

Desirability function approach: A review and performance evaluation in adverse conditions

Nuno R. Costa^{a,c,*}, João Lourenço^{a,b}, Zulema L. Pereira^c

^a Escola Superior de Tecnologia de Setúbal, Campus do IPS, Estefanilha, 2910-761 Setúbal, Portugal

^b INESC-ID, Rua Alves Redol 9, 1000-029 Lisboa, Portugal

^c UNIDEMI/DEMI, Faculdade de Ciências e Tecnologia-Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

ARTICLE INFO

Article history:

Received 12 January 2011

Received in revised form 7 April 2011

Accepted 12 April 2011

Available online 21 April 2011

Keywords:

Bias
Multiresponse
Optimization
Response surface
Robustness
Variance

ABSTRACT

Adverse conditions in terms of quality of predictions and robustness are simulated to evaluate the ability of desirability-based methods for yielding compromise solutions with desired response's properties. The method's solutions are assessed at optimal variable settings with respect to bias, quality of predictions and robustness through optimization measures, and the usefulness of those measures to select the compromise solution is evaluated. Three examples with different features in terms of responses variance are used and the performance of various analysis methods is compared. Results show that a less sophisticated desirability-based method can compete with other methods designed to perform well under adverse conditions and that the optimization measures justify its use in real life problems.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

A widely used methodology for developing, improving and optimizing systems (process and product), the so-called response surface methodology (RSM), consists of the following general phases: 1) screening: experiments are designed with the purpose of discovering the vital few control factors that cause statistically significant effects of practical importance for the goal of the study; 2) modeling: experiments are designed with the purpose of modeling the quality characteristic of interest (response) as a function of control factors; 3) optimization: response model is analyzed to determine the variable settings at which optimum conditions of system property are achieved.

This methodology has inherent sequential experimentation strategy that, if the technical and non-technical issues in the experimental phases are properly managed, leads to a high level of system knowledge. For an understanding of the assumptions and conditions necessary to successfully apply RSM the reader is referred to Refs. [1–3].

This article focuses on the optimization phase and, in particular, on the optimization of multiple responses. These problems are usual in various fields of science and often involve incommensurate and conflicting responses that must be in some sense optimized

simultaneously, because their separate analysis may result in incompatible solutions. This has been a much researched subject and a strategy widely used in the RSM framework consists of converting the multiple responses into a single one by combining the individual responses into a composite function followed by its optimization. Although the desirability function and loss function approaches are popular among practitioners, other approaches have been used for optimizing multiple responses. Among them are those based on Compromise Programming [4], Goal Programming [5,6], Inspection of contour plots [7], Physical Programming [8,9], Probability-based [10], Performance Index [11,12], Neural Networks [13,14], and Vectorial Optimization [15]. All these approaches have their own merits. However, the lack of recommendations for proper use, unavailability of the algorithms employed, and (mathematical/statistical) complexity of some approaches and methods are major reasons by which they are of little practical use or appealing to practitioners, in particular to non-statisticians.

The less sophisticated desirability-based methods are easy to understand, easy to use, and flexible for incorporating the decision-maker's preferences (weights or priorities assigned to responses). Moreover, the most popular desirability-based method, the Derringer and Suich's method [16] or modifications of it [17] are available in many data analysis software packages. This method has been extensively used in practice, namely in chemometrics, such as illustrated in Refs. [18–22]. However, the analyst needs to specify values to four types of parameters (shape factors and weights) to use the so-called Derringer and Suich's method. This is not a simple task

* Corresponding author at: Escola Superior de Tecnologia de Setúbal, Campus do IPS, Estefanilha, 2910-761 Setúbal, Portugal. Tel.: + 351 265790000.

E-mail addresses: nuno.costa@estsetubal.ips.pt (N.R. Costa), joao.lourenco@estsetubal.ips.pt (J. Lourenço), zlp@fct.unl.pt (Z.L. Pereira).

and impacts on the optimal variable settings. So, other less known and used methods that require less subjective information from the user and can yield effective compromise solutions are particularly welcome, namely to non-statistician practitioners who are using statistics more than ever before [23].

This article provides a review on existing desirability-based methods and aims at evaluating the ability of two less sophisticated ones along with appropriate optimization performance measures to give desirable response's properties at optimal variable settings, namely low bias (responses deviation from target) and low variance. This is illustrated with three case studies where adverse conditions in terms of responses variance are simulated. In the first simulated scenario the response's models differ in terms of the quality of predictions (variance due to uncertainty in the regression coefficients). In the second simulated scenario the response's models are characterized by unequal sensitivity to uncontrollable variables (robustness) and high quality of predictions. In the third one, both the quality of predictions and robustness are low.

The remainder of the article is organized as follows: Section 2 provides a review on desirability-based methods; optimization criteria and measures to evaluate the compromise solutions in terms of response's bias, quality of predictions and robustness are introduced in Section 3; Section 4 characterizes the simulated scenarios and includes three examples to illustrate the feasibility of the proposed approach; Section 5 includes the discussion of results and Section 6 presents the final conclusions.

2. Literature review

Desirability function-based approach consists of converting the estimated response models ($\hat{y}$), which usually are second order models, into individual desirability functions (d) that are then aggregated into a composite function (D). This function is usually a geometric or an arithmetic mean, which will be maximized or minimized, respectively.

Next subsections provide a chronological review on the existing desirability-based methods, which are separated into two groups of different mathematical/statistical sophistication levels. The first group includes the less sophisticated approaches, while the second group includes the more sophisticated approaches.

2.1. Less sophisticated approaches

This subsection reviews the methods proposed in Refs. [16,17] and [24–28]. These methods do not require from the analyst high mathematical and statistical expertise as they are easy to understand and can be easily implemented in Excel®.

2.1.1. Derringer and Suich [16]

These authors generalized the exponential type transformations proposed in Ref. [29] to transform the responses into desirability functions and developed more flexible desirability functions. Derringer and Suich proposed individual desirability functions based on three response types as follows:

A. Nominal-The-Best (NTB): the value of the estimated response is expected to achieve a particular target value (T). For this response type, the individual desirability function is defined as

$$d = \begin{cases} \left(\frac{\hat{y}-L}{T-L}\right)^s, & L \leq \hat{y} \leq T \\ \left(\frac{\hat{y}-U}{T-U}\right)^t, & L \leq \hat{y} \leq T \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where s and t are user-specified parameters ($s, t > 0$) that allow practitioners to specify the shape of d . The individual desirability function transforms the response variable into a range of values between 0 and 1, where 1 is most favorable. In the NTB case, $d = 1$ for $\hat{y} = T$, and $d = 0$ for $\hat{y} < L$ or $\hat{y} > U$, where U is the upper and L is the lower specification limit of the response.

