Article:

AlAlaween, W.H., Mahfouf, M. orcid.org/0000-0002-7349-5396 and Salman, A.D. (2021) When swarm meets fuzzy logic: Batch optimisation for the production of pharmaceuticals. Powder Technology, 379. pp. 174-183. ISSN 0032-5910

https://doi.org/10.1016/j.powtec.2020.10.066

When Swarm Meets Fuzzy Logic: Batch Optimisation for the Production of Pharmaceuticals

Wafa' H. AlAlaween[a], Mahdi Mahfouf[b] and Agba D. Salman[c]

[a]*Department of Industrial Engineering, The University of Jordan, Jordan*
[b]*Department of Automatic Control and Systems Engineering, The University of Sheffield, UK*
[c]*Department of Chemical and Biological Engineering, The University of Sheffield, UK*
(E-mail: w.alaween@ju.edu.jo; m.mahfouf@sheffield.ac.uk; a.d.salman@sheffield.ac.uk)

Abstract

The concept of *right-first-time* production is an essential feature for a successful product development process and for companies to be competitive and profitable. However, achieving such a concept is a tricky exercise across a wide spectrum of industrial domains includes the pharmaceutical industry where granulation and tableting processes are considered to be the most critical operations in the production line. Therefore, this research paper presents a new approach that integrates a particle swarm optimization algorithm with a fuzzy logic system in order to implement a new framework by which *right-first-time* production of granules and tablets is ascertained systematically. The proposed approach consists of inverting the models that were previously developed. Through this control technique, one can identify the optimal operating conditions to produce the required granules and tablets, and can minimize the waste and recycling ratios. All frameworks have been successfully validated via real laboratory-scale experiments that include measurement tolerances.

Keywords: Fusion model; Fuzzy logic system; Hybrid model; Incorporated model; *Right-first-time* production.

1. Introduction

Developing a cost-effective product, improving supply chain management, achieving just-in-time inventory, minimizing the waste and recycling ratios and managing a new product development process are the main features that enable companies and industries to be successful, profitable and also retaining leverage in what has become a highly competitive environment. Achieving all the aforementioned features systematically is not a trivial task, particularly for those industrial applications where one can identify the most crucial unit operations that predominantly influence the quality of the final product. In the pharmaceutical industry, for instance, achieving these features is considered to be even a more difficult and challenging task compared to other industries, because of the nature of the final product (i.e. pharmaceutical tablet) and its direct impact on customer health and well-being. Because of its short processing time, mainly due to the fast densification and growth mechanisms, the high shear granulation (HSG) process has been extensively utilized [1]. Although good powder properties are usually obtained and maintained, using such a process makes it difficult to control the production line leading, as a result, to high recycling ratios and inefficient operations [2]. Recently, a good number of books and research papers has been devoted to modelling and control the granulation process [1-4]. For instance, the Quality by Design (QbD) concept has been implemented in order to control the process and the variables affecting it. However, such an implementation was based on experimental studies only [5]. Therefore, developing a model that provides a better understanding of the granulation process and its mechanisms, and predicts the granule properties accurately is necessary to systematically control such a process [6].

Various modelling approaches have hitherto been developed to model the granulation processes [6-9]. These paradigms are either physical based models, data-driven models or models that integrate both of them [6]. Such approaches were utilized to predict the granule properties and to provide a good understanding of the process at different levels. However, to the best of the authors' knowledge, none of these modelling paradigms was exploited in a *right-first-time* framework to systematically control the granulation process by inverting the models that were developed and implemented. The reason for this may be due to the fact that it is not necessarily straightforward to invert nonlinear high dimensional models in order to obtain a unique optimal solution. In this research work, the modelling paradigms; the model presented in [3], which will be referred to as the incorporated model, the hybrid model presented in [6], and the fusion model, which integrates the predictions from the incorporated model with the ones from the hybrid model, as described in [6], are exploited within such a framework to determine the optimal operating conditions for both the granulation and the tableting processes as well as the optimal granule properties to produce a tablet with the desired properties. Since inverting these models may lead to more than one possible set of operating conditions [10], especially when multiple conflicting objective functions need to be taken into consideration, robust optimization algorithms are commonly utilized to define such an optimal set of operating conditions [10].

A plethora of multi-objective optimization algorithms, including but not limited to Evolutionary and Genetic algorithms, have been developed and applied to define a set of optimal solutions, the so-called *Pareto* optimal set, which may also be called non-inferior or non-dominated set [11]. In particular, particle swarm optimization (PSO), as a stochastic population based approach, has been successfully and extensively implemented in many research areas, such as those associated with industrial, academic and medical applications [12-20], this being due to its computational efficiency [11]. Despite the huge body of research that addresses different issues of the multi-objective optimization, it remains a subject of active research. The reason behind this is that there are still many open questions that need to be addressed. One of these issues is the definition of the single optimal solution. Various algorithms have been proposed in the open literature (e.g. the weighted sum method) [21], however, there is hitherto no universally accepted algorithm that can be used to define a single optimal solution for the multi-objective optimization problems (MOPs), instead the concept of *Pareto* optimal is used to propose a set of non-dominated solutions from which the user normally selects one or more pragmatic alternative solutions [11].

In this research paper, a new approach that integrates a multi-objective PSO (MOPSO) algorithm with a fuzzy logic system (FLS) is proposed. First, a MOPSO algorithm is implemented to find the *Pareto* optimal set. Such a choice was motivated by the fact that MOPSO is able to effectively define the optimal solutions for highly nonlinear, mixed integer real-world optimization problems, and also it is considered to be a computationally efficient algorithm (i.e. fast convergence speed and less function evaluations) compared to other algorithms such as Genetic algorithms [12-20]. A FLS is then utilized to determine the membership function values for each solution in the set, the estimation of these values being based on user-defined criteria. The predictive performance of a forward model is considered to be a good criterion for the development of the *right-first-time* modelling framework. This choice is motivated by the fact that the proposed framework will be accurate and reliable only when the predictive performance of the modelling framework is acceptable. In this research paper, the root mean square error (RMSE) and the coefficient of determination (R^2) values, as modelling performance measures, are employed to represent the predictive performance of the forward models. Determining the membership function step is followed by applying the set of fuzzy logic operations to define the single optimal solution, which corresponds to the maximum membership function value.

2. Experimental Work

Calcium Carbonate ($CaCO_3$, D_{50}=85µm) was granulated using the high shear Eirich mixer (1 Litre vertical axis granulator with a top-driven impeller, Maschinenfabrik Gustav Eirich GmbH & Co KG, Hardheim, Germany). Polyethylene Glycol (PEG 1000, with a melting point of approximately 40ºC), as a binder, was added to the moving powder during the granulation process while both the vessel and the impeller were rotating in the same direction (clockwise). It is worth mentioning that the binder was melted, and the powder was pre-heated to approximately 35ºC before the start of the granulation experiment, to make sure that the binder would not solidify before completing the granulation process. The Eirich granulator is equipped with a scrapper and impellers with different shapes, the impellers not being in the centre of the 16cm diameter vessel. Four input parameters were examined, namely, impeller shape (two different impellers were used), impeller speed (from 1000 to 6000 rpm), granulation time (6, 10 and 15 min), and liquid to solid (L/S) ratio (13, 14 and 15%). Such parameters were investigated using a full factorial design of experiments resulting in 108 experiments. It is worth emphasising that the speed of the vessel was kept constant at 170 rpm for all experiments. Once the granulation experiment was completed, the granules were left at room-temperature to allow the binder to solidify. The granules produced were, then, characterized in terms of size, binder content and porosity. The size of the granules was measured using the Retsch Camsizer (Retsch Technology GmbH, Haan, Germany). The porosity and the binder content of the granules were measured for different size classes in the size range (180-2000µm) using a Pycnometer [22] and the method discussed in [23], respectively.

Once the granules produced were classified and characterized, the granules in the size ranges (500-710µm) and (710-1000µm) were compressed using a 10mm diameter set of die and punch to produce tablets. A compression force of approximately 30KN and speed of approximately 0.05mm/min were used during the compression process. All the operating conditions of Instron (3369 dual column, Buckinghamshire, The United Kingdom) were kept constant. Thus, only the effects of the properties of the granules were investigated. Finally, the strength of the tablets produced was measured using Zwick/Roell Z0.5 (Zwick/Roell, Germany). It is worth mentioning at this stage that five tablets were produced from each size class for each experiment, and the average strength value was used in this research.

