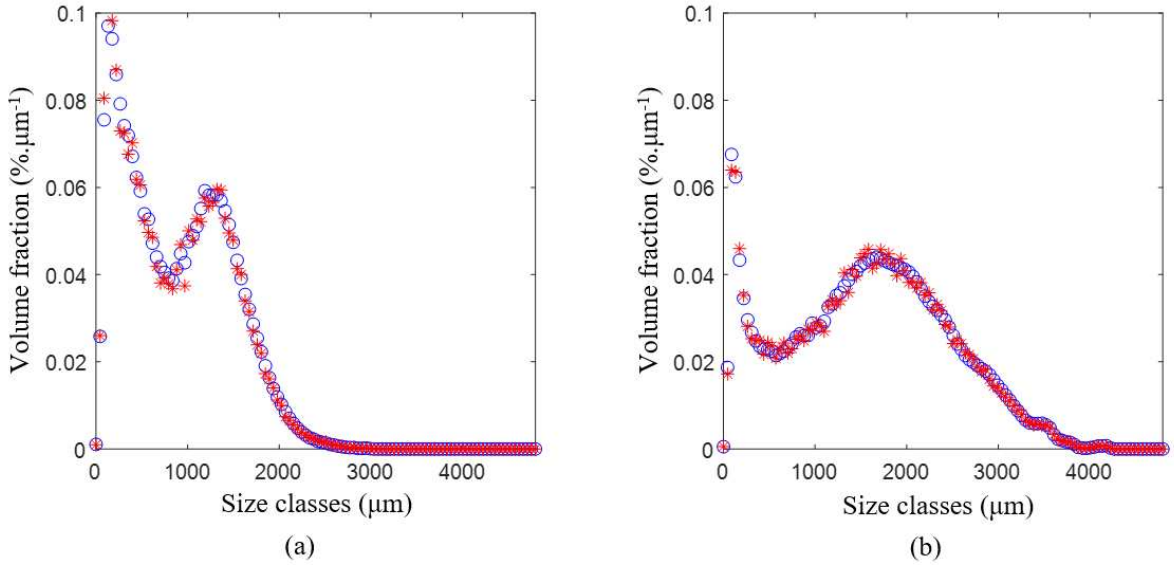
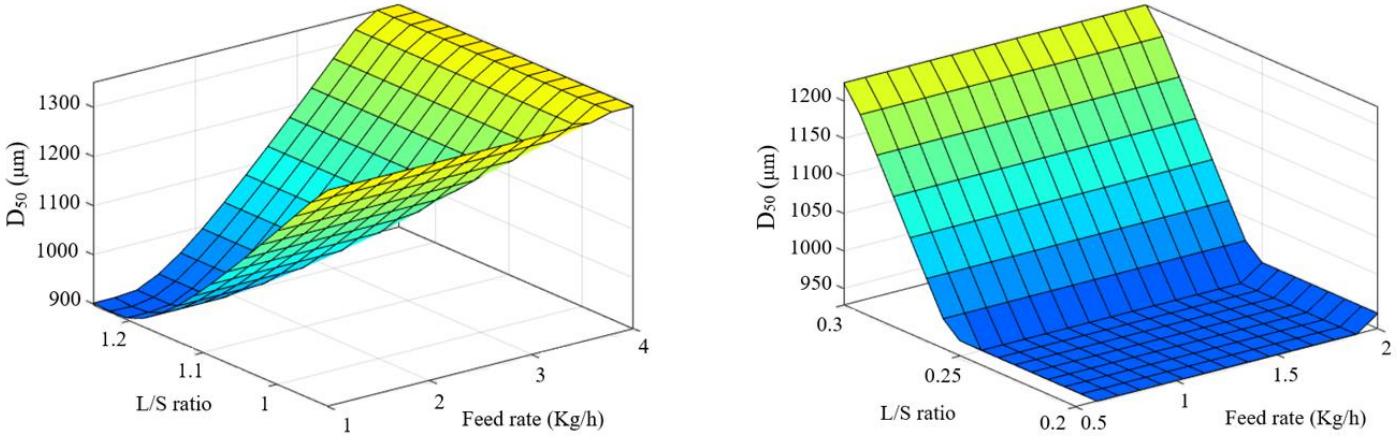


Figure 8. The reduced FLS: the predicted (o) and the experimental (*) distributions for the size (a) using the MCC powder, screw speed=800rpm, screw configuration consists of conveying and 32 kneading elements, L/S ratio=1.25 and feed rate=1Kg/h; (b) using Mannitol powder, screw speed=200rpm, screw configuration consists of conveying elements, L/S ratio=0.3 and feed rate=1Kg/h.



(a) (b)

Figure 9. The response surface for D_{50} with the L/S ratio and feed rate for (a) MCC and (b) Mannitol.

