J Pharm Innov (2010) 5:119–137
DOI 10.1007/s12247-010-9086-y

RESEARCH ARTICLE

Design Space of Pharmaceutical Processes Using Data-Driven-Based Methods

Fani Boukouvala · Fernando J. Muzzio · Marianthi G. Ierapetritou

Published online: 14 September 2010

Abstract
Introduction The identification and graphical representation of process design space are critical in locating not only feasible but also optimum operating variable ranges and design configurations. In this work, the mapping of the design space of pharmaceutical processes is achieved using the ideas of process operability and flexibility under uncertainty.
Methods For this purpose, three approaches are proposed which are based on different data-driven modeling techniques: response surface methodology, high-dimensional model representation, and kriging methodology. Using these approaches, models that describe the behavior of the process at different design configurations are generated using solely experimental data. The models are utilized in mixed integer non-linear programming formulations, where the optimum designs are identified for different combinations of input parameters within the operating parameter and material property ranges.
Results Based on this idea, by defining a desirable output range, the corresponding range of input variables that result to acceptable performance can be accurately calculated and graphically represented.
Conclusions The main advantages of the methodologies used in this work are, firstly, that there is no restriction by the lack of first-principle models that describe the investigated process and, secondly, that the models developed are computationally inexpensive. This work can also be used for the comparative analysis of the use of different surrogate-based methodologies for the identification of pharmaceutical process Design Space.

Keywords Design space mapping · Data-driven models · Kriging · High-dimensional model representations · Response surface · Pharmaceutical processes

F. Boukouvala · F. J. Muzzio · M. G. Ierapetritou (✉)
Department of Chemical and Biochemical Engineering,
Rutgers University,
Piscataway, NJ 08854, USA
e-mail: marianth@soemail.rutgers.edu

Introduction

Definition of Design Space in Pharmaceutical Engineering

The concept of design space was formally introduced in 2005 [1, 2] by the International Society of Pharmaceutical Engineering as one of the fundamental building blocks of the Product Quality Lifecycle Implementation (PQLI) initiative. In simple terms, it is the area of the parametric space within which acceptable product can be produced. The use of risk-based analyses to determine design constraints and then determine appropriate controls has been a popular research area in many other industries. Lately, it has been pointed out that the use of such practices has not been formally applied in the pharmaceutical industry. Working towards an ideal maximum level of quality assurance and performance in the pharmaceutical industry, the design space is directly linked to a well-defined control strategy [3] and criticality assessment [4]. A process variable, material property, process, or step is defined as critical if it causes significant variability in the product quality or process performance within a range of values of the variable that is likely to occur if the variable is not properly controlled. Criticality is highly connected to the results of risk assessment, from which the critical process parameters and quality attributes are identified, and these are the only ones included in the design space.

Finally, a successful control strategy uses the knowledge derived from the criticality assessment and the design space to assure that the process operates within the design space.

Sources for accurately defining the design space of a process include the available literature, experience, and knowledge of the process, first principles, experimental data, empirical models or—most often—some combination of all these methods. However, the choice of the tools used to characterize the design space depends on the availability of these resources. For example, when the science underlying a process is known and well developed, the design space can be obtained by predictive first-principle models. On the other hand, if a process is new and its scientific principles are not well understood, it can be treated as a black-box process. Consequently, the design space is dynamic, since, as additional knowledge and information about a process is obtained (or as new raw materials, evolving specifications, and new technology become available), one can be more certain of its structure and form. Clearly, there is no unique approach for defining a design space of a process, but a general guidance and a series of steps are outlined in [1]. Once the methodology and tools that define the design space are chosen, the design space has to be presented efficiently, in a way that can be easily interpreted.

Figure 1 is a very common representation of the design space, used to explain graphically its relationship to the knowledge space and the normal operating ranges. The larger knowledge space contains all the information about all regions of the process that have been investigated. Subsequently, the difference between the knowledge and design space can be defined as the region of the knowledge space that generates unacceptable product.

Several examples of specific experimental procedures to define the design space of pharmaceutical processes are found in the literature. For example, in [5], the operating window for the process of a fluidized bed granulation has been identified by assessing the impact of critical parameters, such as inlet air humidity, on fluidization behavior and granule size. A similar procedure was followed in [6] in order to define the confidence zone of chromatographic analytical methods. In the present work, three different data-based methodologies are used to generate the graphical representation of the region in which acceptable performance of unit black-box processes can be ensured. The efficiency of these approaches is compared through two pharmaceutical case studies: predicting the design space of a continuous powder mixer and a loss-in-weight feeder of a continuous tablet manufacturing process. The proposed approaches can be considered as building blocks of a general methodology for defining the design space of black-box processes, which operate under uncertainty and for which the only available information is a design of experiment (DOE)-derived experimental data set. The next section of the introduction contains a brief literature review of previous work done in areas of design, optimization, and control of processes under uncertainty, which are used in the present work for defining the design space.

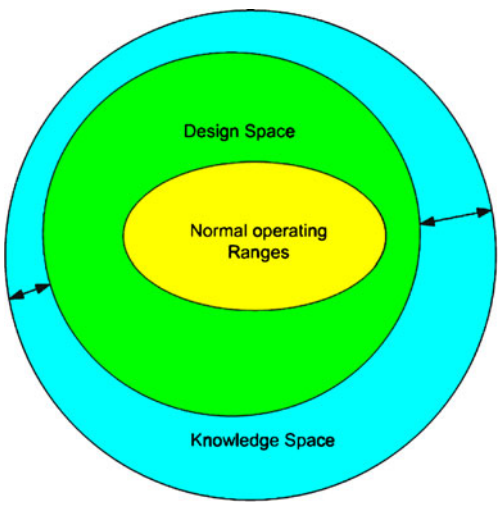


Fig. 1 Link between knowledge space, design space and normal operating ranges

Optimization Approaches to Defining the Design Space

In this work, the construction of the Design Space is considered as the problem of determining the boundaries of a process where feasible, profitable, and acceptable performance of the process is guaranteed. For many years, this has been considered as a major concern in many process industries and a substantial amount of work has been done in order to define concepts such as "operability," "feasibility," and "flexibility" of processes that contain uncertain parameters. Uncertainty can occur for a variety of reasons, most commonly among them is the variability of certain process parameters during plant operation.

Optimal process design under uncertainty was defined as a rigorous formulation in the 1980s [7], where the effects of parameters that contain considerable uncertainty on the optimality and feasibility of a chemical plant were studied. The objective of solving such problems was to ensure optimality and feasibility of operation for a given range of uncertain parameter values, by identifying a measure of the size of the feasible region of operation. According to the methodology introduced in this work, the problem was represented as a max–min–max formulation, where, for a given design and fixed values of the uncertain parameters, the feasible region was calculated. In [8], the flexibility of such processes was defined and quantified by the flexibility index, which represented the maximum allowed deviation of uncertain parameters from their nominal values, such that feasible operation could be guaranteed by changing the

control variables. A series of papers dealing with flexibility analysis and the formulation and optimization of processes under uncertainty were published in the following years; most of them, however, required known process models and rely on particular convexity assumptions [9–12].

More recently, the advantages and the need of integrating process design and control have been recognized for ensuring both optimality and controllability of a process [13, 14]. Specifically, the ability of a process to move from one steady state to another is quantified in order to avoid unacceptable disturbances and minimize unnecessary variability in the final product. Firstly, the desired outputs and the possible expected disturbances are taken into account and then the operability of a process is defined. Operability can be assured by controlling the set of input parameters, while identifying the optimum out of possibly multiple competing designs. The applications of operability analysis can be classified into linear or non-linear and also steady state or dynamic models [15–17]. In [18], an interested reader can find a systematic comparison of the terms of operability and flexibility and their applications in real-life steady-state case problems.

Working in the context of modeling and designing under uncertainty, the recently proposed approach by Banerjee et al. [19] is independent of the mathematical complexity of a process model or any type of convexity assumptions. High-dimensional model representation (HDMR) is used as a method of input–output mapping of black-box processes, and this method is used to model the effect of parameter uncertainty on process design and consequently identify the feasible operation of the process. In [20], the authors extend their work to handle more complex cases where the variability of the optimal solution of mixed integer non-linear programming (MINLP) models with parameter uncertainty is captured through the use of HDMR.

It is by now very common in all areas of science and engineering to utilize mathematical programming techniques. MINLP refers to optimization models that contain both discrete and continuous variables as well as nonlinearities in the problem constraints and objective function. In general, the objective function to be optimized is representative of the performance criteria of the process that is being optimized, for example minimization of cost and output variability or maximization of profit [21]. The solution of an MINLP problem is a set of optimal values of the discrete and continuous variables that simultaneously minimize/maximize the objective(s) and satisfy the constraints.

The following sections are organized as follows. "Data-Driven Modeling Methodologies" consists of a concise description of the surrogate modeling techniques used to model the behavior of processes, for which first principles are not yet available. "Modeling with Discrete Design Variables" clarifies the difficulties of modeling with discrete design variables. The detailed description of the proposed methodologies for constructing the graphical representation of the design space is discussed in "Proposed Approaches," and their applications and results are exemplified with the use of three pharmaceutical case studies in "Case Studies." Small differences in the calculated design spaces are further verified by experimental validation described in "Experimental Validation."

Data-Driven Modeling Methodologies

Three different techniques for mapping the behavior of systems for which first-principle models are not available are used in this work. In a previous work, it has been shown that kriging and response surface methodology (RSM) can be used to produce predictive steady-state models for pharmaceutical unit processes even if some of the experimental data are noisy or missing [22, 23]. The theoretical background of the surrogate-based methodologies is briefly discussed in this section—while available literature is cited.

High-Dimensional Model Representations

High-dimensional model representation methodology was introduced to deal with the exploration of the effect of a high-dimensional input variable space to the output behavior of a system [24–29]. Li et al. [30, 31] have used this method to model chemical laboratory experimental processes. The output $f(x)$ of a process is considered as a finite hierarchical correlated function expansion, in terms of the independent and cooperative forms of input variables $x = \{x_1, x_2, \ldots, x_n\}$. Specifically:

$$f(x) = f_0 + \sum f_i(x_i) + \sum f_{ij}(x_i, x_j) + \cdots + \sum f_{12\ldots n}(x_1, x_2, \ldots, x_n) \quad (1)$$

where f_0 is a constant; $\sum f_i(x_i)$represents the independent contribution of the different inputs to the value of the output; $\sum f_{ij}(x_{i,}x_j)$ represents the sum of the pair-correlated contributions and $\sum f_{12..n}(x_1, x_2, \ldots, x_n)$ represents the sum of nth correlated contributions; n is equal to the number of variables. This expansion is always finite and exact, and most importantly the output is assumed to be well-defined in terms of the input variables. However, the number of terms can be prohibitively large, and it increases rapidly as the input variables increase. This limitation can be overcome since in real typical problems it is observed that interactions of order higher than two are either very small or not statistically significant. Thus, Eq. 1 can be approximated by:

$$f(x) \approx f_0 + \sum f_i(x_i) + \sum f_{ij}(x_i, x_j) \quad (2)$$

Thus, HDMR can be used as a fast algorithm that would circumvent the exponential difficulty of the high-dimensional mapping problem [27]. The HDMR component functions are optimally tailored to each particular $f(x)$ over the entire domain of x. In other words, each component function is obtained by solving an optimization problem—in which the minimum functional under specific orthogonality conditions is calculated.