3. Modelling of the High Shear Granulation Process

3.1. Model Development

Developing a modelling framework for the granulation processes, in particular, the HSG process, is not a trivial task. The reason behind this can be attributed to the complex nature of such a process, and also to the lack of physical representations for its mechanisms [6]. In addition, such a process can be significantly affected by controllable and uncontrollable factors; by uncontrollable factors here one means both the unknown and the very difficult to control factors, these factors may also be interdependent with reverse effects. Therefore, three different new models were previously proposed to predict the properties of the granules produced by the HSG process, and also to provide the necessary understanding of such a process, as presented in [3, 6]. In this research work, these models, which are very effective as what they can be exploited for, will be used to systematically control the granulation and the tableting processes by developing a *right-first-time* approach to produce granules and tablets with the desired properties. Therefore, such models are briefly described in this research paper to help the reader get to grips with the optimization algorithm proposed.

3.1.1 The Incorporated Model

The incorporated model is a data-driven topology, which includes an integrated network and a Gaussian mixture model (GMM) [3]. First, the integrated network is used to represent the complex relationships between the inputs and the outputs of the granulation process. Such a network predicts the granule properties using two phases, as shown in Figure 1. In the first phase, the granulation parameters (x_n) and the target (i.e. experimental) granule properties (y_T) are utilized to train M models, the predicted outputs from these models (y_{P1}, y_{P2}... y_{PM}) and the targets are then utilized to train a single model in the second phase to lead to the final predicted outputs ($\hat{y}_P$). Using an artificial neural network (ANN)-radial basis function (RBF) model in the two phases, the final predicted output can simply be expressed as a combination of composition and superposition of the basis functions as follows [3]:

$$\hat{y}_P = \sum_{k=1}^{K} w_k^{(2)} \phi_k \left(\sum_{i=1}^{I} w_i \phi_i(x) + w_0 \right) + w_0^{(2)} \tag{1}$$

where ϕ_i, w and w_0 are the basis function, the weights and the bias, respectively. The superscript number is utilized to distinguish the parameters of the first phase from the ones of the second phase.

Figure 1. The architecture of the integrated network [3].

The majority of the developed data-driven and physical based modelling approaches, including the integrated network, implicitly assume that the error residuals are normally distributed, which is an invalid assumption in most of the real-world applications. Therefore, the GMM is employed to characterize the error residuals, and, consequently, to improve the performance of the integrated network. Generally, the GMM, as a stochastic model, can be presented as a linear combination of Gaussian functions, where each function has its own parameters (i.e. mean and covariance). The optimal values of these parameters are usually found by maximizing the log likelihood function [24]. The estimated conditional mean can then be added to the predicted output to compensate for the potential bias, as described in [25].

3.1.2 The Hybrid Model

A hybrid model that integrates physical based models with data based modelling topologies was presented previously in [6]. Such a model consists of three separate components, namely; a computational fluid dynamics (CFD) model, a three-dimensional population balance model (PBM) and an ANN-RBF model. The iterative scheme of such a model is depicted in Figure 2. A CFD model is initialized using the granulation input parameters, such as the properties of the materials used and the process operating conditions. The output parameters (e.g. impact velocity) from such a model are then utilized to implement the PBM. Estimating the empirical parameters is also required to implement the PBM. Thus, the ANN-RBF model is developed to estimate these parameters, where each parameter is expressed as a function of the granulation input variables. Once the granule parameters are predicted using the PBM, the size of the granules is then utilized to re-evaluate the parameters estimated by the CFD model. Such a step is followed by re-evaluating the outputs of both the PBM and the ANN-RBF model. These steps are repeated until the difference between two consecutive steps becomes equal to or smaller than a predefined value, or alternatively a satisfactory performance is achieved.

Figure 2. The hybrid model for the HSG process [6].

3.1.3 The Fusion Model

Information fusion is one of the basic concepts of human cognition. Generally, integrating data from different sources should improve the reliability and the degree of accuracy of the information provided to generate the best decisions [26]. Hence, information fusion has been extensively and successfully applied in many areas [26-27]. Various algorithms (e.g. Bayesian inference) for information fusion have been presented in the related literature [27]. However, to develop a more reliable model for information fusion, three different types of uncertainties, namely, uncertainty due to probabilities, uncertainty due to lack of specification and uncertainty due to fuzziness, should be taken into account [27]. Therefore, a fusion model that integrates the Dempster-Shafer (DS) theory with a FLS was previously presented in [6]. Such an approach combines the predicted outputs from the incorporated model with the ones from the hybrid model in order to resolve any potential conflict, and also to improve the prediction performance, in particular, in those areas where the performance was not as good as desired [6].

First of all, the granulation input parameters and the error residuals that result from both the incorporated model and the hybrid one are clustered, followed by determining the membership function value for each data point. In order to integrate the predicted outputs from both models, the DS theory is used, where the mass function values for all the examined hypotheses are estimated using the fuzzy membership function values. The mass function can simply be expressed as follows [27]:

$$m_t = \left(1 - \sum_{\substack{j=1 \\ j \neq t}}^{J_{\max}} \mu_j \times (\eta - \mu_j) \right) \times \mu_t \qquad (2)$$

$$\eta = \arg \max_{1 \le j \le J_{\max}} \mu_j$$

where η is the maximum membership function value, J_{max} is the number of hypothesis, and m_t is the mass function for the t^{th} hypothesis. It is worth mentioning that special cases are usually considered when the number of clusters is either two or three [27]. Then, the hypotheses can be combined by using the Dempster's rule of combination as follows [27]:

$$\begin{aligned} m_{FM} &= \frac{1}{1-K} \sum_{HM \cap IM \neq \phi} m_{HM}\, m_{IM} \\ K &= \sum_{HM \cap IM = \phi} m_{HM}\, m_{IM}\, , \;\; K \neq 1 \end{aligned} \qquad (3)$$

where the subscripts FM, HM and IM are used to distinguish the mass functions for the hypotheses of the fusion, hybrid and incorporated models, respectively. The parameter K, which is a measure of the conflict between two sources, is used to evaluate the normalization factor (1-K). A hypothesis of the incorporated model is commonly merged with the hypotheses of the hybrid model that have better or the same degree of accuracy, and vice versa. It is worth emphasizing at this stage that a high degree of conflict is assumed between a hypothesis and another less accurate one. Consequently, the fusion model should improve the overall predictive performance. Once the combined mass function values are evaluated, the membership function values for the hypotheses of the fusion model can be estimated by solving the set of equation in (2). Finally, the predicted outputs are evaluated using the height deffuzzifier method [28].

3.2 Model Implementation and Results

The models that were briefly described above were used to predict the final granule properties. The incorporated model was employed to represent the granulation process, where an integrated network based on 10 ANN-RBF models having different structures in the first phase, and a single ANN-RBF model in the second phase was developed. For each ANN-RBF model, the final number of RBFs that was selected was the one that led to the minimum error measured via the RMSE index. It is worth mentioning that the scaled conjugate gradient (SCG) algorithm was utilized for training the ANN-RBF models. Once the outputs of the integrated model were estimated, the GMM was employed to characterize the error resulted from the integrated network and improve the predictive performance. The average R^2 and RMSE values of the incorporated model were summarized in [3]. In addition to the incorporated model, the hybrid model briefly described in Section 3 was also employed to represent the granulation process, where the flow of the granules was first studied using the CFD model developed using