B. Larger-The-Best (LTB): The value of the estimated response is expected to be larger than a lower bound ($\hat{y} > L$). For this response type, the individual desirability function is defined as

$$d = \left(\frac{\hat{y}-L}{U-L}\right)^r, \quad L \leq \hat{y} \leq U \quad (2)$$

where r is a user-specified parameter ($r > 0$). In the LTB case one assumes that it is possible to establish a finite target U such that $d = 1$ for $\hat{y} \geq U$ and $d = 0$ for $\hat{y} \leq L$.

C. Smaller-The-Best (STB): The value of the estimated response is expected to be smaller than an upper bound ($\hat{y} < U$). For this response type, the individual desirability function is defined as

$$d = \left(\frac{\hat{y}-U}{L-U}\right)^r, \quad L \leq \hat{y} \leq U \quad (3)$$

under the assumption that it is possible to establish a finite target L such that $d = 1$ for $\hat{y} > L$, and $d = 0$ for $\hat{y} \geq U$.

The values assigned to s , t and r allow changing the shape of d . A large value for s , t or r implies that the individual desirability value is very low unless the response gets very close to its target value. This means that the higher their values are, the greater the importance of response values being closer to the respective target will be. Note that the target value for NTB-, LTB- and STB-type responses are T , U and L , respectively. A small value implies that the individual desirability value is large even when the response is distant from its target value. Illustrative plots for different weights and discussion on their choice are presented in Ref. [30]. Derringer [17] aggregated the individual desirability functions into a weighted geometric mean defined as

$$D = \left((d_1)^{w_1} (d_2)^{w_2} \dots (d_p)^{w_p}\right)^{\frac{1}{\sum w_i}} \quad (4)$$

where d_i is the individual desirability function of the i -th response ($i = 1, \dots, p$), and w_i are user-specified parameters that make possible the assignment of priorities to d_i . The objective is to maximize D , which will be equal to one ($D = 1$) when all responses are on-target ($d_i = 1$), and equal to zero ($D = 0$) when at least one response is outside of the specification limits ($d_i = 0$, for any i). Otherwise, the multiplicative combination of d_i yields a value less than or equal to the lowest value of d_i .

2.1.2. Castillo et al. [24]

The desirability function proposed in Ref. [16] for NTB-type response is non-differentiable at the target points. So, Castillo et al. proposed a piecewise continuous desirability function to correct the non-differentiable points using a local polynomial approximation (a quartic polynomial) so that more efficient gradient-based optimization methods can be used. They defined the individual desirability functions as

$$d = \begin{cases} a_0 + b_0 \hat{y}, & L < \hat{y} \leq T - \delta \\ f(\hat{y}), & T - \delta < \hat{y} \leq T + \delta \\ a_1 + b_1 \hat{y}, & T + \delta < \hat{y} \leq U \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

where $f(\hat{y}) = A + B\hat{y} + C\hat{y}^2 + D\hat{y}^3 + E\hat{y}^4$, $\delta = (U-L)/50$ and $A, B, C, D, E, a_0, b_0, a_1$ and b_1 are constants. The composite function is given by Eq. (4).

The major drawback of this method is that it cannot be applied in problems with STB- and LTB-type responses, which are usual in practice.

2.1.3. Wu and Hamada [25]

Wu and Hamada suggested individual desirability functions in the form of exponential functions, which for NTB-type response are defined as

$$d = \begin{cases} \exp(-k_1|\hat{y}-T|^s) & , \quad -\infty < \hat{y} \leq T \\ \exp(-k_2|\hat{y}-T|^t) & , \quad T \leq \hat{y} < \infty \end{cases} \quad (6)$$

where k_1 and k_2 are scale constants and s and t are shape parameters.

For STB- and LTB-type responses the individual desirability functions are defined by Eqs. (7–8), respectively.

$$d = \exp(-k_3|\hat{y}-\phi|^r) \quad , \quad \phi \leq \hat{y} < \infty \quad (7)$$

$$d = \begin{cases} 1 - \frac{\exp(-k_3\hat{y})}{\exp(-k_3L^r)} & , \quad L \leq \hat{y} < \infty \\ 0 & , \quad \hat{y} < L \end{cases} \quad (8)$$

where k_3 is a scale constant, ϕ represents the target value for STB-type responses, and r is a shape parameter.

As optimization criterion, the authors proposed to maximize the global desirability function D

$$D = (d_1)^{w_1} (d_2)^{w_2} \dots (d_p)^{w_p} \quad (9)$$

with $0 < w_i < 1$ and $w_1 + w_2 + \dots + w_p = 1$. Note that this method requires a number of user-specified parameters larger than the so-called Derringer and Suich's method, which is not an appealing feature.

2.1.4. Kim and Lin [26]

These authors proposed to maximize the overall minimal level of satisfaction with respect to all the responses. To this purpose an exponential desirability functional form that can generate a rich variety of shapes by adjusting its parameters is suggested. It does not require any assumptions regarding the form or degree of the estimated response models, is robust to the potential dependences between response variables, and considers the relative priority of responses. The individual desirabilities are defined as

$$d = \begin{cases} \frac{\exp(k) - \exp(k|z|)}{\exp(k) - 1} & , \quad t \neq 0 \\ 1 - |z| & , \quad t = 0 \end{cases} \quad (10)$$

where k is a constant ($-\infty < k < \infty$) and z is a standardized parameter representing the distance of the estimated response from its target in units of the maximum allowable deviation. This parameter depends on the response type and is defined as

$$z = \begin{cases} \left(\frac{\hat{y}-T}{y^{\max}-y^{\min}} \right) & \text{for NTB} \\ \left(\frac{\hat{y}-y^{\min}}{y^{\max}-y^{\min}} \right) & \text{for STB with } y^{\min} \leq \hat{y} \leq y^{\max} \\ \left(\frac{y^{\max}-\hat{y}}{y^{\max}-y^{\min}} \right) & \text{for LTB} \end{cases} \quad (11)$$

where $y^{\max}$ and $y^{\min}$ represent the maximum and minimum values of the estimated response, respectively. Eq. (11) ranges between -1 and 1 for NTB-type responses and between 0 and 1 otherwise. This modeling approach can take into account the level of responses predictive ability through the adjustment of the parameter k . In this case, it will be defined as

$$k' = k + (1-R^2)(k^{\max}-k) \quad (12)$$

where $k^{\max}$ is a sufficient large value of k such that $d(z)$ with $k^{\max}$ is extremely concave and thus has virtually no effect in the optimization process. R^2 is the coefficient of determination.