In the literature, two different types of HDMR expansions have been introduced: cut and random sampling (RS) HDMR [24–29]. In the cut-HDMR, the values of the component functions depend on a chosen reference point in terms of which the expansion is written, while RS-HDMR functions depend on the average value of $f(x)$ over the entire domain. In the present study, cut-HDMR is employed. RS-HDMR is more computationally expensive, since it requires the calculation of N-dimensional integrals for the constant terms. Cut-HDMR, however, minimizes the computational cost since it requires evaluations of functions at sample points, which we have readily available from the experimental data. Initially, a combination of set-point input variables is chosen as a nominal point, and the output of the process at these operating conditions is the constant term f_0. Higher-order component functions are calculated as cuts (axes, planes, etc.) through the nominal point. For example, the first-order term $f_i(x_i)$ is evaluated for x_i while all other input variables are kept at their nominal values. Next, look-up tables are constructed for the component functions so that component functions at any given arbitrary x_i can be interpolated and added up to obtain the final output. The main advantage of this methodology is that the number of required model runs in order to construct a look-up table grows in a polynomial rate as the number of inputs increases, instead of exponentially (traditional sampling). A more detailed description of the steps of the algorithm of the interpolation technique is included in Appendix 3.

Response Surface Methodology

RSM was first introduced by Box and Wilson in 1951 [32] and is a tool that has been widely employed for the optimization of noisy processes. RSM is a local optimization technique whereby an optimum is found after sequential optimization of localized sampling-based models [33]. There are three basic steps to the algorithm: (1) specification of a sampling set within the local region, usually accomplished with DOE tools, (2) construction of a local model centered at a nominal sampling point, and (3) model optimization with respect to the local region in order to determine the location at which process improvement is maximized. There are two major questions associated with model construction, namely (1) spatial location of sampling points and (2) quantification of model uncertainty. In order to address the first issue, DOE tools are applied. The experimental design is defined as the specification of a number of treatment levels for each input variable, the experimental units by which responses are measured, and the mechanism by which treatments are assigned to units.

For a problem containing n continuous input variables, an n-dimensional quadratic polynomial is used as the local model since quadratic behavior describes the mathematical geometry in the neighborhood of an optimum. Model accuracy can be improved if bilinear terms capturing the interaction effects between two inputs are also incorporated into the local model. A general second-order response surface model has the following form:

$$z = \beta_0 + \sum_j \beta_j x_j + \sum_{i<j} \beta_{ij} x_i x_j + \sum_j \beta_{jj} x_j^2 \tag{3}$$

where x_j are input variables, β_0 β_j, β_{ij}, and β_{jj} are model coefficients, and z represents the response that describes the predicted output behavior.

Kriging Methodology

Kriging was first developed as an inverse distance weighting method to describe the spatial distribution of mineral deposits [34, 35]. This method has attracted a lot of attention recently due to its capability of modeling complex functions with also providing error estimates [22]. The prediction f_{pred} is expressed as a weighted sum of the observed function values at sampling points that fall within a set interval around the point that is predicted. The basic idea of kriging methodology is that a function value for a sampling point located close to the test point is generally weighted more heavily in contrast to the function value corresponding to a sampling point located farther away. Lower weight is placed on function values whose sampling points are clustered together in order to minimize the possibility of generating biased estimates. Since a variance for each test point is also calculated, regions where subsequent sampling is required can be linked to a high variance at the regional points.

The first step of the kriging methodology is the determination of variogram coefficients from an experimental sampling set consisting of N sampling points. The variogram is a quantitative descriptive statistic that graphically characterizes data set roughness (and the information obtained complements that which is obtained using histograms and other common descriptive statistics). The variogram coefficients are determined by fitting an optimum variogram model to the data. Data smoothing is often used to improve the fit by replacing clustered scatter points falling within an interval with average values defined as semivariances. Variogram model coefficients are then obtained from

regression of the semivariance scatter points to one of the five elementary types: spherical, Gaussian, exponential, power, or linear. It is considered that the one that captures the trend of the semivariance best is the one whose least square error is the lowest. The covariance function which is a complementary function of the semivariance is used to calculate the kriging weights (w_i) by solving the following system:

$$\begin{bmatrix} \mathrm{Cov}(d_{1,1}) & \cdots & \mathrm{Cov}(d_{1,M}) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \mathrm{Cov}(d_{M,1}) & \cdots & \mathrm{Cov}(d_{M,M}) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \times \begin{bmatrix} w_1 \\ \vdots \\ w_M \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathrm{Cov}(d_{1,k}) \\ \vdots \\ \mathrm{Cov}(d_{M,k}) \\ 1 \end{bmatrix} \tag{4}$$

where d_{ij} is the distance between sampling point x_i and sampling point x_j and d_{ik} is the distance between sampling point x_i and test point x_k. Similarly, Cov(d_{ij}) and Cov(d_{ik}) represent the modeled covariances between sampled function data whose corresponding input vectors are a distance x_i–x_j or x_i–x_k apart, respectively. The kriging prediction f_k is then evaluated by the following form:

$$f_k = \sum_{i=1}^{M} w_i f_i \tag{5}$$

where w_i and f_i represent the weight and observed value at sampling point i, respectively. For each test point x_k, a variance σ_k is also obtained as follows:

$$\sigma_k^2 = \sigma_{\max}^2 - \sum\nolimits_{i=1}^{M} w_i \mathrm{Cov}(d_{ik}) - \lambda \tag{6}$$

where Cov (d_{ik}) corresponds to the right-hand side of Eq. 4.

Through this procedure, it is suggested that additional experiments are conducted at test points which have any of the following characteristics: (1) high variance, (2) high difference between kriging predictions for consecutive iterations, and (3) minimum prediction values, or the set of points in which the model estimate is lower than the corresponding estimates obtained for model predictions at a set of nearest-neighbor sampling vectors. The kriging procedure stops when the average kriging prediction values for the current iteration do not change significantly relative to the average kriging value for the previous iteration. A more detailed description of the steps of the algorithm of the interpolation technique is included in Appendix 3.

Modeling with Discrete Design Variables

Solving process design problems often involves two different types of decisions: (1) selection of components such as processes, type of equipment, etc. which are represented by discrete variables and (2) selection and determination of the operating conditions [20, 36] which are usually defined by continuous variables. Thus, process design is usually formulated as MINLP problem, where discrete/binary variables are introduced for representing existence of units and composing different design configurations. Modeling and optimization of pharmaceutical unit operations often includes modeling with discrete—sometimes even non-numerical—input variables. These parameters often represent design variables such as the use, size, or configuration of a specific part of a piece of equipment (e.g., screw size of a feeder, design of a nozzle). Up to now, however, data-driven modeling of pharmaceutical processes not often treats these types of variables as integer decision variables. Conversely, discrete or non- numerical variables are usually represented in a multivariate data set by coded values based on which the final response surface is fitted [37, 38]. Based on the methodologies proposed in the present work, individual models are produced for alternative process designs, which are complemented by the assignment of a decision variable for each design. A statistical analysis of the design variables is always performed initially to identify the variables that are statistically significant. Once variables are found to be significant, using the proposed modeling approach enables the formulation of the mathematical optimization problem that can lead to the optimum design for different values of operating conditions. This is the basic advantage of the proposed method since this cannot be achieved if the design is modeled as an additional input variable. Consequently, the design space can be constructed separately and more accurately for each different design configuration while the optimum design is identified for each combination of operating conditions.

Proposed Approaches

The purpose of this work is to develop a general methodology that will construct a graphical representation of the region bounded by the limits within which acceptable product or process performance is achieved. For this purpose, the only available knowledge on the system consists of a multivariate experimental data set of a desired output measured at different operating conditions for different design configurations. This is very often the problem faced in processes for which a physics- based model does not exist. The goal of designed experiments is to collect the necessary data in order to define relationships between input and output variables. The three surrogate-based modeling approaches described in "Data-Driven Modeling Methodologies" are used to model the effects of the input parameters on the output and the Design Spaces produced by each methodology are compared.

Mapping the Design Space Using RSM

The processes that are considered in this work have two shared characteristics: the absence of a reliable first-principle model and the existence of discrete type variables. Thus, in order to identify the Design Space using RSM, the following procedure is proposed:

Step 1: Second-order quadratic functions (Eq. 3) are fitted to the data of each design configuration using Response Surface methodology, producing k models; where k represents the total number of competing designs.

Step 2: The Knowledge Space of continuous input variables is discretized within their investigated ranges.

Step 3: For each combination of input variables, the optimum design and output are calculated. The existence of discrete type variables introduces a certain amount of complexity in the process of Design Space Mapping because for each combination of operating conditions and material properties, a different design configuration might be identified as optimal. Thus, a set of k binary variables is introduced in the objective functions which are associated with each design. Moreover, an additional constraint must ensure that only one design can be chosen for each combination of operating parameters.

Step 4: Finally, an acceptable range is set for the value of the output in order to identify the corresponding feasible operating ranges and material properties which assure the output is kept within these limits.

The detailed formulation of this concept is illustrated through its application to the case studies in the next section. It is clear that this procedure might create a discontinuous Design Space if for different values of operating conditions there is a change in the optimum value of the discrete input variable. The implementation of the described algorithm is achieved by the link between GAMS and MATLAB software [39].

Mapping the Design Space Using Cut-HDMR

In this proposed method, HDMR ("High-Dimensional Model Representations (HDMR)") methodology is used for modeling the behavior of the process.

Step 1: A nominal point for each design configuration is chosen. Next, for each of the k competing designs, second-order HDMR component functions are calculated, resulting to the production of k HDMR look-up tables.

Step 2: The Knowledge Space of continuous input variables is discretized within their investigated ranges.

Step 3: The output of each function at a specific combination of input variables is calculated by interpolating the corresponding component functions and adding them up to the second-order term.

Step 4: For each of the k calculated outputs, the optimum is identified by simply comparing the value of each design and storing the minimum or maximum-depending on the objective. The next step comprises of identifying the optimum operating conditions and design configurations within the entire range of the available knowledge space of input variables.