ANSYS software (ANSYS Inc., US, Release 16.1). In such a model, the gas-solid flow was investigated be employing a two-fluid model inspired from the kinetic theory behind the granular flow model. In such a model, the materials properties were chosen to simulate the properties of air and granules produced using 500gm of $CaCO_3$ and three mass values of PEG 1000. The rest of the parameters were selected to simulate the actual experiments as described in Section 2 (e.g. the vessel speed was kept constant at 170rpm clockwise). Once the CFD model converged, a three-dimensional PBM was developed by discretising the three-dimensional population into a number of bins represented as finite volumes. The parameters of the PBM (e.g. aggregation kernel) were estimated to match the experimental results, then, they were mapped to the granulation inputs by using a single ANN-RBF model. Using the estimated parameters, the properties of the granules were predicted, and utilized to update the parameters of the CFD model. The steps of the hybrid model were repeated until the difference between the predictions for two consecutive steps became very small. Finally, the fusion model was utilized to combine the predicted outputs of the hybrid model with that of the incorporated model. Such a task was performed by, first, clustering the input variables and the error residuals of both the hybrid and the incorporated models into a number of clusters using the K-means clustering algorithm. Such a step was followed by determining the parameters of each cluster [6]. Since the number of clusters is subjective and depends significantly on the process investigated, in this research work, the best number of clusters that was selected was the one that led to the minimum RMSE (i.e. the maximum predictive performance). The membership functions were then determined and the predicted outputs of the hybrid and incorporated models were combined using Equations (2) and (3). It is worth mentioning that the set of Equations in (3) was solved numerically, since the analytical solution was computationally expensive, particularly, when a large number of clusters was used. The average R^2 and RMSE values of the hybrid and fusion models were summarized in [6].

The incorporated model was also utilized to represent the tableting process. First, the integrated network based on 10 ANN-RBF models in the first phase and a single ANN-RBF model in the second phase was established. The granules properties and the strength of the tablets produced were used as the inputs and the output of the integrated network, respectively. Such a model was followed by including the GMM framework, where the binder content and the porosity of the granules and the error residuals were utilized to develop such a model. The incorporated model performance for the strength is R^2 (training, testing) = [0.84, 0.82], as shown in Figure 3. It is worth noting that most of the predictions (85%) lie within a 90% confidence interval. The predictive performance for the strength was actually worse than the one for the granule size, but better than the ones for the binder content and porosity. Such a relatively low performance was actually expected because of the uncertainties in both the inputs and the output of the model that was developed for the tableting process. To elucidate further, the heterogeneous distributions of the binder content and air (i.e. porosity) of the granules seemed to affect only the outputs of the models that represent the granulation process, this being due to the fact that the inputs of the models developed for the granulation process are the process parameters (i.e. impeller shape, impeller speed, granulation time and L/S ratio) and the outputs are the granule size, the binder content and porosity. Accordingly, the uncertainties that stemmed from the heterogeneous distribution of the binder content and porosity affected the outputs of the models developed for the granulation process. However, the inputs of the model developed for the tableting process are the outputs of the granulation process (i.e. the granule size, the binder content and porosity) and the output of such a model is the tablet strength that is mathematically represented as a function of the inputs. Accordingly, the uncertainties that stemmed from the heterogeneous distribution of the binder content and porosity affected both the inputs and the output of the models developed for the tableting process. It is also worth mentioning at this stage that the performance of the incorporated model measured by the R^2

value was approximately four times and twice the performances of the single ANN-RBF model and the ensemble model, which was developed using 10 ANN-RBF models, respectively. Thus, these models were not utilized to model the tableting process.

Figure 3. The incorporated model for the tablet strength: (a) Training and (b) Testing (with 10% band).

4. The *Right-First-Time* Framework

4.1 Model Development

Recently, nature-inspired optimization algorithms have attracted a great deal of interest [11-12]. Several algorithms, such as Genetic and Evolutionary algorithms, have been proposed and applied to tackle nonlinear, multimodal and discontinuous real-world problems [12-21]. In particular, particle swarm optimization (PSO) algorithms, which mimic the competitive and cooperative behaviours of fish schooling and bird flocking, have been widely used, this being due to their computational efficiency (i.e. fast convergence speed) and, more often than not, their effectiveness in locating the global optima [11]. Various PSO algorithms have been presented in the related literature to solve single optimization problems (SOPs) [11]. The promising results that were obtained via these algorithms were the main motivation behind extending the PSO to deal with multi-objective optimization problems (MOPs) [20-21]. Unlike SOPs, solving MOPs leads to a set of alternative solutions. Such a set is optimal, in the wider sense, no other set of solutions is superior to it when all the objective functions are taken into consideration [10]. Such a set is usually known as *Pareto* optimal, non-inferior or non-dominated set [11].

Many algorithms, such as ranking objective functions and weighted sum methods, have been utilized to identify a single optimal solution for the MOPs [11, 21]. However, as already stated there is hitherto no universally accepted definition for such a single solution. In this research work, a new algorithm that integrates the PSO with a fuzzy logic system is proposed to solve MOPs.

Generally, a continuous and unconstrained MOP can simply be written as follows [11]:

$$\underset{x \in \Omega}{Min}\ F(x) = \left(f_1(x), f_2(x), \ldots f_N(x) \right) \tag{4}$$

where x is a l-dimensional vector of decision variables, which are usually bounded by the decision space (Ω). In many cases, the decision vector must satisfy a set of equality and inequality constraints [10, 21].

To identify the set of *Pareto* optimal solutions, a PSO algorithm is utilized. In the PSO, particles are randomly initialized within the feasible (i.e. search) space. Each particle represents a potential solution. A particle i is usually described by its position, $x_i = (x_{i1}, x_{i2}, \ldots, x_{il})$, and velocity, $v_i = (v_{i1}, v_{i2}, \ldots, v_{il})$. During the search, the position and the velocity of each particle are updated according to the particle's experience and the experience of the neighbour particles. For the i^{th} particle, these parameters can be updated using the following set of equations [20]:

$$\begin{aligned} v_{ij}^{t+1} &= v_{ij}^{t} + c_1 r_1 \left(p_{ij}^{t} - x_{ij}^{t} \right) + c_2 r_2 \left(p_{gj}^{t} - x_{ij}^{t} \right) \\ x_{ij}^{t+1} &= x_{ij}^{t} + v_{ij}^{t+1} \end{aligned} \tag{5}$$

where r_1 and r_2 are random numbers generated uniformly in the range of 0 to 1. The parameter t stands for the iteration number. The parameters c_1 and c_2, the so-called learning factors, represent the degree of influence of the local and global best solutions, respectively. The velocity value is usually bounded within a predefined range of $v_{\min}$ to $v_{\max}$ to prevent the particles from flying out of the feasible space [11].

The set of equations in (5) is only applicable when the decision variables are continuous. For discrete variables, the velocity vector is first transformed into a probability one through the sigmoid function as follows [20]:

$$\lambda_{ij}^{t} = \frac{1}{1+e^{-v_{ij}^{t+1}}} \tag{6}$$

where λ_{ij}^{t} represents the probability that the value of the j^{th} element of the position vector is 1, and v_{ij} represents the particle velocity as per (5). Hence, the position of the discrete particle can be updated as follows [20]:

$$x_{ij}^{t} = \begin{cases} 1 & \text{if } R < \lambda_{ij}^{t} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where R is a random number in the range of 0 to 1.

Once the *Pareto* optimal set is defined, the single optimal solution can be found using the FLS. First of all, and based on predefined criteria (e.g. the predictive performance of a model), a number of clusters can be defined in the feasible space. The membership function value for each point in the set can then be estimated as follows [28]:

$$\mu_i = \exp\left(\frac{-(P_{x_i} - M)^2}{2 \times \sigma^2}\right) \tag{8}$$

where μ_i represents the membership function value for the i^{th} solution. The parameters M and σ stand for the mean and the standard deviation of a cluster, respectively. Because of the continuity and smoothness of the Gaussian function, it is utilized, in this research work, to estimate the membership function. The parameter P_{x_i} represents the performance of the i^{th} solution estimated using the predefined criteria. It is worth noting at this stage that the number of criteria may not necessarily be equal to the number of the objective functions considered.

Once the membership function values are calculated, they can be combined for each solution using the set of fuzzy logic operations. Various operations, such as minimum and product, can be used for this purpose. Selecting the appropriate set of operations depends on the problem under investigation.