The composite function to maximize the overall minimal level of satisfaction with respect to all the responses is

$$D = \max \min(d_1, d_2, \dots, d_p) \quad (13)$$

This *maximin* scheme has a serious disadvantage; it only considers the response with the lowest degree of satisfaction so the degrees of satisfaction associated with all the other responses are ignored, which may lead to an unreasonable decision in some cases. As an example, the approach would suggest a solution with satisfaction levels equal to (0.6, 0.6, 0.6, 0.6) instead of (0.9, 0.9, 0.9, 0.58).

2.1.5. Ch'ng et al. [27]

Minimization of an arithmetic mean is the proposal of Ch'ng et al. for finding the optimal variable settings under the assumption of normality and homogeneity of error variances. They proposed individual desirability functions of the form

$$d = \frac{2\hat{y} - (U+L)}{U-L} + 1 = \frac{2}{U-L}\hat{y} + \frac{-2L}{U-L} = m\hat{y} + c \quad (14)$$

with $0 \leq d \leq 2$. As optimization criterion, the authors minimize a composite function defined as

$$D = (\sum_{i=1}^p w_i |d_i - d_i(T_i)|) / p \quad (15)$$

where $d_i(T_i)$ represents the value of the i -th individual desirability function for $\hat{y}$ equal to the response target ($\hat{y} = T_i$), w_i represents the weight (importance or priority) assigned to response i , p is the number of responses, and $\sum_{i=1}^p w_i = 1$.

Distinctive features of this method are the following: individual desirabilities do not have any breakpoint; can be easily understood and implemented; and subjective information required from the analyst is minimal. Moreover, the number of user-specified parameters (weights) is just equal to the number of responses, which is an appealing feature. The theoretical drawback is that it does not consider the variances and correlations existing among the responses.

2.1.6. Wu [28]

In contrast to previous authors, Wu [28] proposed individual and composite desirability functions that consider the variances and correlations among the responses. The individual desirability function of the i -th response is defined as

$$d_{ii} = \exp(-F_i) \quad (16)$$

where $F_i = \left[k_i \left((\hat{y}_i - T_i)^2 + \hat{\sigma}_i^2 \right) \right]$, k_i is the loss coefficient and $\hat{\sigma}_i^2$ is the estimated variance model for the i -th response (y_i). The correlated desirability function (d_{ij} , $i \neq j$) is defined as

$$d_{ij} = \begin{cases} 1 & , \quad F_{ij} < 0 \\ \exp(-F_{ij}) & , \quad F_{ij} \geq 0 \end{cases} \quad (17)$$

where $F_{ij} = k_{ij}(\hat{\rho}_{ij}\hat{\sigma}_i\hat{\sigma}_j + (\hat{y}_i - T_i)(\hat{y}_j - T_j))$, k_{ij} is the correlated loss coefficient and ρ_{ij} is the correlation coefficient of responses y_i and y_j . As optimization criterion they maximize the global desirability shown in Eq. (18).

$$D^* = \left[\prod_{i=1}^p \prod_{j=i+1}^p d_{ii}d_{ij} \right]^{1/p} \quad (18)$$

The major drawback of this method is the number and computation of the user-specified parameters (denoted by k_i and k_{ij} , which represent the loss coefficient of the i -th response and correlated loss coefficient of responses i and j) to which the analyst needs to assign values. These parameters play a decisive role in the optimization process. However, its number is always larger than the number of responses and the information necessary to compute them is not easy to define or readily available. For details on the computation of the loss and correlated loss coefficients the reader is referred to Wu and Chyu [31].

2.2. More sophisticated approaches

This subsection provides a brief summary of some desirability-based methods that require from the user a higher level of expertise on mathematics and statistics, which may limit their use in practice. It includes the methods proposed in Refs. [32–37].

2.2.1. Ribardo and Allen [32]

Arguing that there is often an interaction between the mean and standard deviation with regard to their effects on process yield and therefore on profitability, Ribardo and Allen [32] proposed a desirability-based method that explicitly accounts for the combined effect of the mean and dispersion of responses to determine the steepness of the desirability functions and to regulate how far inside the limits the mean should be positioned. These authors employ a criterion that is based on estimates of the process yield, namely on the predicted fraction of conforming parts. However, to preserve the interpretation of the desirability functions as the yield associated to responses, the individual and composite desirabilities become hard to interpret and use.

2.2.2. Ortiz et al. [33]

These authors show that if multiresponse optimization problems grow even moderately in either the number of factors or the number of responses, conventional optimization algorithms may fail to find the optimal variable settings. Thus they proposed an approach that uses a genetic algorithm in conjunction with an unconstrained desirability function, enabling the algorithm to differentiate between far-from-feasible and nearly feasible solutions. Delineation is accomplished by incorporating a penalty function proportional to the magnitude of the constraint violation into the overall desirability value of each design space point.

2.2.3. Pasandideh and Niaki [34]

To integrate the desirability function and simulation approach with a genetic algorithm was the proposal of Pasandideh and Niaki, which consists in: modeling the multiresponse problem by desirability function models; generating the required input data from a simulated system; optimizing the composite function through a genetic algorithm. Four genetic algorithms that are different in structure, especially in controlling the stochastic nature of the problem, are compared and they concluded that no statistical difference exists between them.

2.2.4. Lee and Kim [35]

A new type of desirability function is proposed by these authors. The so-called expected desirability function is defined as the average of the

conventional desirability values based on the probability distribution of the predicted response variable. It considers the dispersion effects and location effects of the responses, does not require a separate desirability function for the standard deviation response, and can be used for (moderately) asymmetric NTB-type responses.

2.2.5. Das and Sengupta [36]

In situations where customer perceptions are used as metrics/measurements of critical performance characteristics of the process or product, the scales usually take values between $-k$ and $+k$ for some characteristic(s) and so the individual and composite desirability functions above reviewed are not appropriate. As alternative, Das and Sengupta proposed desirability functions that either do not reflect or avoid using the negative and zero values, because the negative exponential transformation used to modify Gatzka's desirability function can accommodate such values.

2.2.6. He et al. [37]

According to these authors, most of the recent work in industrial optimization does not deal with the quality of predictions and most practitioners follow up optimization calculations without considering this responses property. So He et al. [37] proposed a creative procedure to define individual desirability functions that account for quality of predictions (QoP) and a composite desirability function that makes balance between QoP and bias.