Step 5: An acceptable range is set for the value of the output in order to identify the corresponding feasible operating ranges and material properties which assure that the output is kept within the tolerable limits.

The main difference of this method from the RSM-based methodology is that there is no need to pre-postulate the response surface (i.e., quadratic form). The only assumption made is that second-order interactions of input variables are sufficient to predict the desired output. The described algorithm is implemented using MATLAB.

Kriging-Based Approach for Mapping the Design Space

This approach is based on the kriging algorithm ("Kriging Methodology") which—similar to HDMR—is an interpolating methodology. The steps of this proposed approach are the following:

Step 1: The Knowledge Space of continuous input variables is discretized within their investigated ranges.

Step 2: At each combination of operating conditions if output variable is experimentally measured, this measurement is used as an output. Otherwise the Kriging algorithm is called having as an input sampling set the entire experimental data set. The test set is simply the single point that is being handled in the current iteration. This procedure is continued iteratively until the output is measured for all possible combinations of operating conditions and material properties. This set of iterations is performed for each of the k competing designs.

Step 3: For each of the k calculated outputs, the optimum is identified by simply comparing the value of each design and storing the minimum or maximum—depending on the objective. The next step comprises of identifying the optimum operating conditions and design configurations within the entire range of the available knowledge space of input variables.

Step 4: Finally, an acceptable range is set for the value of the output in order to identify the corresponding feasible operating ranges and material properties which guarantees that the output is kept within the tolerable limits.

Steps 3–4 are identical to the HDMR-based methodology. The basic difference here is the interpolation technique, since each predicted point is calculated as a weighted function of the available experimental data points. For each predicted point, a variance is calculated—which provides a measure of prediction quality. The described algorithm is also implemented in MATLAB.

Case Studies

Continuous Mixer

In this case study, the performance of a commercial continuous powder mixer (Gericke GCM 250) is studied as a function of operating parameters and design variables. The parameters that have been identified as significant—through preliminary statistical analysis of variance (ANOVA)—involve the impeller rotation rate, the powder flow rate, and the design variables are the different Blade configurations ("All forward" and "Alternate Blades"). Forward direction imposes a forward flow along the axis of the mixer; backward direction imposes powder flow in the opposite direction. The first two variables are continuous whereas the last one has to be represented as a discrete variable. A detailed description of the experimental setup can be found in [40], while specifications of the equipment design can be found in Appendix 1.

The output measured to characterize the performance of the mixer is the relative standard deviation (RSD) of the concentration of the active pharmaceutical ingredient (API), which is in this case acetaminophen. Initially, pure acetaminophen (API) was pre-blended with a small amount of silicon dioxide (0.25%) in a V-blender and finally Avicel PH-200 was blended with acetaminophen (APAP). The RSD measurements are a result of the ratio of the SD of acetaminophen concentration over the average concentration of a large number of samples (Eq. 7). To determine the homogeneity of the stream coming out of the mixer, samples were retrieved from the outlet stream and the RSD, which is the most common mixing index used in industry, was computed based on Eq. 7. For each experimental run, 20 samples were collected from the outlet. Concentration of acetaminophen in each sample was measured using a NIR spectroscopy analytical method.

$$\mathrm{RSD} = \frac{s}{\overline{C}} = \frac{\text{Standard deviation}}{\text{Average concentration}} s = \sqrt{\frac{\sum_1^N \left(C_i - \overline{C}\right)^2}{N-1}} \tag{7}$$

In Eq. 4, $\overline{C}$ is the average concentration of the total samples (N) collected in each mixing run and C_i is the concentration of each sample. s is the standard deviation between the sample concentrations. The RSD index characterizes the uniformity of any mixture. In the context of continuous blenders, variability in the feeding rate of the materials often contributes significantly to the variability of the final mixture. For further details of the experimental procedure, refer to [40].

Firstly, the experimental data is used to produce different regression models for the two different design configurations (k=2): "All-Forward Blades" and "Alternate Blades", which result in the following equations:

All-forward blades:

$$\mathrm{RSD}_f = a_{of} + a_{1f} fr + a_{2f} rr + a_{11f} fr^2 + a_{22f} rr^2 + a_{12f} fr \cdot rr \tag{8}$$

Alternate blades:

$$\mathrm{RSD}_a = a_{oa} + a_{1a} fr + a_{2a} rr + a_{11a} fr^2 + a_{22a} rr^2 + a_{12a} fr \cdot rr \tag{9}$$

Where fr is the powder flow rate $\left(\frac{kg}{hr}\right)$ and rr is the impeller rotation rate (RPM).

The investigated operating space of different impeller rotation rates and powder flow rates is discretized and the minimum output is calculated at each node within this mesh, by solving the following MINLP problem:

$$\begin{aligned} \min\ & y_f \mathrm{RSD}_f + y_a \mathrm{RSD}_a \\ \text{s.t}\ & y_f + y_a = 1 \\ & fr_{\min} \le fr \le fr_{\max} \\ & rr_{\min} \le rr \le rr_{\max} \end{aligned}$$

where $fr_{\min}$, $rr_{\min}$, $fr_{\max}$, and $rr_{\max}$ represent the minimum and maximum values of the input variables (flow rate and rotation rate) in their experimentally investigated range (Knowledge Space). Discrete variables y_f and y_a can only take values of either 0 or 1 (binary variables) and are linked to the two different design configurations: all-forward and alternate blade configuration, respectively. Based on the formulation of the problem, a value of 1 indicates the use of the corresponding design. For example if y_f is equal to 1, then the all-forward design configuration is the optimum design for the specific conditions.

Using the kriging methodology, the experimental data set is used each time to estimate the output at the same combinations of design and operating variables. Look-up tables are produced for the two designs—using the HDMR-based method—from which the optimum output and design is also calculated within the created mesh.

Using all the above approaches, the calculated outputs within the knowledge space are compared in Fig. 2. Setting the upper limit threshold of $\mathrm{RSD}_{\max}$ of output acetaminophen

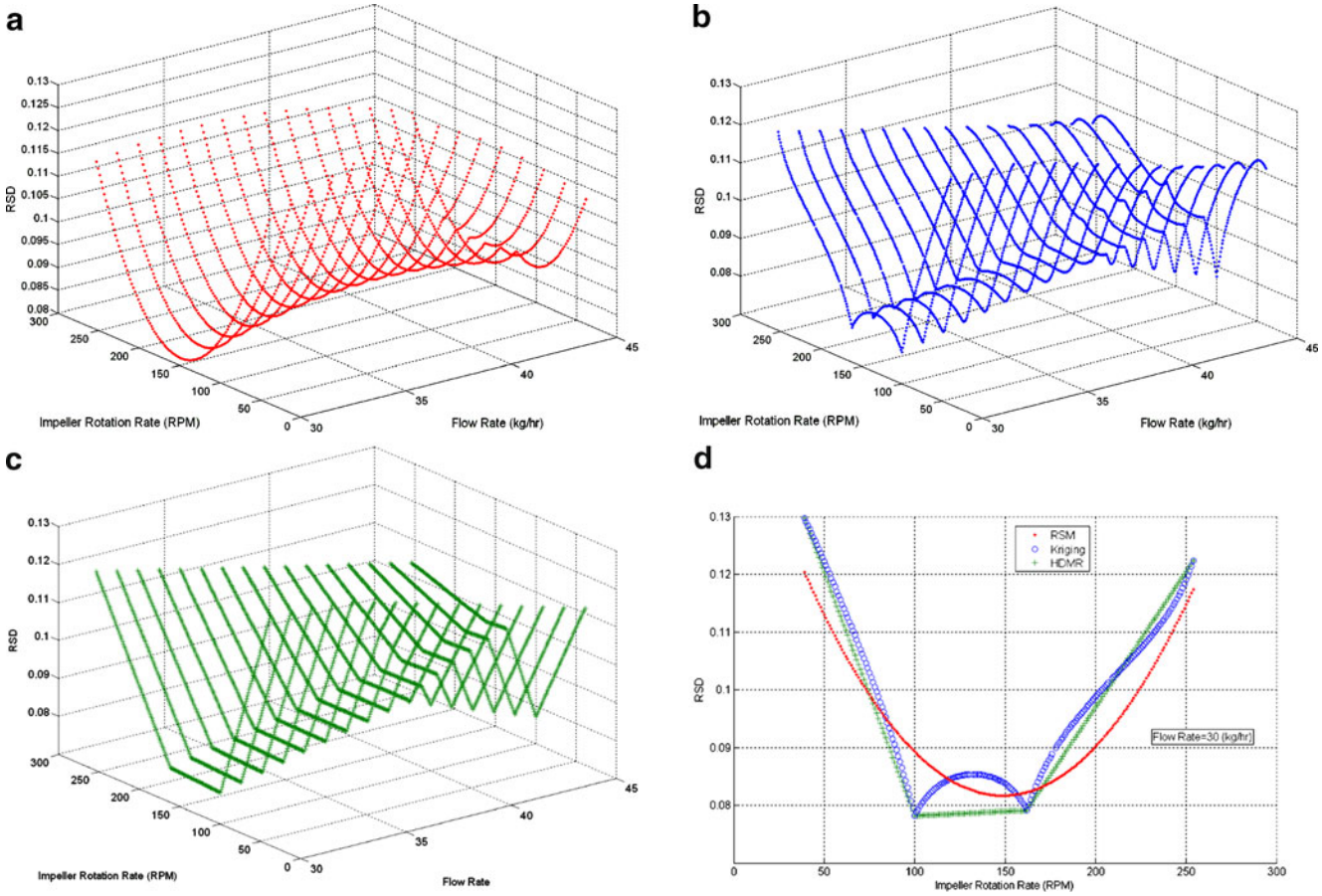


Fig. 2 Estimated output *RSD* within the range of investigated flow rate and impeller rotation rate values (**a**) predicted by RSM approach, (**b**) predicted by Kriging approach, (**c**) predicted by HDMR approach, and (**d**) predicted by RSM, Kriging, and HDMR approaches projected in 2D for a flow rate of 30 kg/h

concentration to be 10%, the acceptable design spaces calculated using all methods are mapped in Fig. 3. As a side note, we mention that the relatively high value of the RSD mentioned here (the usual acceptability limit for powder blends is 5%) is an artifact of the on-line spectroscopic technique used to quantify composition; RSD's measured by dissolving tablets generated by the process range around 2.5%.