4.2 Model Implementation and Results

The models, which were introduced previously [3, 6] and briefly described above, were exploited within a *right-first-time* framework that was used to facilitate the implementation of the proposed optimization algorithm. Figure 4 illustrates the development of this framework for the granulation and tableting processes. As shown in Figure 4, such a framework was implemented in two stages. In the first stage, the target properties of the tablets were used to define the best granule properties, which were utilized to identify the optimal operating conditions of the HSG process in the second stage of the framework.

Figure 4. Modelling and optimization frameworks for the granulation and tableting processes.

Since the main property of the tablet that was considered in this research work is the strength, the input parameters for the tableting process (i.e. the properties of the granules produced) were identified using a single objective PSO optimization method. The optimization problem for the tableting process can be written as follows:

$$\begin{aligned} &\mathit{Minimize} \quad J_{Tablet} = \left| \frac{S_P}{S_T} - 1 \right| \\ &\mathit{Subject\,to}: \\ &\qquad\qquad x_{ij}^{\min} \le x_{ij} \le x_{ij}^{\max} \qquad \forall j \end{aligned} \tag{9}$$

where S represents the strength of the tablet, and x_{ij} represents the j^{th} element of the position vector of the i^{th} particle. The subscripts are used to distinguish the predicted strength value (P) from the target (i.e. desired) one (T). The constraints ensure that all of the granule properties, namely, size (500-1000μm), binder content (8-17%) and porosity (4-20%), are located between their minimum and maximum values. To elucidate, suppose that the target strength value of the tablet is 0.75 MPa. To illustrate, the optimization model in (9) can now be rewritten as follows:

$$\begin{aligned} &\mathit{Minimize} \quad J_{Tablet} = \left| \frac{S_P}{0.75} - 1 \right| \\ &\mathit{Subject\,to}: \\ &\qquad 500 \le x_{i1} \le 1000 \\ &\qquad 8 \le x_{i2} \le 17 \\ &\qquad 4 \le x_{i3} \le 20 \end{aligned} \tag{10}$$

The optimization problem presented in (10) was solved using an SOPSO algorithm, where 300 particles were randomly initialized. It is worth mentioning that the maximum number of iterations was 300. The optimal granule properties that were obtained by the SOPSO algorithm are: the granule size is in the range of 710 to 1000μm, the binder content is 11.75% and the porosity is 10.53%. Such properties were used as final target properties that need to be obtained from the granulation process.

For the granulation process, two objective functions should be considered; minimizing the difference between the target and the predicted values for both the binder content and the porosity, and maximizing the amount of the granules in the desired size range, in other words, minimizing the waste and the recycling ratio. The optimization problem for such a process can be mathematically described as follows:

$$\text{Minimize} \quad J_{properties} = \sum_{i=1}^{I} \left| \frac{\gamma_P^i}{\gamma_T^i} - 1 \right| \tag{11}$$

$$J_{waste} = 1 - \upsilon_T$$

$$\text{Subject to:}$$

$$x_{ij}^{\min} \le x_{ij} \le x_{ij}^{\max} \qquad \forall j, j \ne 1$$

$$x_{i1} = \begin{cases} 0 & \text{if impeller type I is used} \\ 1 & \text{if impeller type II is used} \end{cases}$$

where γ_i represents the i^{th} property of the granules, and x_{ij} represents the j^{th} element of the position vector of the i^{th} particle. Similarly to above, such subscripts should distinguish the predicted (P) from the target (T). The parameter υ_T represents the volume fraction of the granules that are in the required size range. The first set of constraints ensures that all the input variables of the granulation process are within the investigated ranges, as listed in Section 2, whereas the second constraint ensures that the impeller type is considered as a discrete or binary variable. To illustrate, the optimization model in (11) can now be rewritten as follows:

$$\text{Minimize} \quad J_{properties} = \left| \frac{\gamma_P^{BC}}{11.75} - 1 \right| + \left| \frac{\gamma_P^{Porosity}}{10.53} - 1 \right|$$

$$J_{waste} = 1 - \upsilon_{(710-1000)} \tag{12}$$

$$\text{Subject to:}$$

$$x_{i1} = \begin{cases} 0 & \text{if impeller type I is used} \\ 1 & \text{if impeller type II is used} \end{cases}$$

$$1000 \le x_{i2} \le 6000$$

$$6 \le x_{i3} \le 15$$

$$13 \le x_{i4} \le 15$$

where BC stands for Binder Content. Figure 5 shows the behaviour of the two objective functions in the search area of the three continuous parameters, namely, impeller speed, granulation time and L/S ratio. It is noticeable that the two objectives are in conflict, in other words, any improvement in one of the objectives leads to a deterioration in the other one. For instance, when both the impeller speed and the granulation time are medium, the first objective function is increasing while the second one is decreasing.

Figure 5. The 3D surfaces relating to the two objective functions: (a) Minimizing the waste and recycling ratio, and (b) Obtaining the pre-defined properties (the arrows show the location of the optimal solution, as selected by the fuzzy system, for a tablet with a 0.75 strength value).

The optimization approach presented in this research paper was implemented to find the single optimal solution. First, the MOPSO algorithm was applied to find the *Pareto* optimal set for the model presented in (12). The MOPSO process was run using 300 particles, which were initialized randomly, for 300 iterations. It is worth mentioning at this stage that the incorporated model was used to estimate the objective function values for the 300 iterations, whereas the hybrid and the fusion models were used on the last iteration only, this being due to the computational effort that is required by the hybrid model [6]. Since the size of the *Pareto* optimal set can be potentially large, the results are rather displayed on a Cartesian plot, such a set being shown in Figure 6. This figure proves that the two objectives are indeed conflicting.

Figure 6. The normalized best performance obtained by MOPSO algorithm (Tablet strength is 0.75 MPa, the final single optimal solution, as selected by the fuzzy system, is highlighted (o)).

Defining the *Pareto* set was followed by selecting the single optimal solution. Various criteria can be selected based on the designer's priorities and preferences. However, in the *right-first-time* framework, one may consider the performance of the predictive models and their behaviours at different areas in the feasible space. In this research paper, the predictive performances of the fusion model for the three granule properties were chosen as criteria. Therefore, the granulation input variables and the error residuals that resulted from the fusion model were utilized to determine the predictive performance of such a model in the optimization feasible space. It is worth emphasizing, at this stage, that the predictive performance of the fusion model depends on the predictive performance values of the incorporated and the hybrid models. Based on the number of clusters that was defined by the fusion model as described in Section 3 and by using 5 clusters, Figure 7 illustrates how the fusion model performed in one of the space areas of the binder content. Such a figure shows that the performance of the fusion model is good when the impeller is of type I, the impeller speed is medium, the granulation time is small and the L/S ratio is medium.

Figure 7. An example of the fusion model performance in the space area of the binder content.

The membership function values for each point in the *Pareto* optimal set were estimated using Equation (8). For each cluster, each solution had membership function values for the size, binder content and porosity. These values were combined using the set of fuzzy logic operations. The minimum operation was used to combine the binder content and porosity. The product operation was then utilized to combine the minimum value with the membership function value for the size. The solution that corresponds to the maximum membership function value was finally selected. The best set of the operating conditions for the granulation process is: impeller type is of type I, impeller speed is 3879 rpm, granulation time is 6 minutes and L/S ratio is 13.74%. The single optimal point is highlighted in Figure 6. The location of this point is also highlighted by arrows in Figure 5, where it can be seen that any shift in this point may lead to a deterioration in one of the objective functions. An actual granulation experiment was conducted, where the input variables were assigned to be as close as possible to the optimal set of the operating conditions. The granules produced were characterized in terms of size, binder content and porosity. This was followed by producing tablets from the granules in the size range of 710 to 1000μm, and measuring the strength of these tablets. The target and the experimental properties of the granules and the tablets are summarized in Table 1.

Table 1. The target and the predicted values of the granule and tablet properties.

To prove the efficiency and the effectiveness of the proposed algorithm at different areas of the feasible space, two strength values, 0.45 MPa and 0.6 MPa, were also used to test the *right-first-time* framework. It is worth mentioning that these values, in addition to the 0.75 strength value, are examples that were selected from different areas of the feasible space. The experimental and the target values of the properties for these values are listed in Table 1.