3. Proposed approach

Progress in research on desirability functions has been focused on flexibility, minimization of the impact in decision-making of certain types of dependencies caused by data limitations, consideration of response's variance, and use of genetic algorithms. Approaches that require minimum subjective information from the user and are easy to understand and implement are especially appealing to practitioners and useful if they provide compromise solutions with desired response's properties. Therefore, we evaluated the ability of two less sophisticated methods along with optimization performance measures to yield compromise solutions with respect to desired response's properties under adverse variance conditions. The goal is to identify the best performer, if at all possible. The methods under evaluation are those of Ch'ng et al. [27] and Wu [28].

Ch'ng et al.'s method is adapted to accommodate LTB- and STB-type responses by using $d_i(U_i)$ and $d_i(L_i)$ in Eq. (15) instead of $d_i(T_i)$, under the assumption that it is possible to establish specification limits U and L to those responses. If these limits are not readily available, it is reasonable to use the maximum and minimum values of the respective response model.

Ch'ng et al.'s method is not designed to take into account the responses variance levels and correlation information, so it may lead to less reasonable solutions when the response's correlation and variance due to low QoP and uncontrollable factors are important issues in practice. To minimize this problem, the estimated variance models in addition to estimated mean models for the responses are considered in Eq. (15) and then optimization measures are used to help the analyst evaluate and select a compromise solution.

Seemingly Unrelated Regression (SUR) method is used to estimate the regression models. This method can lead to a more precise estimate of the control factors setting, in particular if responses are correlated [38,39]. This is corroborated in Ref. [10] where it is showed that SUR modeling can make a noticeable difference.

3.1. Optimization performance measures

The solutions obtained from the methods proposed in Refs. [27,28] are assessed in terms of bias, quality of predictions, and robustness through the optimization measures proposed in Ref. [40].

The bias is assessed with an optimization measure that considers the response types, response's specification limits and response deviations from target. This measure, named cumulative bias (B_{cum}), is defined as

$$B_{cum} = \sum_{i=1}^p W_i |\hat{y}_i^* - \theta_i| \quad (19)$$

where $\hat{y}_i^*$ represents the value of the i -th estimated response at “optimal” variables setting, θ_i is the response target value and W_i is a parameter that takes into account the specification limits and response type of the i -th response. This parameter makes responses dimensionless and is defined as follows: $W = 1/(U-L)$ for STB and LTB-type responses; $W = 2/(U-L)$ for NTB-type response.

The cumulative bias gives an overall result of the optimization process instead of focusing on the value of a single response, which prevents the analyst from taking unreasonable decisions in some cases, such as what Kim and Lin [26] showed. Nevertheless, the bias of each response is also available from Eq. (19).

To assess method's solutions in terms of quality of predictions is used a measure defined as follows:

$$QoP = \text{trace} \left[\phi \left(x_j^T [X^T Q^{-1} X]^{-1} x_j \right) \right] = \text{trace} [\phi \Sigma_{y^*}(x)] \quad (20)$$

where x_j is the subset of independent variables consisting of the $K \times 1$ vector of regressors for the i -th response with N observations on K_i regressors for response, X is an $N \times K$ block diagonal matrix and $Q = \sum \otimes I_N$. An estimate of Σ is $\hat{\sigma}_{ij} = \hat{e}_i^T \hat{e}_j / N$, where $\hat{e}$ is the residual vector from the OLS (Ordinary Least Squares) estimation of the i -th response; I_N is an identity matrix and $\otimes$ represents the Kronecker product. The matrix $\Sigma_{y^*}(x)$ is calculated for x_j equal to optimal settings and to make it dimensionless this matrix is multiplied by matrix ϕ , whose diagonal and non-diagonal elements are $\phi_{ii} = 1/(U_i - L_i)^2$ and $\phi_{ir} = 1/(U_i - L_i)(U_r - L_r)$ for $i \neq r$, respectively.

This measure is defined under the assumption that Seemingly Unrelated Regression (SUR) method is used to estimate the regression models. If the Ordinary Least Squares is used, the reader is referred to Ref. [41] where variants of Eq. (20) are presented for the cases of regression models with equal and different terms.

The robustness is assessed by

$$\text{Rob} = \text{trace} [\phi \hat{\sigma}_{y^*}] \quad (21)$$

where $\hat{\sigma}_{y^*}$ represents the variance–covariance matrix of the “true” responses at “optimal” variable setting. Note that replications of the experimental runs are required to calculate the variance–covariance matrix and matrix ϕ only considers the specification limits assigned to responses variance. Although the replicates increase the time and cost of experimentation, they may provide significant improvements in robustness that overbalance or at least compensate the time and cost spent.

The optimization measures defined in Eqs. (19–21) do not exclude other measures from also being used. As concerns their results, the lower the values are, the better the compromise solution will be. In practice, the presented measures take values greater than or equal to zero, being the most favorable the zero value.

3.2. Select a compromise solution

Choosing the best compromise solution from a set of competing alternatives in multiresponse problems is not a trivial task. In the RSM framework the general procedure consists of selecting that solution based on the output (maximum or minimum value) of the objective function used. However, the widely used global desirability and expected loss functions give inconsistent and incomplete information to the analyst in terms of the merit of the final solution, namely with

respect to desired response's properties. If the analyst only focuses on the result of the composite function used for making decisions he/she may ignore a solution of interest, because that value depends on the weights assigned to responses and a satisfactory solution may have less favorable desirability values. Such as what Costa et al. [40] showed, the composite functions yield different results in cases where the solutions are equal or have slight changes in the response values due to the different weights or priorities assigned to responses. This is a relevant shortcoming that the analyst must be aware of and may overtake by using the aforementioned optimization measures, because they do not depend on the priorities assigned to responses. However, another problem arises in choosing a compromise solution because some of them may be better with respect to bias but worse with respect to variance (QoP and/or Rob). In these instances, unless technical or economical reasons arise to justify another option, the cumulative value of B_{cum} , QoP and Rob measures, denoted by Cum, is considered in this article to select the compromise solution.

4. Examples

Myers and Montgomery [42] reported an experimental study where the objective was to maximize the conversion of a polymer and minimize the thermal activity by setting the variables reaction time (x_1), reaction temperature (x_2), and amount of catalyst (x_3) at appropriate levels. The ranges for the conversion (y_1) and thermal activity (y_2) responses are [80, 100] and [55, 60], respectively. Assuming that y_1 is a LTB-type response, its target value is set equal to 100; y_2 is a NTB-type response and its target is set equal to 57.5.

This example supports the design of three case studies to evaluate and compare the performance of several analysis methods, which were coded in Matlab®. Adverse variance conditions are simulated for each case study as follows: Case 1 – low quality of predictions; Case 2 – high quality of predictions and unequal robustness; and Case 3 – low quality of predictions and unequal robustness.