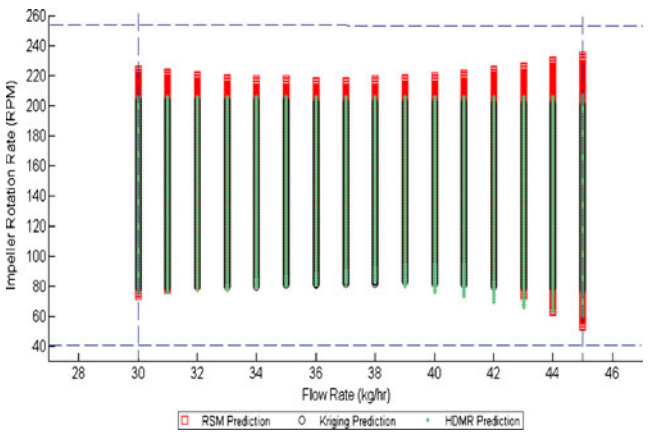


Fig. 3 Design space (flow rate and impeller rotation rate ranges) of continuous mixer for achieving $RSD_{max}=10\%$

Comparing the predicted values in Fig. 2, the similarity of the results of the three methods can be observed. As expected the RSM predictions result in smooth regions (Fig. 1d) due to the quadratic nature of the proposed model in this case. In the same Figure, the two sharp points of the kriging response are sampled points since kriging is an interpolating technique that uses the actual experimental values, when they are sampled. The form of the response predicted between the two sampled points using kriging depends highly on the number of nearby sampled points used in the weighted function. This is one parameter that can be tuned according to the nature of the response that is being modeled. In general, kriging has shown its best performance when space-filling designs are used, but factorial design samples were available in this case. The scope of the work, however, is to compare the performance when the same data are available. As shown in Fig. 3, which compares the Design Spaces, all three methods predict similar feasible operating ranges. It is clear that a feasible operation can be achieved for the entire range of

flow rates, but only for a mid-range of the impeller rotation rates. The outer dashed lines represent the limits of the investigated Knowledge Space.

The results obtained by the three data-driven methodologies agree with the expected results based on the existing knowledge of mixing process. In general, rotation rate is an important factor in continuous mixing [40]. Increasing the rotation rate leads to a higher degree of dispersion of powder in the mixer. However, higher impeller rotation rate also decreases the time available for mixing (lower residence time). Understanding these opposing effects is the key to achieving optimal mixing and can explain the fact that lowest RSD is achieved in mid-range rotation rates.

In contrast, flow rate is found to be the least significant factor through ANOVA (see Appendix 2) and this can explain the fact that for the investigated rotation rates, feasible operation can be achieved for all flow rates in the range of 30–45 kg/h. However, this variable must be included in the design space since it is a significant factor for feeder performance. Specifically, increasing the flow rate can decrease the output flow variability of the feeder. The fact that this variable does not affect the variability in concentration of this specific case study implies that the mixer can efficiently filter out the feeder variability, but this might not be the case for other feeding–mixing integrated systems.

Finally, in order to compare the effect of the different design configurations, the best design must be identified. This is an advantage of the proposed approaches, since for each combination of operating parameters, the design that can result in better performance can be identified (Fig. 4). For almost the entire design space, all three methods predict better performance when using the "Alternate Blades" configuration, while for high flow rates and low rotation rates, the "All-forward" design results to better mixing performance. This feature of the proposed approaches can be used in order to identify the best design configuration for a given set of operating parameters.

The efficiency of the three methods can be tested by setting a stricter lower output limit $RSD_{max}=9\%$ where the calculated design spaces are compared in Fig. 5. The shape of the design space changes significantly with this slight modification of the acceptable output range, but all three methods result in very similar design spaces. Large differences in the acceptable operating range when the performance criterion becomes more demanding (lower RSD), illustrates the sensitivity of the response to the changes in the input variables. More specifically, based on the models developed here, a small change in the acceptable variability significantly limits the design space area.

The comparison of different methodologies in this work aims at identifying areas where there is higher certainty that the output is within the acceptable ranges, and these are the regions for which all the methodologies jointly predict.

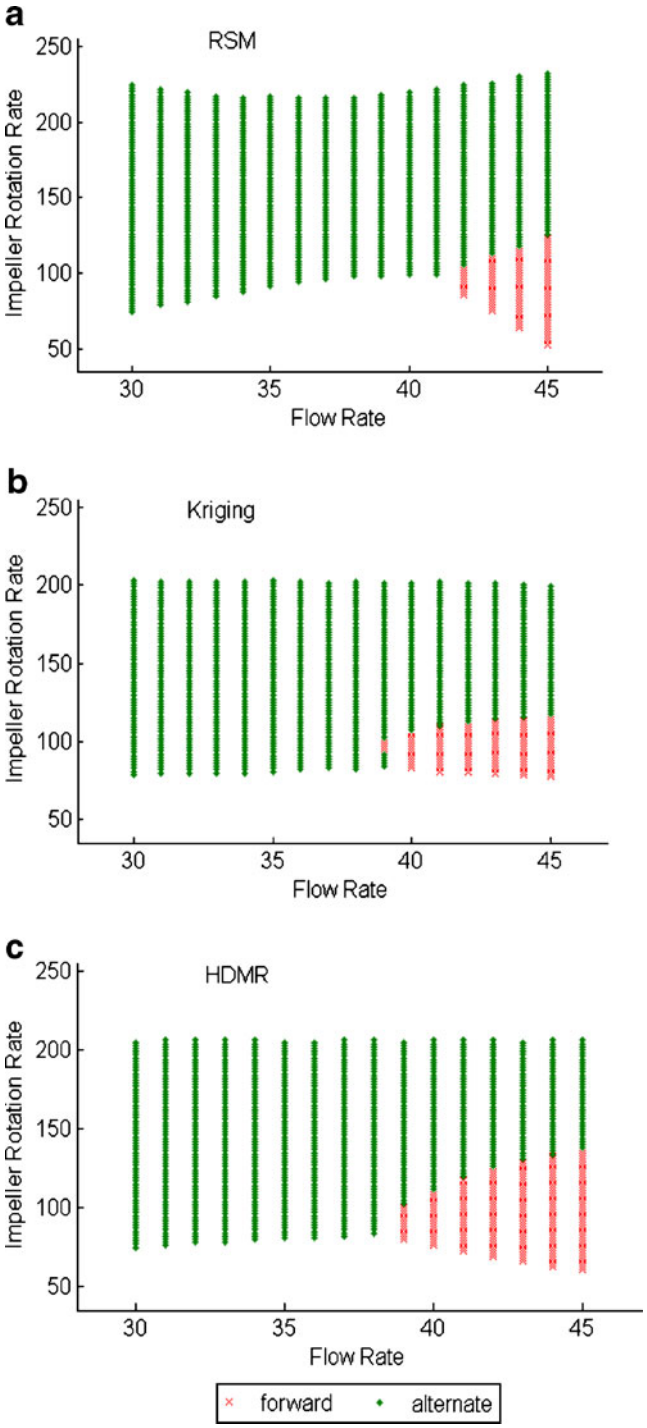


Fig. 4 Design space produced by (**a**) RSM approach, (**b**) Kriging approach, (**c**) HDMR approach showing the optimum design configurations for each point

Loss-in-Weight Feeder Case Study

In this case study, the performance of a Gericke Loss-in-Weight feeder is investigated. This piece of equipment is typically the first unit operation of an integrated system for

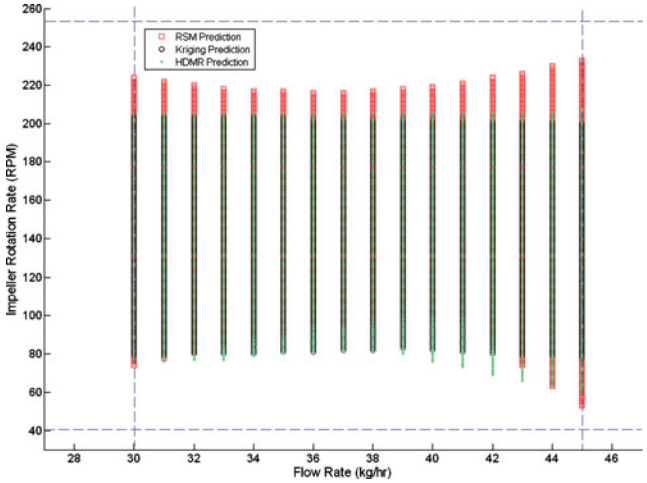


Fig. 5 Design space (flow rate and impeller rotation rate ranges) of continuous mixer for achieving RSD_{max}=9%

continuous manufacturing of pharmaceutical products (e.g., tablets) and its purpose is the feeding of APIs, excipients and lubricants, all of them powder materials, into the mixer. As it is common in many powder-based manufacturing processes, the knowledge of the physics that the material undergoes inside the feeder is not enough to conduct first-principles design, thus its performance can be characterized by data- driven models [22]. The significant input variables that affect the performance of this type of powder feeder are identified through ANOVA (see Appendix 2) as: Screw Speed inside the feeder (RPM), Screw size (sizes "2", "3", and "4"), Open or Closed Helix configuration (see Appendix 1) and the material property of Flow Index of the processed powder mixture. The property of Flow Index is connected to the flowability of the powder mixtures, where a high value can be translated to a poorly flowing powder [41]. The output established to characterize and quantify the performance of this process is the flow rate RSD, defined once again as the standard deviation of the output flow rate (quantified through a suitable "moving window" average) divided by the mean flow rate at the exit of the feeder. This measurement is chosen because it embodies the variability of the flow rate at the output of the feeder, which is important to be minimized since it will be transferred to the next process operations and consequently into the final product performance. This measurement is the RSD of the output flow and will be represented as RSD_{feed} for differentiation from the previous case study.

Following the RSM-based approach; the experimental data is used to generate different regression models for each design configuration (k=6): "Open size2", "Open size3", "Open size4", "Closed size2", "Closed size3" and "Closed size4". Details and specifications of these designs are included in Appendix 1. The investigated operating space of different Screw Speeds and Flow Indexes is discretized, and the optimum (minimum) output is calculated for each possible combination. In this case study, six quadratic models are generated (similar to Eqs. 7 and 8) and consequently, the MINLP formulation contains six binary variables from which only one can be equal to 1 for each point. Using all the proposed approaches, the calculated outputs are graphically represented in Fig. 6. By setting the upper threshold of acceptable output to be: RSD_{feed}=1.5% (arbitrary since it depends on the width of the moving window), the design spaces calculated using all three methods are mapped in Fig. 7.