It is noticeable that there is a difference between the target and the predicted values for all the investigated properties. Such a difference is relatively high for the strength of the tablets (approximately 10%) when compared to the differences for the granule properties. The actual reasons behind this can be attributed to one or more of the following reasons:

- The heterogeneity of the granules: It has been demonstrated in the open literature that the granules from the same batch but different size classes are heterogeneous [29-30]. In addition, the granules from the same size class may not be homogeneous, particularly when the size class is wide.

- Measurement uncertainties: Laboratory apparatus and techniques, which were used to characterize the granules and the tablets, involved uncertainties. Although each measurement was repeated a number of times, such uncertainties could not be eliminated. Moreover, the propagation of uncertainties might affect the models, in particular, the model that was developed for the tableting process.

- Simplifying the optimization model: As mentioned previously, only the incorporated model was used to estimate the objective function values for the 300 iterations, whereas the hybrid and the fusion models were used on the last iteration only. Since the predictive performances of the incorporated model for the binder content and porosity were not as good as the ones for the hybrid and the fusion models, this assumption might negatively affect the performance of the *right-first-time* framework.

For comparison purposes and in order to prove the effectiveness and efficiency of the proposed algorithm, the well-known PSO algorithm that was incorporated with a fitness assignment approach was utilized to identify the optimal granules' properties and the optimal operating conditions for the granulation process to produce tablets with the same predefined strength values. Then, actual granulation and tableting experiments were carried-out, and the granules' properties and the strength of the tablets were measured. The target and the experimental results are summarized in Table 1, where it is obvious that the proposed algorithm outperforms the PSO one, this is being due to its ability to take account of the predictive performance of the modelling framework.

In summary, the MOPSO algorithm and the FLS play complementary roles in the optimization process, where the former was utilized to define the set of Pareto optimal solutions and the latter was used to identify the single optimal solution by considering the predictive modelling performances of the forward models. It is worth mentioning at this stage that the FLS alone cannot be used to identify the *Pareto* optimal solutions for an MOP. On the other hand, the PSO algorithm was not able to consider the predictive performances of the forward models. Therefore, integrating these two approaches has combined the strengths of these approaches in a way that the single optimal solution can be successfully identified. The proposed approach was successfully implemented to produce granules and tablets right from the first time. Such an approach can be further improved to be more robust to the uncertainties and the propagation of uncertainties. A type-2 fuzzy set can be integrated with the PSO algorithm. The choice of such a set is motivated by the fact that the type-2 fuzzy set can deal with uncertainties more efficiently compared to its counterpart type-1 fuzzy set. However, it is well known that a type-2 fuzzy system is computationally demanding when compared to type-1. Therefore, further investigation should be performed to explore the advantages and the potential limitations of integrating such a more complex topology.

5. Conclusions

In the pharmaceutical industry, producing granules and tablets right from the first time may be a 'tricky' exercise that researchers in both academia and industry strive to ascertain, this being due to the complex nature of the many processes involved, namely, granulation and tableting processes, which are key to the production line of these pharmaceutical tablets, and also due to the uncertainties and challenges that surround these processes, including access to measurement points. Therefore, in this research paper, a new approach that integrated a particle swarm optimization (PSO) algorithm with a fuzzy logic system (FLS) was proposed. The ultimate aim of this new approach was to facilitate the development of the *right-first-time* framework by which the concept of the optimal production of granules and tablets was achieved. This framework inverted the non-linear models that were previously elicited to represent the high shear granulation (HSG) and the tableting processes [3, 6]. By implementing such a framework, the optimal granule properties and the optimal operating conditions to produce a tablet with predefined properties were successfully achieved. Moreover, waste and recycling ratios were significantly minimized.

This *right-first-time* optimization framework is original in that it circumvented the direct mathematical inversion of the complex models via multi-objective optimization which, more often than not, is not possible due to the fact that such inverse does not exist or simply is not unique; furthermore, it considered more than one objective function which made the definition of optimality not so straightforward and instead the concept of *Pareto*-optimality was relied upon whereby a set of non-dominated solutions was obtained. In this case study, it was possible to only include two objectives without loss of interpretability or accuracy but even if more objectives need to be incorporated in the future, such as those associated with the environment, for instance, then these can be taken into account within a hierarchical framework without compromising the concept of *Pareto*-optimality.

In summary, the results achieved in this research were truly promising and showed that *right-first-time* production of the granules and tablets should no longer remain just a 'myth'. By considering some aspects (e.g. scaling-up methods), the modelling frameworks can be implemented correctly in the pharmaceutical industry (i.e. on a relatively large scale). The advantages of such an implantation can include, but not limited to, developing a cost-effective product, meeting the stringent regulations imposed on the pharmaceutical industry and minimizing new product development time.

Acknowledgement

The Authors wish to thank Maschinenfabrik Gustav Eirich GmbH & Co KG (Germany) for providing the Eirich mixer that has been used for all the experimental work.

References

[1] 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.

[2] P.R. Mort, S.W. Capeci, J.W. Holder, Control of agglomerate attributes in a continuous binder-agglomeration process, Powder Technology, 117 (2001) 173-176.

[3] 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.

[4] W.H. AlAlaween, B. Khorsheed, M. Mahfouf, G.K. Reynolds, A.D. Salman, An interpretable fuzzy logic based data-driven model for the twin screw granulation process, Powder Technology, 364 (2020) 135-144.

[5] D. Barrasso, A. El 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.

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

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

[8] C.F.W. Sanders, A.W. Willemse, A.D. Salman, M.J. Hounslow, Development of a predictive agglomeration model, Powder Technology, 138 (2003) 18-24.

[9] X. Yu, M.J. Hounslow, G.K. Reynolds, A. Rasmuson, I.N. Bjorn, P.J. Abrahamsson, A compartmental CFD-PBM model of high shear wet granulation, AIChE. 63 (2017) 438-458.

[10] M. Mahfouf, M. Jamei, D.A. Linkens, J. Tenner, Inverse modelling for optimal metal design using fuzzy specified multi-objective fitness functions, Control Engineering Procatice, 16 (2008) 179-191.

[11] J.F. Kennedy, R.C. Eberhart, Y. Shi, Swarm Intelligence, San Francisco, London, 2001.

[12] P. Shanmugavadivu, K. Balasubramanian, K., Particle swarm optimized multi-objective histogram equalization for image enhancement, Optics and Laser Technology, 57 (2014) 243-251.

[13] Y.M. da Silva Veloso, M.M. de Almeida, O.L.S. de Alsina, M.L. Passos, A.S. Mujumdar, M.S. Leite, Hybrid phenomenological/ANN-PSO modelling of a deformable material in spouted bed drying process, Powder Technology, 366 (2020) 185-196.

[14] W.H. AlAlaween, B. Khorsheed, M. Mahfouf, I. Gabbott, G.K. Reynolds and 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.

[15] Y.M. da Silva Veloso, M.M. de Almeida, O.L.S. de Alsina, M.S. Leite, Artificial neural network model for the flow regime recognition in the drying of guava pieces in the spouted bed, Chemical Engineering Communications, 207 2020 549-558.

[16] H.M. Zawbaa, S. Schiano, L. Perez-Gandarillas, C. Grosan, A. Michrafy, C.Y. Wu, Computational intelligence modelling of pharmaceutical tabletting processes using bio-inspired optimization algorithms, Advanced Powder Technology, 29 (2018) 2966-2977.

[17] E.H. Houssein, M.E. Hosney, D. Oliva, W.M. Mohamed, M. Hassaballah, A novel

hybrid Harris hawks optimization and support vector machines for drug design and discovery, Computers & Chemical Engineering, 133 (2020) 106656.

[18] D.F. Viana, G.R. Salazar-Banda, M.S. Leite, Electrochemical degradation of Reactive Black 5 with surface response and artificial neural networks optimization models, Separation Science and Technology, 53 (2018) 2647-2661.

[19] C. Qi, Q. Chen, X. Dong, Q. Zhang, Z.M. Yaseen, Pressure drops of fresh cemented paste backfills through coupled test loop experiments and machine learning techniques, Powder Technology, 361 (2020) 748-758.

[20] A. Unler, A. Murat, A discrete particle swarm optimization method for feature selection in binary classification problems, European Journal of Operational Research, 206 (2010) 528-539.