Cases 2–3 assume unequal robustness, which implies that, in addition to estimated models for the mean response ($\hat{\mu}_1$ and $\hat{\mu}_2$, respectively), replicates are needed to estimate the variance models for polymer conversion ($\hat{\sigma}_1$) and thermal activity ($\hat{\sigma}_2$). Based on the central composite design with six axial and six center points ran by Myers and Montgomery (see Table 1), five replicates ($i = 1, 2, \dots, 5$) are generated, like in Ref. [43], using the model

$$\begin{pmatrix} y_{1i}(x) \\ y_{2i}(x) \end{pmatrix} = \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix} + e_i \quad (22)$$

where $y_1(x)$ and $y_2(x)$ represent the response values presented in Ref. [42], and the error term e_i is defined as follows:

$$e_i \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11}(x) & \sigma_{12}(x) \\ \sigma_{21}(x) & \sigma_{22}(x) \end{pmatrix} \right] \quad (23)$$

with $i = 1, 2, \dots, 5$, $\sigma_{11}(x) = \exp(3 - x_1^2 - 3x_3^2)$, $\sigma_{22}(x) = \exp(2 - 2x_1^2 - x_3^2)$, and $\sigma_{12}(x) = \sigma_{21}(x) = 0.03\sigma_{11}(x)\sigma_{22}(x)$. These models make the responses variance due to uncontrollable variables larger for operation conditions near the origin, that is, the closer the operating conditions for x_1 and x_3 are to the origin, the worst the responses robustness will be. The responses models are fitted using SUR method and the range for the input variables is $-1.682 \leq x_j \leq 1.682$, $j = 1, 2, 3$.

Ch'ng et al.'s and Wu's methods are evaluated in three case studies and the feasibility of their solutions compared with those of methods designed to consider the quality of predictions and robustness. The method proposed by Vining [41], which allows specifying the directions of economic importance (weights or priorities to responses) while considering the variance–covariance structure of the estimated responses (quality of predictions) is used in Case 1. The method proposed by Pignatiello [44], which puts emphasis on improving

Table 1
Experimental design and simulation results.

x_1	x_2	x_3	Polymer conversion (y_1 -LTB)				Thermal activity (y_2 -NTB)			
−1	−1	−1	74.125	74.009	74.500	73.016	74.432	53.303	53.485	54.188
1	−1	−1	51.203	51.690	51.490	50.321	50.159	62.654	63.173	63.678
−1	1	−1	88.613	87.718	88.961	88.761	87.587	54.174	53.311	53.608
1	1	−1	69.542	69.520	69.448	70.643	70.083	61.789	63.312	62.873
−1	−1	1	70.456	71.564	71.063	69.522	70.094	56.746	58.295	57.090
1	−1	1	89.714	90.387	89.818	90.689	90.366	68.327	68.373	67.927
−1	1	1	66.475	65.967	66.505	65.991	66.539	59.721	59.248	60.622
1	1	1	96.119	97.510	96.819	96.950	96.318	67.419	67.591	67.333
−1.682	0	0	77.632	78.137	76.570	74.430	76.951	59.055	59.087	58.867
1.682	0	0	78.549	76.373	78.309	78.017	78.619	66.128	65.928	66.064
0	−1.682	0	85.650	84.369	86.418	84.901	85.768	62.055	59.375	62.376
0	1.682	0	100.255	97.683	94.842	102.277	93.825	58.975	63.333	60.095
0	0	−1.682	55.136	54.895	55.020	55.102	54.936	57.114	57.532	57.368
0	0	1.682	81.008	80.989	80.983	80.916	81.090	63.455	62.834	63.315
0	0	0	86.978	78.373	70.546	85.439	90.903	62.974	58.608	56.190
0	0	0	82.010	84.358	78.765	72.300	79.464	61.435	56.218	57.124
0	0	0	74.151	73.247	77.608	81.747	77.671	63.677	64.414	60.810
0	0	0	91.657	80.442	83.103	80.349	86.053	61.235	61.968	61.927
0	0	0	72.136	81.935	80.230	78.625	82.035	63.178	60.127	60.793
0	0	0	88.702	90.861	86.640	85.786	90.213	61.782	56.693	55.698
										58.563
										62.406

robustness, by considering the variance–covariance structure of the “true” responses and minimizing bias, is used in Case 2. In Case 3, the method proposed by Ko et al. [43] is used because it is designed to consider both quality of predictions and robustness in addition to bias. This method is built on the methods presented in Refs. [41] and [44], and yields better results than these two methods when both quality of predictions and robustness are important issues in practice [43].

The expected loss function proposed in Ko et al. [43] includes the expected losses of Vining [41] and Pignatiello [44] as special cases and is defined as follows:

$$E[L(y(x), \theta)] = (\hat{y}(x) - \theta)^T C (\hat{y}(x) - \theta) + \text{trace}[C \Sigma_{\hat{y}(x)}] + \text{trace}[C \Sigma_y(x)] \quad (24)$$

where the first term is an imposed penalty due to bias, the second term is the imposed penalty due to poor quality of predictions, and the third term is the imposed penalty due to poor robustness. The first and second terms constitute the Vining's expected loss function; the first and third terms constitute the Pignatiello's expected loss function. The variance–covariance matrix of the predicted responses at x is denoted by $\Sigma_{\hat{y}(x)}$ while $\Sigma_y(x)$ is the variance–covariance matrix of the responses at x . C is a cost matrix and their elements represent the costs of non-optimal variable setting. If C is a diagonal matrix then each element represents the relative importance assigned to the respective response, that is, the penalty (cost) incurred for each unit

of response value deviates from its optimum. If C is a non-diagonal matrix, the off-diagonal elements represent additional costs incurred when pairs of responses are off-target simultaneously.

4.1. Case 1 – low quality of predictions

Some insignificant regressors (x_2 , x_2^2 , x_3^2 , x_1x_2 , x_1x_3 , x_2x_3) were added to the model fitted to thermal activity so that the predicted response has a lower quality of predictions when x_i values are farther from the origin. The estimated models are:

$$\hat{\mu}_1 = 81.0943 + 1.0290x_1 + 4.0426x_2 + 6.2060x_3 - 1.8377x_1^2 + 2.9455x_2^2 - 5.2036x_3^2 + 2.1250x_1x_2 + 11.3750x_1x_3 - 3.8750x_2x_3$$

$$\hat{\mu}_2 = 59.8505 + 3.5855x_1 + 0.2547x_2 + 2.2312x_3 + 0.8360x_1^2 + 0.0742x_2^2 + 0.0565x_3^2 - 0.3875x_1x_2 - 0.0375x_1x_3 + 0.3125x_2x_3$$

According to Vining [41] the objective of this problem is to achieve a value for $\hat{\mu}_1$ as large as possible, but the product must be sellable. This occurs when the value of thermal activity is inside the specification limits, although the ideal situation is to have $\hat{\mu}_2$ on target. To bring thermal activity closer to the target, Vining assigned a larger weight to $\hat{\mu}_2$ than to $\hat{\mu}_1$. At this condition his method yielded a

Table 2
Results – Case 1 (LQP).