Figure 6a shows that the calculated output of the feeder for different designs and within the Knowledge Space is similar when calculated with the three methods. However, there is a differentiation of the kriging predicted output variability, especially in the mid ranges of flow index and screw speed, which needs further investigation ("Experimental Validation"). In Fig. 6b the predicted RSD of the three methodologies is projected in two dimensions for a constant screw speed. In Fig. 7 different designs are represented by different symbols in order to highlight the advantage of the proposed approaches to identify the optimum design for a set of operating parameters and material properties. The calculated design spaces obtained when the acceptance limits of output variability is lowered to 0.014 are shown in Fig. 8. Large differences in the acceptable operating range when the performance criterion becomes more demanding (lower RSD), signifies the sensitivity of the response to the changes in the input variables. More specifically, based on the models developed here, a small change in the acceptable variability significantly limits the design space area.

In this case study, differences in the design spaces calculated with different approaches are observed. All methods, predict that screw sizes must be either 3 or 4 while size 2 screws always result to higher output variability. In both Figs. 7 and 8, size 2 screw is not included in the design space because it was not found to be optimal in any region within the feasible operating region. This illustrates that the larger size screws always give lower output variability when they are compared with smaller size screws. This conclusion is considered advantageous since by using a larger screw size (larger diameter, see Appendix 1) can result to higher flow rates. The fact that all methods predict performance at lower flow indexes is expected since higher flow index implies a poor flowability of the powder mixture [41]. The proposed approaches, however, have the advantage of identifying the optimal design configuration and screw speed ranges if a specific powder mixture must be processed. In order to validate the different approaches two additional experiments described in the next section were performed.

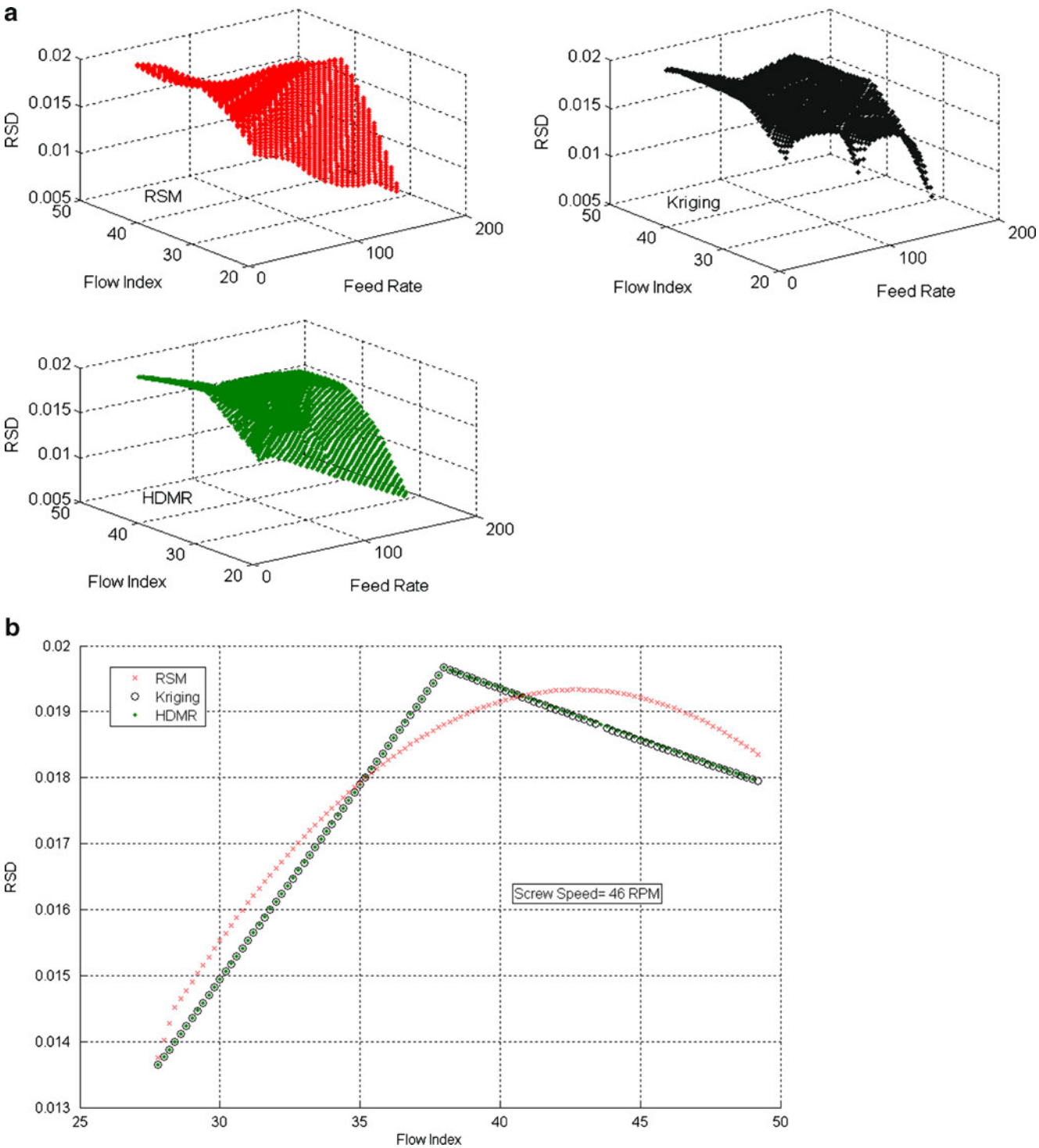


Fig. 6 Predicted output variability (RSD_{feed}) of a loss- in weight feeder within the investigated ranges of Flow Index and Screw Speed (RPM) (**a**) predicted by RSM, Kriging and HDMR approach and (**b**) predicted by RSM, Kriging and HDMR projected in 2D for Screw Speed 46RPM

Experimental Validation

The results of the feeder case study show that there are some differences in the predicted design spaces when using different methodologies. This ambiguity can be eliminated through experimental validation of the results. Experiments were carried out at the following operating conditions and a constant design configuration of Open Screw size 4. The powder used has a flow index of 27.8 and the two screw speeds investigated are: Screw Speed1=70 rpm and Screw Speed2=140 rpm.

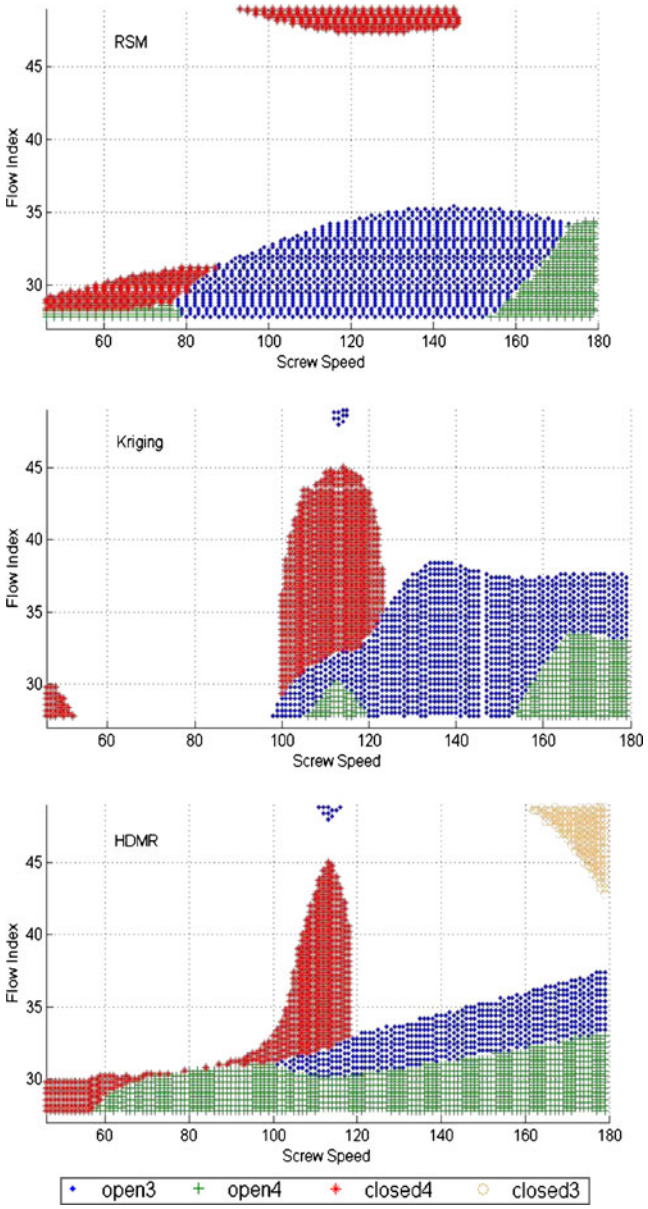


Fig. 7 Design space (flow index and screw speed ranges) of a loss-in-Weight feeder for acceptable output $RSD_{feed\_max}=1.5\%$

The two different experimental conditions were chosen in order to compare the accuracy of the three methodologies used (Fig. 8). Specifically operating at Screw Speed1 can verify the results of kriging approach, compared with the other two methods. Operating at Screw Speed2 can similarly compare the results of HDMR to the results of the remaining two approaches. The measured RSD_{feed} at Screw Speed 1 is 2.95% (>1.4%) which suggests that kriging is successful in not predicting feasible operation. Operation at Screw Speed 2 has a resulting Rsd_{feed} of 1% (<1.4%). This result proves that the HDMR predicted design space is again not validated; on the contrary RSM and kriging approaches are proven to be more accurate.