[21] T.T. Vu, H.H. Kha, T.Q. Duong, N.S. Vo, Particle Swarm Optimization for Weighted Sum Rate Maximization in MIMO Broadcast Channels, Wireless Personal Communications, 96 (2017) 1-15.

[22] J. Fu, Y. Cheong, G. Reynolds, M. Adams, A. Salman, M. Hounslow, An experimental study of the variability in the properties and quality of wet granules, Powder Technology, 140 (2004) 209–216.

[23] P.C. Knight, T. Instone, J.M.K. Pearson, M.J. Hounslow, An investigation into the kinetics of liquid distribution and growth in high shear mixer agglomeration, Powder Technology, 97 (1998) 246–257.

[24] C. Bishop, Pattern recognition and machine learning, Springer, New York, 2006.

[25] Y.Y. Yang, M. Mahfouf, G. Panoutsos, Probabilistic Characterisation of Model Error Using Gaussian Mixture Model– with Application to Charpy Impact Energy Prediction for Alloy Steel, Control Engineering Practice, 20 (2012) 82-92.

[26] A. Frikha, H. Moalla, Analytic hierarchy process for multi-sensor data fusion based on belief function theory, European Journal of Operational Research, 241 (2015) 133-147.

[27] A. Boudraa, A. Bentabet, F. Salzenstein, L. Guillon, Dempster-Shafer's basic probability assignment based on fuzzy membership functions, Electronic Letters on Computer Vision and Image Analysis 4 (2004) 1-9.

[28] J.M. Mendel, Uncertain rule-based fuzzy logic systems: introduction and new directions, Upper Saddle River, Prentice Hall, 2001.

[29] G.K. Reynolds, C.A. Biggs, A.D. Salman, M.J. Hounslow, Non-uniformity of binder distribution in high-shear granulation, Powder Technology, 140 (2004) 203-208.

[30] A.C. Scott, M.J. Hounslow, T. Instone, Direct evidence of heterogeneity during high-

shear granulation, Powder Technology, 113 (2000) 205-213.

Figures

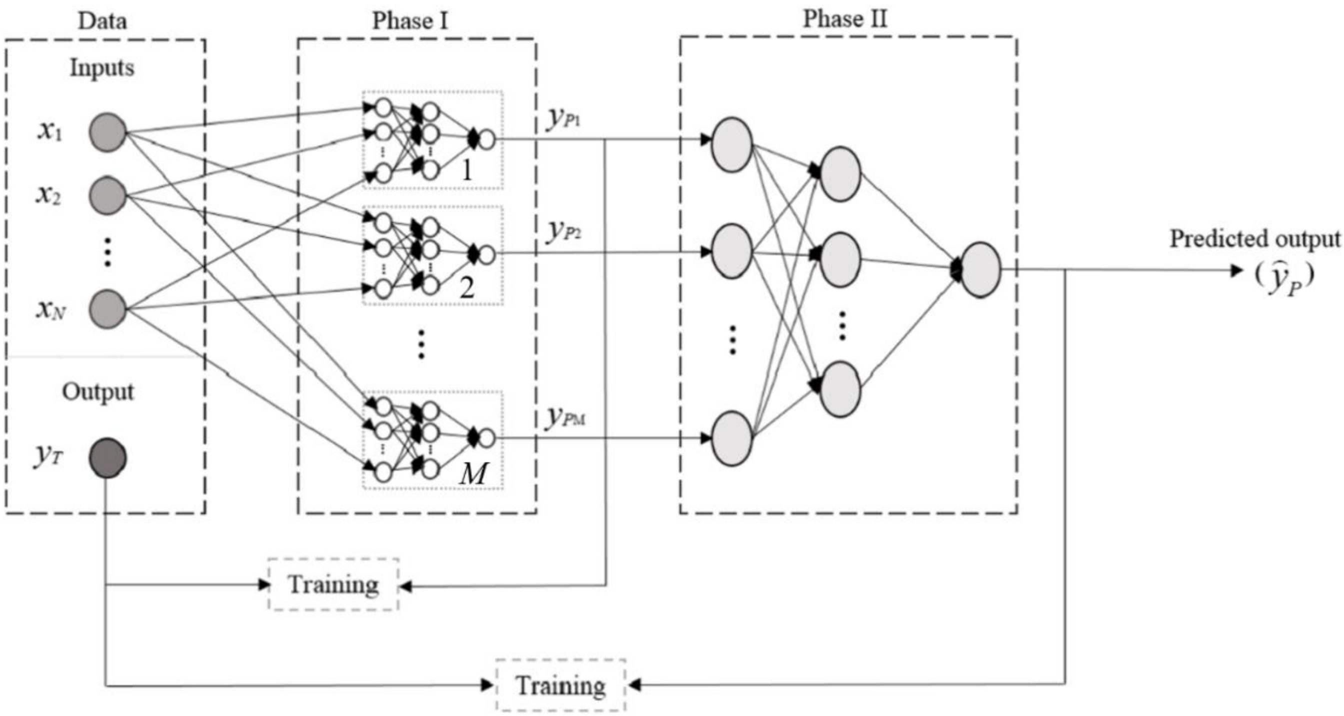

Figure 1. The architecture of the integrated network [3].

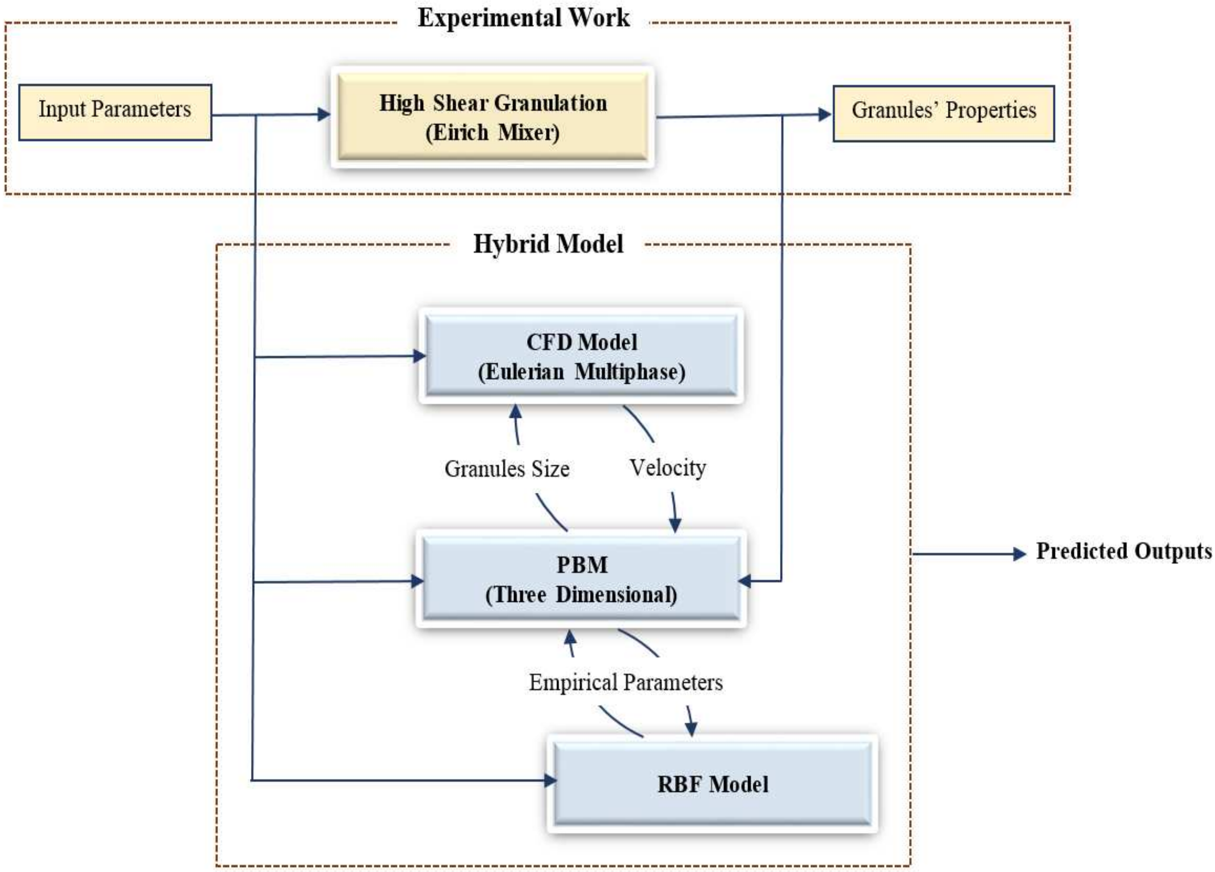

Figure 2. The hybrid model for the HSG process [6].