Solution	Weights	x_i	$\hat{\mu}_1, \hat{\mu}_2$	Result	B_{cum}	QoP	Cum
<i>Vining</i>							
V11	(1.0, 0.00, 50.0)	(−0.4898, 1.682, −0.5778)	95.19, 57.67	$E_{Loss} = 98.41$	0.548	0.066	0.614
V21	(0.1, 0.025, 0.50)	(−0.3781, 1.682, −0.4871)	95.24, 58.16	$E_{Loss} = 3.73$	0.742	0.059	0.801
V31	(0.01, 0.04, 0.30)	(−0.3376, 1.607, −0.4243)	94.18, 58.39	$E_{Loss} = 0.50$	0.938	0.049	0.987
V41	(0.02, 0.01, 0.80)	(−0.4776, 1.682, −0.5654)	95.19, 57.73	$E_{Loss} = 1.68$	0.572	0.065	0.637
V51	(0.01, 0.04, 0.80)	(−0.4299, 1.526, −0.4871)	92.95, 57.97	$E_{Loss} = 1.17$	0.892	0.046	0.938
<i>Ch'ng et al.</i>							
C11	(0.22, 0.78)	(−0.5434, 1.682, −0.5984)	95.21, 57.50	$D = 0.053$	0.479	0.068	0.547
C21	(0.89, 0.11)	(0.0221, 1.682, −0.2019)	96.13, 60.00	$D = 0.227$	1.387	0.049	1.436
C31	(0.86, 0.14)	(−1.682, 1.682, −1.058)	98.03, 55.00	$D = 0.155$	1.197	0.243	1.440
C41	(0.01, 0.99)	(−0.6576, 1.682, −0.5163)	95.04, 57.50	$D = 0.002$	0.496	0.070	0.566
<i>Wu</i>							
W11	(0.01, 0.01, 0.30)	(−0.5436, 1.682, −0.5982)	95.21, 57.50	$D^* = 0.891$	0.479	0.069	0.548
W21	(0.08, 0.01, 0.40)	(−0.2125, 0.346, 0.3163)	82.64, 60.00	$D^* \approx 0.000$	2.736	0.013	2.749

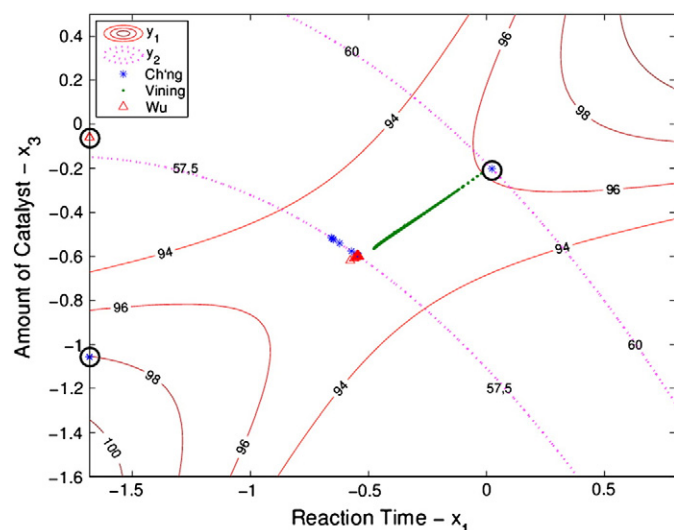


Fig. 1. LQP.

solution denoted by V11 in Table 2. However, a solution with B_{cum} , QoP and Cum values in close agreement with V11, denoted by V41, was achieved without the need of assigning a large weight to $\hat{\mu}_2$, which justifies the difference in their expected loss (E_{Loss}) values.

The set of solutions yielded by Vining's method is presented in Fig. 1. None of the solutions have $\hat{\mu}_2$ on target and those with better $\hat{\mu}_2$ values have worse $\hat{\mu}_1$ values and vice-versa. Except V11, all the other solutions have $\hat{\mu}_2 \geq 57.73$ with $93.77 < \hat{\mu}_1 < 95.99$.

These results were achieved by testing the loss coefficient values as follows: $c_1 = \{0.01, 0.02, \dots, 0.2\}$; $c_{12} = \{0.01, 0.02, 0.03, 0.04\}$; $c_2 = \{0.3, 0.4, \dots, 0.8\}$. This set of values include the weights used by Vining [41], $c_1 = 0.1$, $c_{12} = 0.025$ and $c_2 = 0.5$, which lead to the solution V21. This solution gives worse QoP and E_{Loss} values than V31 and V51, which are the solutions with the lowest E_{Loss} and QoP values, respectively. Moreover, V21 gives worse E_{Loss} , B_{cum} and Cum values than V41. These situations, like that one highlighted at the end of the first paragraph, points out how critical is to set values to loss coefficients.

In the Ch'ng et al.'s method, the change or increment for the weights was set equal to 0.01 units, under the constraint that the sum of e_i equals one. Most solutions (76 out of 99) have the same x_i values. This compromise solution, denoted by C11 in Table 2, gives the lowest Cum value and is characterized by having $\hat{\mu}_2$ on target. In the other solutions, a degradation of $\hat{\mu}_2$ value occurs, that is, $\hat{\mu}_2$ is moved to the specification limit, while $\hat{\mu}_1$ is closer to the target. Examples are the solutions C21 and C31, which are displayed in Fig. 1

rounded by a circle. Solution C21 gives the lowest QoP value among the solutions yielded by Ch'ng et al.'s method, and its value is equal to that of V31.

The solution C41 gives the lowest composite function (D) value, although is slightly worse than C11 with respect to Cum. Note that C11 gives a better Cum value than all the solutions yielded by Vining's method. The solutions of Ch'ng et al.'s method are displayed in Fig. 1 as asterisks, but some of them, like the C11, are not visible because they are hidden by Wu's method results, which are displayed as triangles. In fact, some solutions (146 out of 480) yielded by Wu's method, where the response weights were set equal to those used in the Vining's method, are identical or in close agreement with C11. These solutions give the largest desirability value and present a Cum value lower than all the solutions yielded by Vining's method. The remaining solutions have $\hat{\mu}_1$ and $\hat{\mu}_2$ close or equal to the upper specification limit. Examples are the solution rounded by a circle and the solution denoted by W21, which give an unacceptable D value. Solutions like W21 (318 out of 480, with $2.52 < Cum < 2.84$) are not shown in Fig. 1, because only solutions with $x_2 = 1.682$ are displayed.