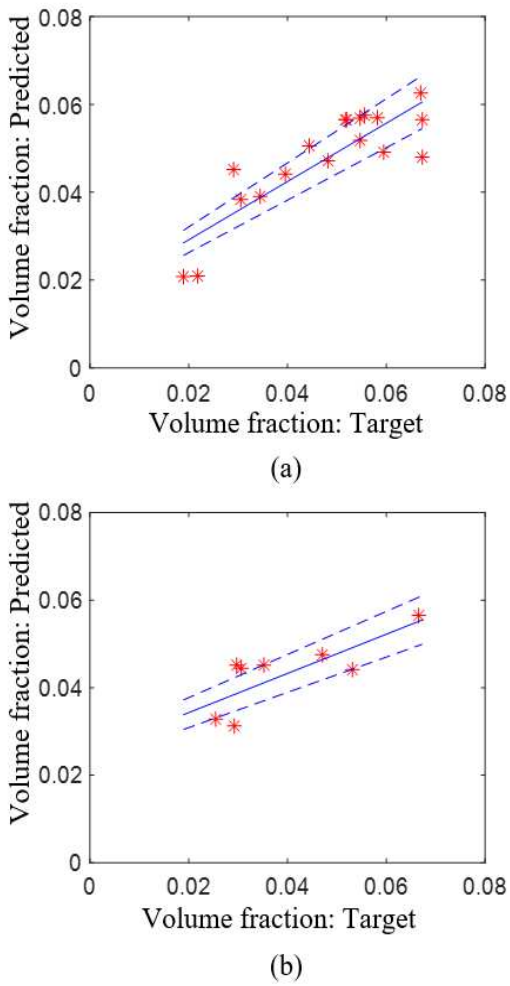


Figure 10. The traditional FLS for the volume fraction (%.μm^{-1}) for the size class (1188-1232μm): (a) Training (R^2=0.75), (b) Testing (R^2=0.63) (with 10% bands).

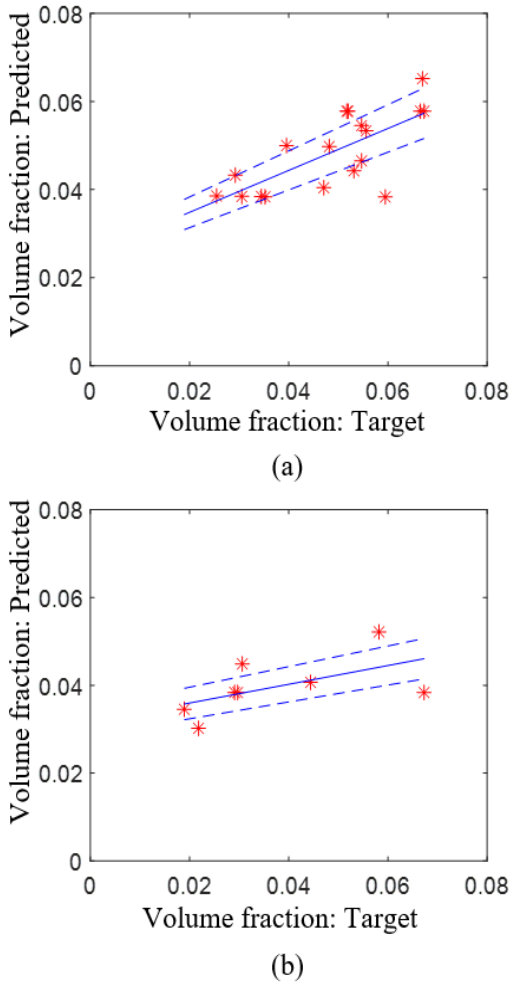


Figure 11. The artificial neural network for the volume fraction (%.μm^{-1}) for the size class (1188-1232μm): (a) Training, (b) Testing (with 10% bands).

Table 1. The input parameters of the TSG process.

Inputs	Inputs' levels	
	MCC	Mannitol
Screw speed (rpm)	200, 400 and 800	200, 400 and 800
L/S ratio	0.93, 1.10 and 1.25	0.20, 0.25 and 0.30
Powder feed rate (Kg/h)	1, 2 and 4	0.5, 1 and 2

Table 2. The correlation coefficients.

Inputs	MCC			Mannitol		
	D_{10}	D_{50}	D_{90}	D_{10}	D_{50}	D_{90}
Screw speed	-0.25	-0.19	-0.19	0.21	-0.10	-0.12
L/S ratio	0.46	-0.32	-0.52	0.20	0.23	0.12
Powder feed rate	0.09	-0.11	-0.10	-0.10	0.53	0.55

Table 3. The performances of the models represented by RMSE* and R^2.

Models	Output	D_{10}		D_{50}		D_{90}		Overall size
		Train	Test	Train	Test	Train	Test	
Artificial neural network	R^2	0.69	0.60	0.63	0.73	0.44	0.59	0.66
	RMSE	116.10	185.60	235.46	389.45	523.52	483.20	0.006
Fuzzy logic system	R^2	0.68	0.71	0.71	0.74	0.70	0.71	0.72
	RMSE	111.32	108.40	205.46	230.54	387.29	388.22	0.005
The proposed structure based on fuzzy logic system	R^2	0.90	0.91	0.92	0.90	0.89	0.87	0.93
	RMSE	49.91	47.73	75.66	83.37	160.24	165.35	0.002

*The RMSE values have different units: for the diameter and the overall all size the units are (μm) and (%.μm^{-1}).