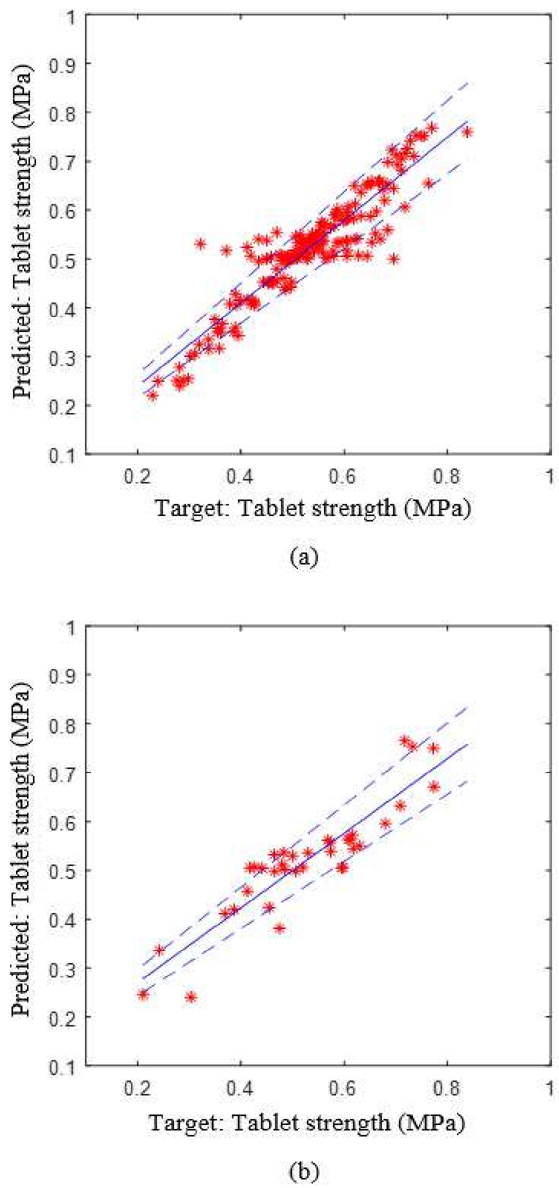

Figure 3. The incorporated model for the tablet strength: (a) Training and (b) Testing (with 10% band).

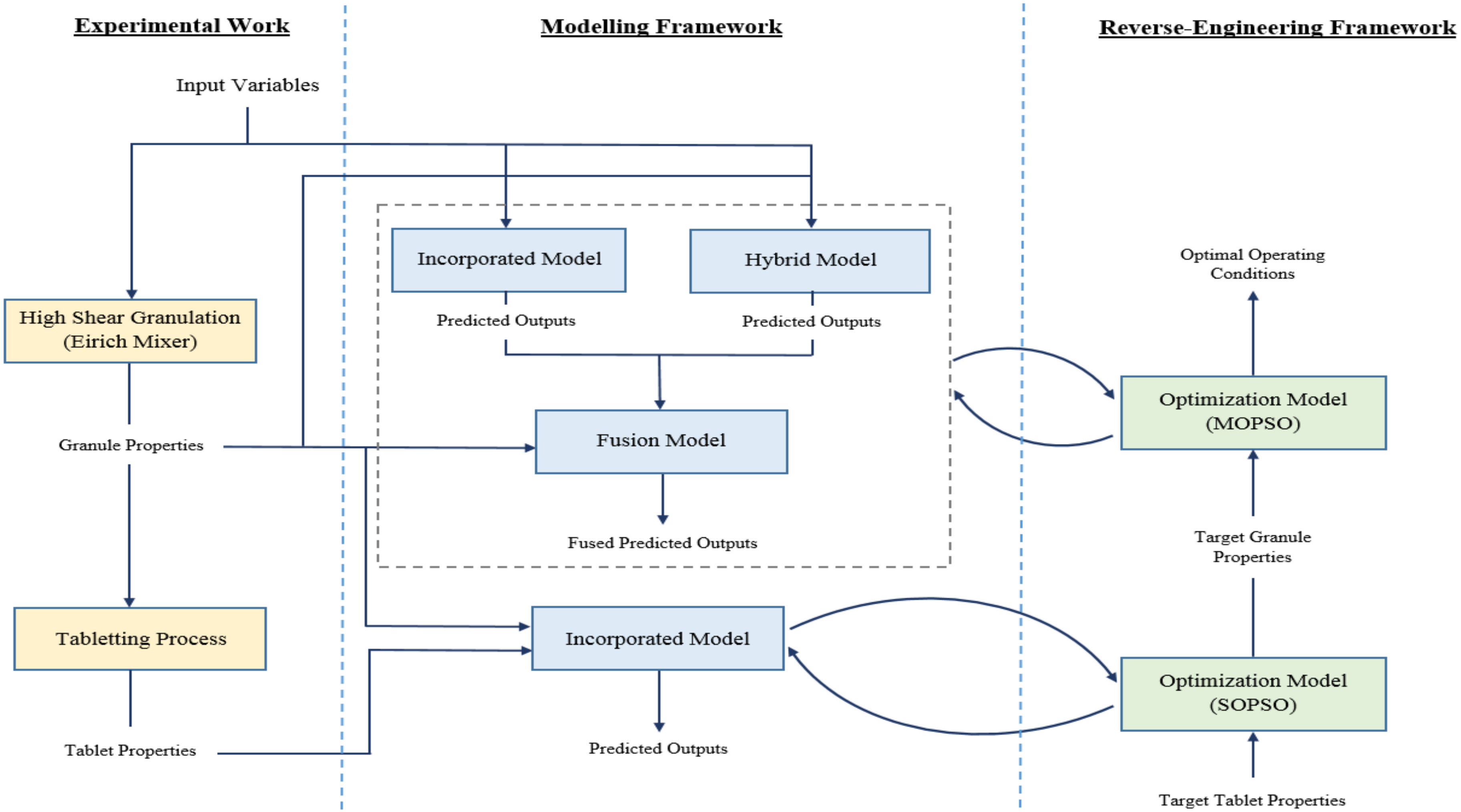

Figure 4. Modelling and optimization frameworks for the granulation and tableting processes.