4.2. Case 2 – high quality of predictions and unequal robustness

Besides the two estimated models for the mean response ($\hat{\mu}_i$, $i = 1, 2$), two models are fitted for standard deviation ($\hat{\sigma}_i$) so that responses robustness is considered. The models are as follows:

$$\hat{\mu}_1 = 81.7863 + 0.8234x_1 + 4.1361x_2 + 6.2127x_3 + 1.9954x_1^2 + 2.8346x_2^2 - 5.5585x_3^2 + 1.6981x_1x_2 + 11.4114x_1x_3 - 3.8694x_2x_3$$

$$\hat{\sigma}_1 = 4.4903 - 1.2366x_1^2 - 0.8904x_2^2 - 1.6466x_3^2$$

$$\hat{\mu}_2 = 60.7977 + 3.5894x_1 + 2.1794x_3$$

$$\hat{\sigma}_2 = 2.0583 - 0.6512x_1^2 - 0.7015x_3^2$$

The specification limits and targets for $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are: $\hat{\sigma}_1 \leq 10$ and $\hat{\sigma}_2 \leq 5$ with targets equal to zero. The response weights and increments are the same as in Case 1.

Pignatiello solutions with the lowest E_{Loss} , Cum and Rob values are denoted in Table 3 by P12, P22 and P32, respectively. As regards the Cum value, the best Pignatiello solution (P22), delimited by a hexagon in Figs. 2 and 3, is in close agreement with the best ones yielded by Ch'ng et al.'s method (C12 and C22). These solutions also have the lowest B_{cum} value. The solution P32 has a B_{cum} value slightly higher than P22, whereas its Rob value is lower, because x_1 and x_3 values are

Table 3
Results – Case 2 (HQPUR).

Solution	Weights	x_i	$\hat{\mu}_1, \hat{\mu}_2$	$\hat{\sigma}_1, \hat{\sigma}_2$	Result	B_{cum}	QoP	Rob	Cum
<i>Pignatiello</i>									
P12	(0.01, 0.04, 0.30)	(−0.3886, 1.682, −0.672)	95.70, 57.94	1.04, 1.64	$E_{Loss} = 0.599$	1.471	0.03	0.078	1.579
P22	(0.01, 0.02, 0.80)	(−0.4058, 1.682, −0.8589)	95.07, 57.47	0.55, 1.43	$E_{Loss} = 1.403$	1.189	0.03	0.064	1.283
P32	(0.18, 0.01, 0.30)	(−0.8529, 1.682, −0.8067)	96.65, 55.98	0.00, 1.13	$E_{Loss} = 3.381$	1.395	0.03	0.045	1.475
<i>Ch'ng et al.</i>									
C12	(0.1, 0.7, 0.1, 0.1)	(−0.3652, 1.682, −0.9117)	94.60, 57.50	0.44, 1.39	$D = 0.0296$	1.183	0.03	0.061	1.274
C22	(0.25, 0.25, 0.4, 0.1)	(−0.3307, 1.682, −0.9686)	94.05, 57.50	0.29, 1.33	$D = 0.0563$	1.184	0.03	0.057	1.271
C32	(0.1, 0.1, 0.1, 0.7)	(−0.8820, 1.125, −1.2076)	87.87, 55.00	0.00, 0.53	$D = 0.0923$	2.425	0.03	0.021	2.476
<i>Wu</i>									
W12	(0.01, 0.01, 0.3)	(−0.4720, 1.682, −0.9084)	95.16, 57.12	0.34, 1.33	$D^* = 0.66$	1.236	0.03	0.057	1.323
W22	(0.01, 0.04, 0.8)	(−0.3051, 1.682, −1.0617)	93.20, 57.39	0.00, 1.21	$D^* = 0.43$	1.207	0.03	0.048	1.289
W32	(0.01, 0.01, 0.5)	(−0.3800, 1.682, −1.0430)	93.86, 57.16	0.00, 1.20	$D^* = 0.56$	1.230	0.03	0.048	1.308

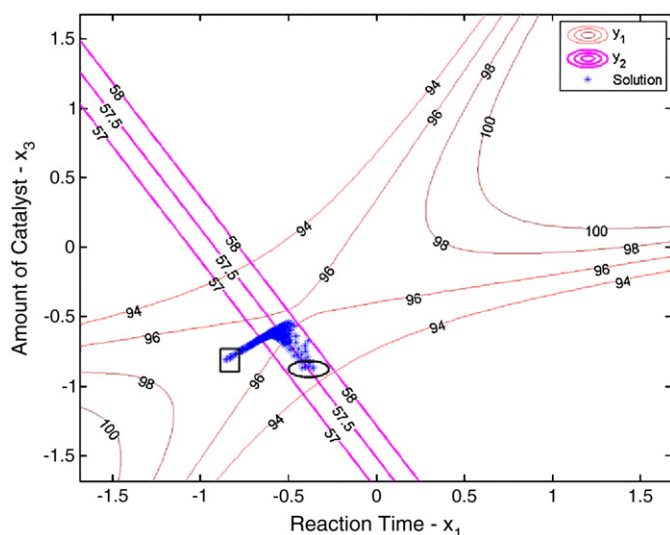


Fig. 2. Pignatiello: HQPUR—mean responses.

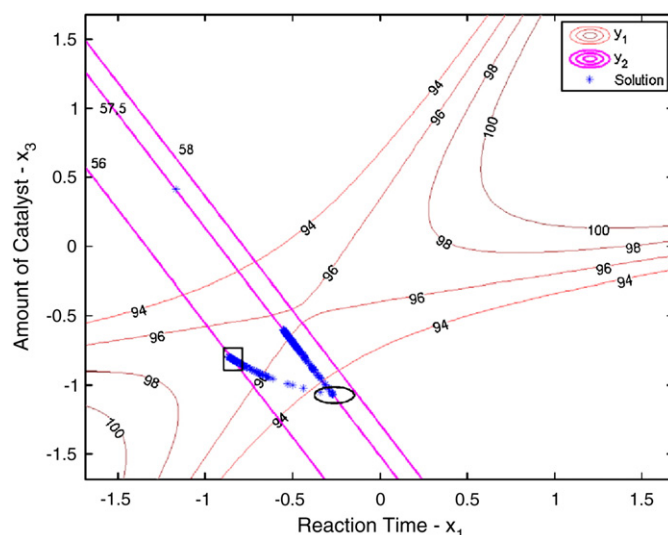


Fig. 4. Ch'ng et al.: HQPUR—mean responses.

farther from the origin. This solution and others with the lowest Rob value are delimited by a square in Figs. 2 and 3.