The experimental validation procedure used in this case study serves as a method for identifying the differences between traditional quadratic surface fitted functions, based on Design of Experiment theories (i.e., RSM) and stochastic interpolating techniques (i.e., kriging). It has been previously shown in the literature [42] that quadratic surfaces can fail to identify global optima of more complex, non-smooth functions. Regression techniques, such as RSM, focus entirely on the estimation of the regression coefficients while making simplistic assumptions about the errors (indepen-

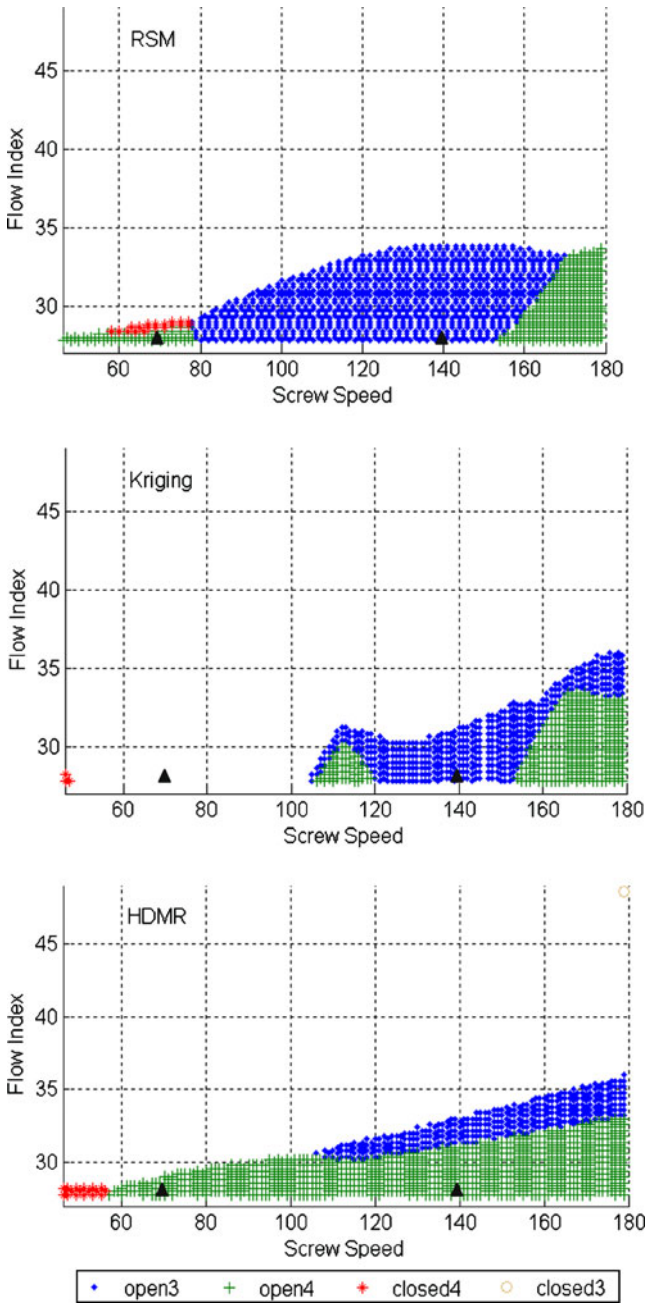


Fig. 8 Design space (flow index and screw speed ranges) of a loss-in-weight feeder for acceptable output $RSD_{feed\_max}=1.4\%$

dence, normality). They also require a predefined form of the response surface on which the regressors are fitted. On the other hand, interpolating methods such as kriging focus more on the correlation structure of the errors. Using such approaches also provides an estimate of prediction accuracy. In fact, a point which is close to a sampled point (with prediction uncertainty equal to zero) has a small variance value. If a point is far from a sampled point, its variance will be high. Comparing kriging and HDMR methodologies is interesting, since they are both not based on a predefined model; they are both interpolating techniques. The only difference is that in the kriging approach the predicted output is calculated as a weighted sum of Euclidean distances from nearby sampling points. On the other hand, in HDMR, component function values are calculated and simple linear interpolation is used to calculate the output at intermediate operating conditions.

Furthermore, one of the aforementioned advantages of the kriging algorithm is the calculation of the uncertainty (variance) for each estimated point. This represents the range within which the estimated point may lie with a certain probability, and it can identify areas with high variance that more experimentation is needed. These variances are plotted in Fig. 9 for both case studies described in the previous section. In this Figure, it is clear that the variances are higher in the feeder case study. Specifically, as the predicted point is further away from the points where there exists an experimental measurement, the variances are higher. These regions can be identified as high uncertainty regions where possibly further experimentation is necessary. This can be considered as the reason for which the results in case study 2 are not consistent between the different methodologies; as opposed to case study 1, where the calculated design spaces are very similar.

Conclusions and Future Work

In this work, three different approaches are used to graphically map the Design Space of pharmaceutical

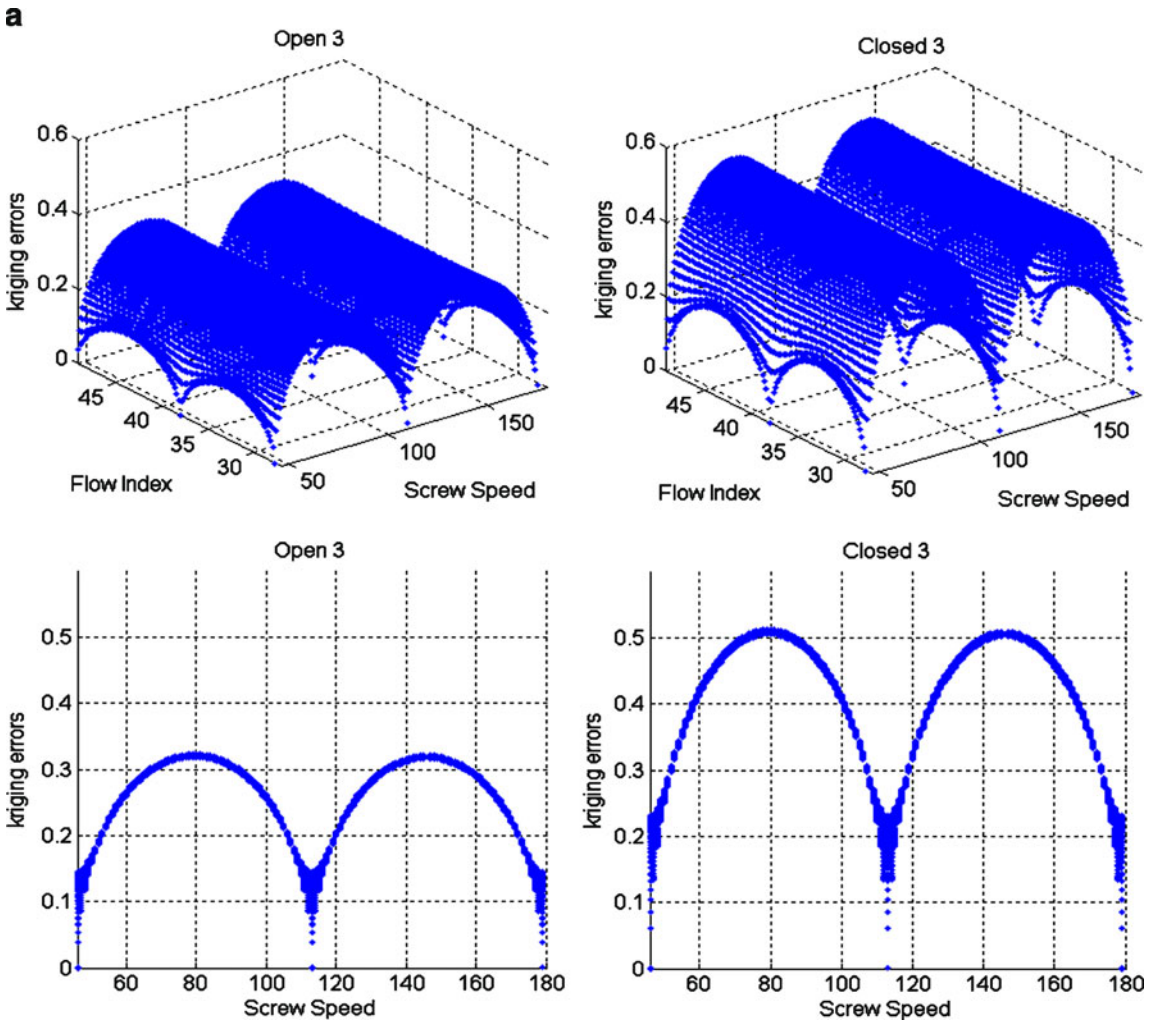


Fig. 9 Kriging variances (**a**) for feeder case study and (**b**) for mixer case study plotted in 3D and projected in 2D

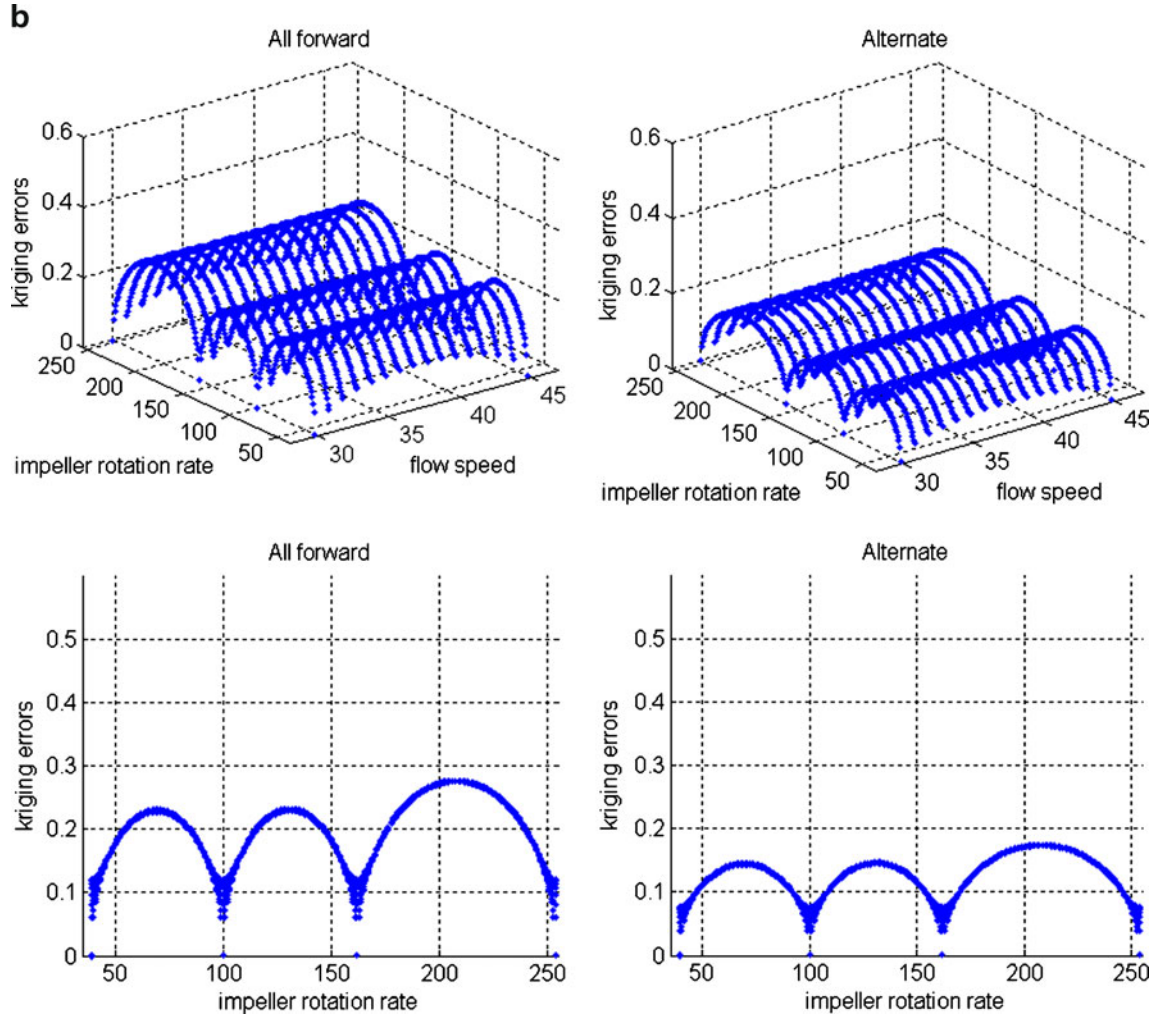


Fig. 9 (continued)

process unit operations in order to identify the ranges of the continuous input variables as well as the design configurations that result to an output that satisfies certain qualitative or quantitative constraints. The methodologies used are not restricted by the lack of first-principle models that characterize the performance of a process and can be used for any black-box process operation. In [1], the steps of the procedure of design space identification—right from the initial drug conceptualization stage, to the final commercialization of the product—are described. This work aims to propose a set of tools and enabling technologies to simultaneously identify desirable operating regions and optimal design configurations of black-box pharmaceutical processes. More specifically, three different methodologies are proposed as the means to extract useful information from an experimental data set of a process in order to be able to operate in regions which result to desired performance with confidence. The novelty lies in the use of optimization programming techniques to compare performance of different possible designs within the normal operating ranges.