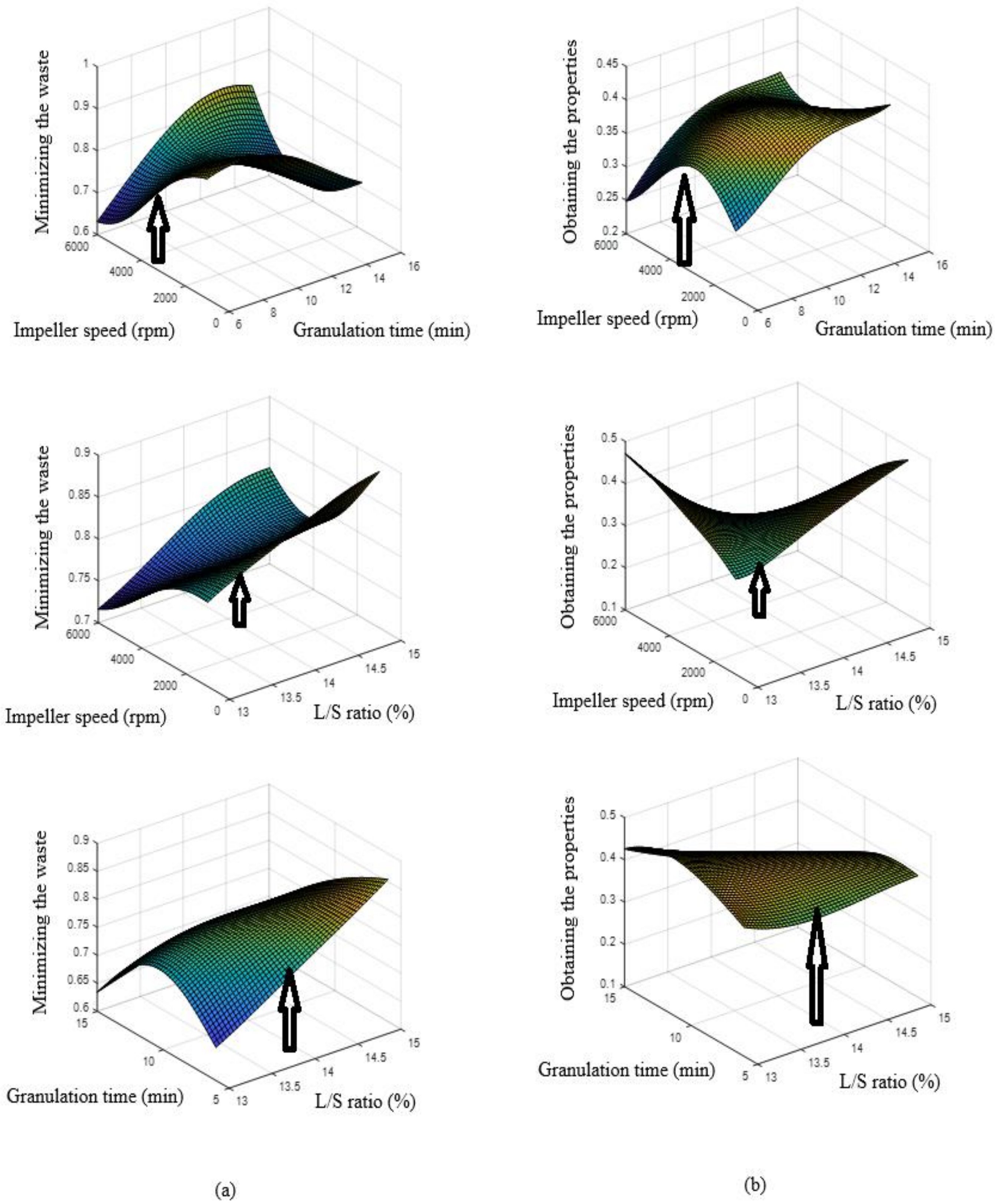

Figure 5. The 3D surfaces relating to the two objective functions: (a) Minimizing the waste and recycling ratio, and (b) Obtaining the pre-defined properties (the arrows show the location of the optimal solution, as selected by the fuzzy system, for a tablet with a 0.75 strength value).

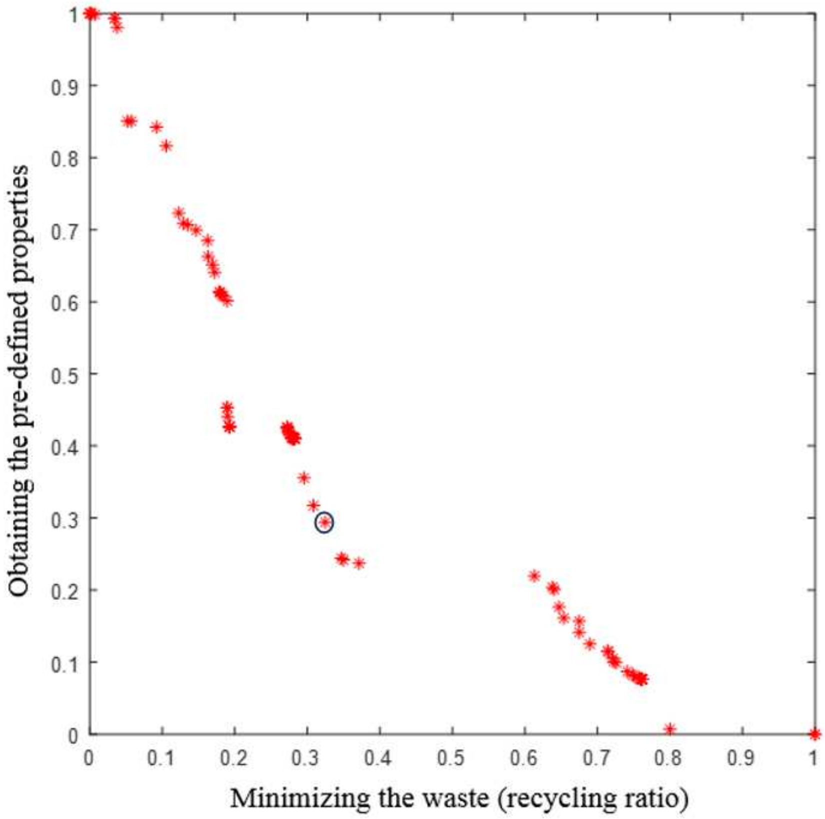

Figure 6. The normalized best performance obtained by MOPSO algorithm (Tablet strength is 0.75 MPa, the final single optimal solution, as selected by the fuzzy system, is highlighted (o)).

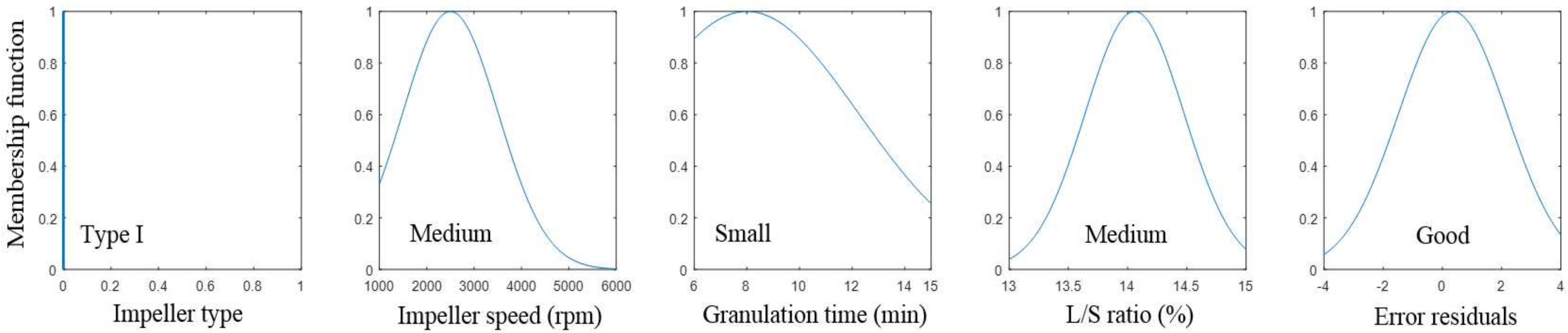

Figure 7. An example of the fusion model performance in the space area of the binder content.

Table 1. The target and the predicted values of the granule and tablet properties*.

	Target tablet strength (MPa)		0.45		0.6		0.75	
			PSO and FLS	PSO	PSO and FLS	PSO	PSO and FLS	PSO
Operating conditions	Impeller shape		Type I	Type I	Type II	Type I	Type I	Type II
	Impeller speed (rpm)		5978.88	3363.57	1878.15	5237.32	3879.45	5832.24
	Granulation time (min)		6.06	15	12.69	6.17	6	6.02
	L/S ratio (%)		14.03	13.84	14.8	14.65	13.74	13.27
Granule properties	Binder content (%)	**Target**	13.71		8.75		11.75	
		Experimental	12.86±0.53	11.81±0.42	9.72±0.95	10.44±0.48	11.63±0.37	12.88±0.34
	Porosity (%)	**Target**	8.58		10.94		10.53	
		Experimental	9.24±0.91	9.44±0.95	10.04±0.63	13.02±0.43	9.72±0.66	11.80±0.71
Experimental Tablet strength (MPa)		**Experimental**	0.412 ±0.03	0.407±0.03	0.533±0.07	0.68±0.04	0.814±0.05	0.858±0.03

*The intervals were calculated using a 95% confidence interval.