Response Surface Methodology is shown to produce a good representation of the design space when it is used to generate models for different design configurations which jointly form an objective function of a mixed integer optimization problem. High-Dimensional Model Representation second-order function look-up tables are also used for interpolation of the desired output at any given combination of input variables and it is shown to produce similar design space representations. Both methodologies, however, do not provide a measure of prediction uncertainty. On the contrary, when kriging is used to model the effects of input parameters to the output of the process, not only a design space representation is produced (comparable to the ones produced from the aforementioned approaches) but also a prediction variance is calculated for

each predicted point. This additional information can help identify regions in which additional sampling is required. Experimental validation of the results in uncertain areas proves that the kriging-based approach is more accurate in predicting the design space. Additionally, all three methodologies are shown to be responsive to changes in the desirable output ranges—always following the same trends. In this work, the differences are assessed between the three methods, given the same experimental data, which is obtained from a common design of experiment sampling procedure. In the mixer case study, all three methods show good agreement. In the feeder case study however, some differences are observed between the results of the different methods. This fact increases the uncertainty in the results and this is why further experimental validation is performed. The kriging error, which is a function of the distance between sampled areas, illustrates the benefit of this technique to identify regions of high uncertainty as candidate regions for further experimentation. In the literature, kriging has shown to have a better performance when space-filling designs are used for sampling. Future work will include investigating the effects of sampling and the minimization of the necessary sampling for accurate results, since this will also lead to minimization of sampling cost.

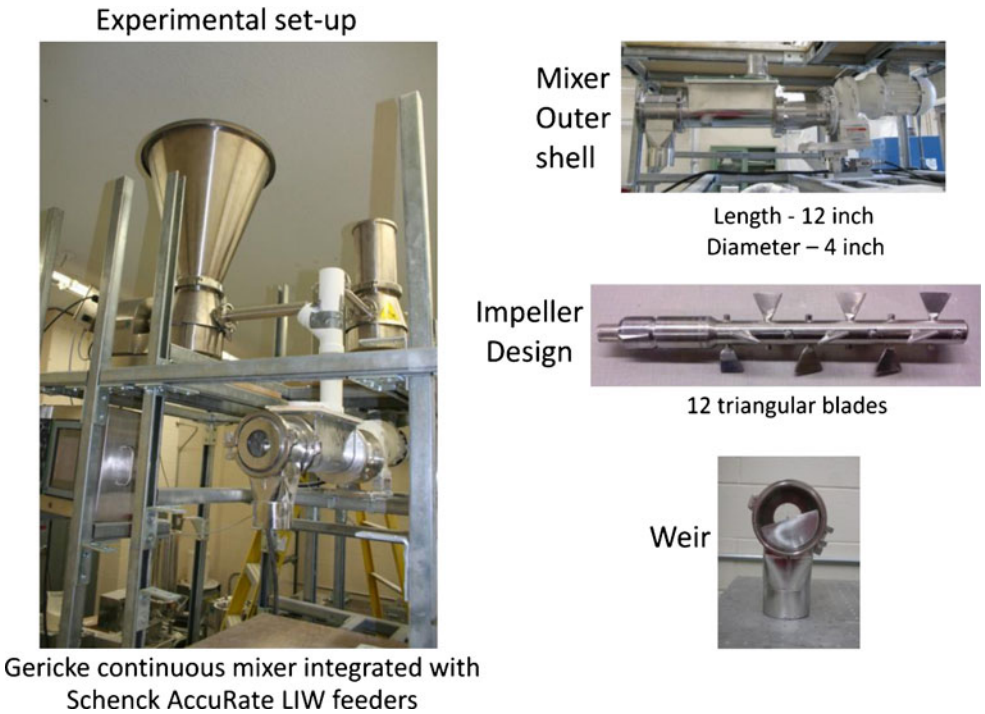
Future work will also focus on employing the proposed methods to integrated process systems (e.g., continuous tablet manufacturing process consisting of feeders, mixer, tablet press etc.) so that the design space will include parameters from all unit operations while the specified output will be connected to acceptable final product properties. Current work addresses the use of the proposed methodologies for modeling and defining the design space of dynamic systems. All the above will provide a general framework for mapping the design space of any process operation for which first-principle models are not yet available.

Acknowledgments This work was supported by the ERC-SOPS (NSF-0504497, NSF-ECC 0540855). Also special thanks to Bill Englisch and Aditya Vanarase for providing the experimental data.

Appendix 1—Equipment Specifications

Gericke GCM 250 Mixer Specifications

Length	0.3 m
Diameter	0.1 m
Impeller	12 triangular equally spaced shaped blades



Gericke Loss-in-Weight Feeder Specifications

The following picture shows (from left to right) the main agitator piece that is used for all screws, the size 2 screw dependant agitator piece, open helix screw and closed auger screw, followed by sizes 3 and 4 sets.

The following is a picture of the feeder with the tooling before the nozzle plate is attached.

Appendix 2—Experimental Data and Analysis of Variance (ANOVA)

Table 1 Continuous mixer case study

Flow rate (kg/h)	Blade configuration (all forward=1, alternate=2)	Impeller rotation rate (RPM)	RSD
30	1	39	0.131324
30	1	100	0.103559
30	1	162	0.103234
30	1	254	0.179275

Table 1 (continued)

Flow rate (kg/h)	Blade configuration (all forward=1, alternate=2)	Impeller rotation rate (RPM)	RSD
30	1	39	0.138515
30	1	100	0.112030
30	1	162	0.098612
30	1	254	0.160198
30	2	39	0.137095
30	2	100	0.090748
30	2	162	0.094803
30	2	254	0.146097
30	2	39	0.122224
30	2	100	0.065687
30	2	162	0.063477
30	2	254	0.098588
45	1	39	0.098759
45	1	100	0.073322
45	1	162	0.119089
45	1	254	0.180556
45	1	39	0.123905
45	1	100	0.080957
45	1	162	0.103170
45	1	254	0.191361
45	2	39	0.132929
45	2	100	0.099215
45	2	162	0.084938
45	2	254	0.096425
45	2	39	0.125567
45	2	100	0.096007
45	2	162	0.108194
45	2	254	0.110720

Table 2 ANOVA table for continuous mixer case study

Source	*df*	SS	MS	*F*	*P*
Flow rate	1	0.0000129	0.0000129	0.03	0.868
Blade configuration	1	0.0033038	0.0033038	7.15	0.013
Impeller rotation rate	3	0.0159447	0.0053149	11.50	0.000
Error	26	0.0120158	0.0004621		
Total	31	0.0312773			

S=0.0214976; R^2 =61.58%; R^2 (adj)=54.20%

Table 3 Loss-in-weight feeder case study

RPM	Flow Index	Open/closed	Size	SD/Mean
46	38	O	2	0.037529295
113	38	O	2	0.024045867
179	38	O	2	0.020045903
46	38	C	2	0.039382739
113	38	C	2	0.022863443
179	38	C	2	0.015842134
46	38	O	3	0.026314162
113	38	O	3	0.018842017
179	38	O	3	0.015267703
46	38	C	3	0.022191219
113	38	C	3	0.014792139
179	38	C	3	0.012924242
46	38	O	4	0.020567412
113	38	O	4	0.023541078
179	38	O	4	0.017691507
46	38	C	4	0.01966698
113	38	C	4	0.014381375
179	38	C	4	0.021190405
46	27.8	O	2	0.020769979
113	27.8	O	2	0.011146501
179	27.8	O	2	0.012334908
46	27.8	C	2	0.03079855
113	27.8	C	2	0.020359721
179	27.8	C	2	0.018491395
46	27.8	O	3	0.015730878
113	27.8	O	3	0.011410378
179	27.8	O	3	0.008650336
46	27.8	C	3	0.016490942
113	27.8	C	3	0.011924303
179	27.8	C	3	0.013519275
46	27.8	O	4	0.014521154
113	27.8	O	4	0.010099635
179	27.8	O	4	0.00593691
46	27.8	C	4	0.013643659
113	27.8	C	4	0.014780312
179	27.8	C	4	0.017804938
46	49.2	O	2	0.044822316
113	49.2	O	2	0.028094832
179	49.2	O	2	0.019076613
46	49.2	C	2	0.0376905
113	49.2	C	2	0.017728798
179	49.2	C	2	0.013834081
46	49.2	O	3	0.026672266
113	49.2	O	3	0.014458669
179	49.2	O	3	0.022003755
46	49.2	C	3	0.027508537
113	49.2	C	3	0.022332606
179	49.2	C	3	0.02316389
46	49.2	O	4	0.024397465
113	49.2	O	4	0.020911612
179	49.2	O	4	0.021890465
46	49.2	C	4	0.017947927
113	49.2	C	4	0.015328795
179	49.2	C	4	0.015000022

Table 4 ANOVA table for feeder case study

Source	*df*	SS	MS	*F*	*P*
Screw Speed	2	0.00085605	0.00042802	17.27	0.000
Flow Index	2	0.00065948	0.00032974	13.30	0.000
Screw Type	1	0.00000050	0.00000050	0.02	0.888
Screw size	2	0.00052281	0.00026141	10.55	0.000
Error	46	0.00114023	0.00002479		
Total	53	0.00317906			

S=0.00497871; R^2 =64.13%; R^2 (adj)=58.68%

Appendix 3—Interpolation Techniques

A. High-Dimensional Model Representation Algorithm

The HDMR algorithm steps are the following:

1. Obtain input- output data $g\ (x_1,\ x_2,...,x_n)$ for n input variables. (Appendix 2)
2. Choose nominal point conditions $x^N = \left(x_1^N, \ldots, x_n^N\right)$ to calculate f_0 term as:

$$f_0 = g\left(x_1^N, \ldots, x_n^N\right) \qquad (3A-1)$$

3. Calculate the first-order component function $f_i\ (x_i)$ values by keeping all other input variables at their nominal values only varying x_i, and always subtracting f_0. The number of component function values computed are equal to the experimentally sampled conditions along the x_i axes.

$$f_i(x_i) = g\left(x_i, x_j^N\right) - f_0, i \neq j \qquad (3A-2)$$

4. Calculate the second-order component functions $f_{i,j}\ (x_i,\ x_j)$ values by keeping all other variables other than $x_i,\ x_j$ at their nominal values and subtracting the corresponding lower order function values based on the following equation:

$$f_{i,j}\left(x_i, x_j\right) = g\left(x_i, x_j, x_k^N\right) - f_i(x_i) - f_j\left(x_j\right) - f_0, k \neq i,j \qquad (3A-3)$$

5. Create a look-up table of component functions based on at different values of input variables. Once the look-up tables are created, the HDMR predicted output can be calculated as the sum of component functions up to the second-order term (Eq. 3A-4). If the output at an unsampled position needs to be calculated, linear interpolation is performed for each component function and finally the predicted value is calculated based on the sum of all terms from Eq. 3A-4.

$$f_{HDMR}(x_1,\ldots,x_n) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{1\le i\le j\le n} f_{i,j}(x_i,x_j) \qquad (3A-4)$$

B. Kriging Algorithm

The Kriging Algorithm steps are the Following:

1. Obtain input – output data (Appendix 2) $f(x_i)$ for N sampling points x_i
2. For any combination of a pair of sampling points x_i - x_j, calculate the distance between them as:

$$d_{i,j} = \sqrt{\left(x_{1,i}-x_{1,j}\right)^2 + \cdots + \left(x_{n,i}-x_{n,j}\right)^2} \text{ for } i,j = 1,\ldots,N, i \neq j \qquad (3B-1)$$

In total there will be $N(N–1)/2$ sampling pairs x_i–x_j.

3. For all sampling pairs, the squared difference between the corresponding output values is calculated next as:

$$f_{i,j}^2 = \left(f(x_i) - f(x_j)\right)^2 \qquad (3B-2)$$

4. The next step is to plot the corresponding squared differences (Eq. 3B-2) with respect to their distances (Eq. 3B-1) to form a scatterplot that will give us the semivariogram function $\gamma(h)$. A function that best describes the scattered data is fitted out of the following possible functions: exponential, Gaussian, spherical, linear and power models. The semivariogram model is necessary for the computation of the kriging weights. If needed, data smoothing is first applied to the Scatterplot before the semivariogram fitting. The best model is identified as the one that has the minimum least squares error.
5. The peak value of the chosen semivariogram model is identified as $\sigma^2_{\max}$ and it is used to calculate the covariance function which -based on definition- is given by:

$$\mathrm{Cov}(h) = \sigma^2_{\max} - \gamma(h) \qquad (3B-3)$$

6. Based on Eq. 3B-3, the covariance between any pair of sampling points x_i–x_j can be obtained. Given a new unsampled point, first a number of sampled points that will affect the predicted value of the test point are chosen. In this algorithm, the seven closest points to the test point are always chosen. The kriging prediction of the unsampled point is expressed as a weighted sum of the kriging weights multiplied by their corresponding sampled point output values (Eq. 3B-4). The kriging weights depend on the covariances between sampling point pairs as well as the covariances between sampling points and test point and are given by the solution of the system of equations given by Eq. 5. The sum of the weights for any test point must be equal to 1.

$$\widetilde{f}(x_k) = \sum_{i=1}^{7} w_i f(x_i) \qquad (3B-4)$$

7. The final step is the calculation of the variance estimate of the test point which is given by Eq. 6.

References

1. Lepore J, Spavins J. PQLI design space. J Pharmaceut Innovation. 2008;3(2):79–87.
2. Garcia T, Cook G, Nosal R. PQLI key topics - criticality, design space, and control strategy. J Pharmaceut Innovation. 2008;3(2):60–8.
3. Davis B, Lundsberg L, Cook G. PQLI control strategy model and concepts. J Pharmaceut Innovation. 2008;3(2):95–104.
4. Nosal R, Schultz T. PQLI definition of criticality. J Pharmaceut Innovation. 2008;3(2):69–78.
5. Lipsanen T, Antikainen O, Räikkönen H, Airaksinen S, Yliruusi J. Novel description of a design space for fluidised bed granulation. Int J Pharm. 2007;345(1–2):101–7.
6. Lebrun P, Govaerts B, Debrus B, Ceccato A, Caliaro G, Hubert P, et al. Development of a new predictive modelling technique to find with confidence equivalence zone and design space of chromatographic analytical methods. Chemom Intell Lab Syst. 2008;91(1):4–16.
7. Halemane KP and Grossmann IE. Optimal process design under uncertainty. 1987 cited. Available from: http://hdl.handle.net/1903/4569
8. Swaney RE, Grossmann IE. An index for operational flexibility in chemical process design. Part I: Formulation and theory. AIChe J. 1985;31:621–30.
9. Floudas CA, Gumus ZH. Global optimization in design under uncertainty: feasibility test and flexibility index problems. Ind Eng Chem Res. 2001;40(20):4267–82.
10. Grossman IE, Floudas CA. Active constraint strategy for flexibility analysis in chemical processes. Comput Chem. 1987;11:675–93.
11. Vishal G, Marianthi GI. Determination of operability limits using simplicial approximation. AIChE J. 2002;48:2902–9.
12. Vishal G, Marianthi GI. Framework for evaluating the feasibility/operability of nonconvex processes. AIChe J. 2003;49:1233–40.
13. Georgakis C, Uztürk D, Subramanian S, Vinson DR. On the operability of continuous processes. Control Eng Pract. 2003;11(8):859–69.
14. Vinson DR, Georgakis C. A new measure of process output controllability. J Process Control. 2000;10(2–3):185–94.
15. Subramanian S, Georgakis C. Steady-state operability characteristics of reactors. Comput Chem Eng. 2000;24(2–7):1563–8.

16. Subramanian S, Georgakis C. Steady-state operability characteristics of idealized reactors. Chem Eng Sci. 2001;56(17):5111–30.
17. Subramanian S, Uzturk D, Georgakis C. An optimization-based approach for the operability analysis of continuously stirred tank reactors. Ind Eng Chem Res. 2001;40(20):4238–52.
18. Lima F, Jia Z, Ierapetritou M, Georgakis C. Similarities and differences between the concepts of operability and flexibility: The steady-state case. AIChe J. 2010;56:702–16.
19. Banerjee I, Ierapetritou MG. Design optimization under parameter uncertainty for general black-box models. Ind Eng Chem Res. 2002;41(26):6687–97.
20. Banerjee I, Ierapetritou MG. Parametric process synthesis for general nonlinear models. Comput Chem Eng. 2003;27(10):1499–512.
21. Floudas CA. Nonlinear and mixed-integer optimization: fundamentals and applications. New York: Oxford University Press; 1995.
22. Jia Z, Davis E, Muzzio FJ, Ierapetritou MG. *Predictive modeling for pharmaceutical processes using kriging and response surface*. JPI, 2009;4:174.
23. Boukouvala F, Muzzio F, Ierapetritou M. Predictive modeling of pharmaceutical processes with missing and noisy data. AIChe J, 2010 cited; Available from: http://onlinelibrary.wiley.com/doi/10.1002/aic.12203/full.
24. Chowdhury R, Rao BN. Assessment of high dimensional model representation techniques for reliability analysis. Probab Eng Mech. 2009;24(1):100–15.
25. Genyuan Li S-WW, Herschel Rabitz. High dimensional model representations (HDMR): concepts and applications. cited. Available from: http://www.ima.umn.edu/talks/workshops/3-15-19.2000/li/hdmr.pdf.
26. Pistek M. High dimensional model representation. cited. Available from: as.utia.cz/files/phdws06/fullpaper_file_53.pdf
27. Rabitz H, Aliş Ö. General foundations of high–dimensional model representations. J Math Chem. 1999;25(2):197–233.
28. Rabitz H, Alis ÖF, Shorter J, Shim K. Efficient input–output model representations. Comput Phys Commun. 1999;117(1–2):11–20.
29. Sobol IM. Theorems and examples on high dimensional model representation. Reliab Eng Syst Saf. 2003;79(2):187–93.
30. Li G, Rosenthal C, Rabitz H. High dimensional model representations. J Phys Chem A. 2001;105(33):7765–77.
31. Li G, Wang S-W, Rabitz H. Practical approaches to construct RS-HDMR component functions. J Phys Chem A. 2002;106(37):8721–33.
32. Box GEP, Wilson KB. On the experimental attainment of optimum conditions. J R Stat Soc B Methodol. 1951;13(1):1–45.
33. Raymond HM, Douglas CM. Response surface methodology: process and product in optimization using designed experiments. New York: Wiley; 1995. p. 728.
34. Cressie N. Statistics for spatial data (Wiley Series in Probability and Statistics). New York: Wiley; 1993. p. 1993.
35. Isaaks E. SR, *Applied Geostatistics*. New York: Oxford University Press; 1989.
36. Grossmann IE. Mixed-integer nonlinear programming techniques for the synthesis of engineering systems. Res Eng Des. 1990;1(3):205–28.
37. Myers RH, Classical and modern regression with applications (second edn). The Duxbury advanced series in statistics and decision sciences, ed. D. Press. PWS-KENT, Boston, MA, 1990
38. Myers RH, Montgomery DC. Response surface methodology: process and product in optimization using designed experiments. New York: Wiley; 1995. p. 728.
39. Ferris MC. MATLAB and GAMS: interfacing optimization and visualization software. 2005 cited; Available from: http://pages.cs.wisc.edu/~ferris/matlabgams.pdf.
40. Vanarase AU, Muzzio F. Effect of operating conditions and design parameters in a continuous powder mixer. Adv Powder Tech, 2010; (in press).
41. Faqih A, Chaudhuri B, Alexander AW, Davies C, Muzzio FJ, Silvina Tomassone M. An experimental/computational approach for examining unconfined cohesive powder flow. Int J Pharm. 2006;324(2):116–27.
42. Jones DR, Schonlau M, Welch WJ. Efficient global optimization of expensive black-box functions. J Glob Optim. 1998;13(4):455–92.