The spread out of the estimated mean ($\hat{\mu}_1$ and $\hat{\mu}_2$) and standard deviation ($\hat{\sigma}_1$ and $\hat{\sigma}_2$) values for Pignatiello's and Ch'ng et al.'s methods presented in Figs. 2–5 allows stating that they yielded distinct and equal or in close agreement solutions. An example of a distinct solution is C32 which has the lowest Rob value. However, C32 and similar ones with lower variance values (8 out of 424), which are not displayed in Fig. 2 because $x_2 \neq 1.682$, have $\hat{\mu}_1$ and $\hat{\mu}_2$ too far from target, leading to B_{cum} and Cum values higher than those of all the other solutions. The solutions with the lowest B_{cum} and Cum values yielded by Ch'ng et al.'s method (C12 and C22) are delimited by a hexagon in Figs. 4 and 5, while those ones with lower Rob values displayed in these figures are delimited by a square.

As regards the results from Wu's method, Table 3 shows that solution W22 gives B_{cum} , Rob and Cum values at least in close agreement with W12 and W32. Its Cum and B_{cum} values are marginally less favorable than those of C12, C22 and P22. This is not surprising, because Wu's method did not yield $\hat{\mu}_2$ values on-target (see Fig. 6). On the other hand, W22, delimited by an ellipse in Figs. 6 and 7, gives a slightly lower Rob value than C12, C22 and P22.

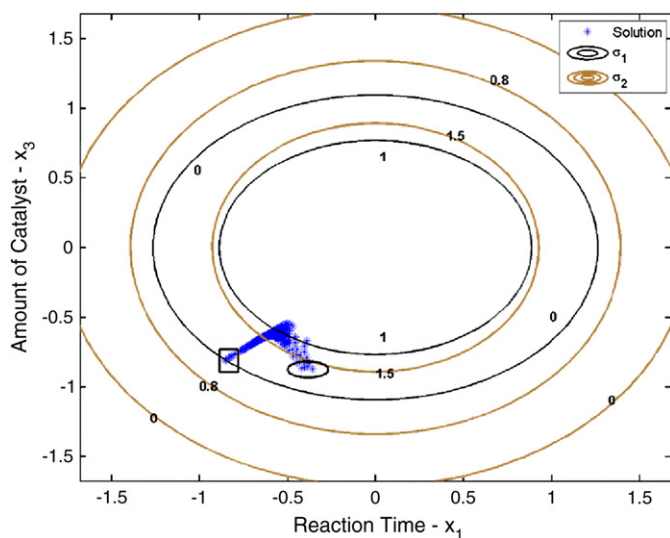


Fig. 3. Pignatiello: HQPUR—standard deviations.

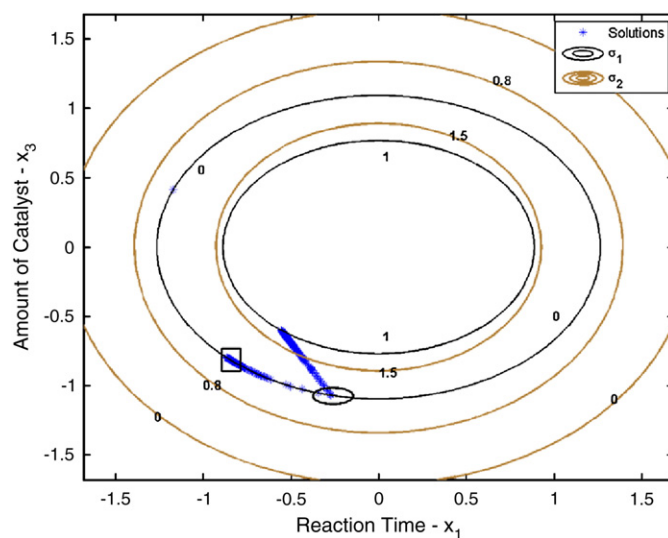


Fig. 5. Ch'ng et al.: HQPUR—standard deviations.

It is also important to highlight that Wu's method yields 328 out of 480 solutions with $x_2 \neq 1.682$ and high Cum values (Cum = 4.736), whereas Ch'ng et al.'s method yielded 8 out of 424 with $2.02 < \text{Cum} < 2.48$. These solutions are not displayed in Figs. 4–7, because these figures only show solutions with $x_2 = 1.682$.

4.3. Case 3 – low quality of predictions and unequal robustness

As in Case 1, some insignificant regressors are added to the model fitted to $\hat{\mu}_2$. This makes the QoP fluctuate more severely when x_i is farther from the origin, whereas the robustness for the conversion (y_1) and thermal activity (y_2) responses improves. The estimated models are as follows:

$$\begin{aligned}\hat{\mu}_1 &= 81.8380 + 0.8234x_1 + 4.0717x_2 + 6.2127x_3 - 2.1774x_1^2 \\ &\quad + 2.8731x_2^2 - 5.4906x_3^2 + 1.7698x_1x_2 + 11.4084x_1x_3 - 3.9496x_2x_3\end{aligned}$$

$$\hat{\sigma}_1 = 4.5033 - 1.2540x_1^2 - 0.8972x_2^2 - 1.6415x_3^2$$

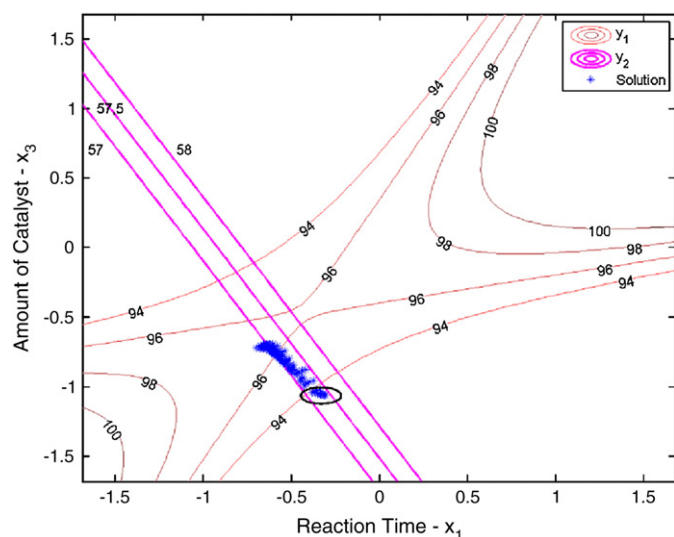


Fig. 6. Wu: HQPUR—mean responses.

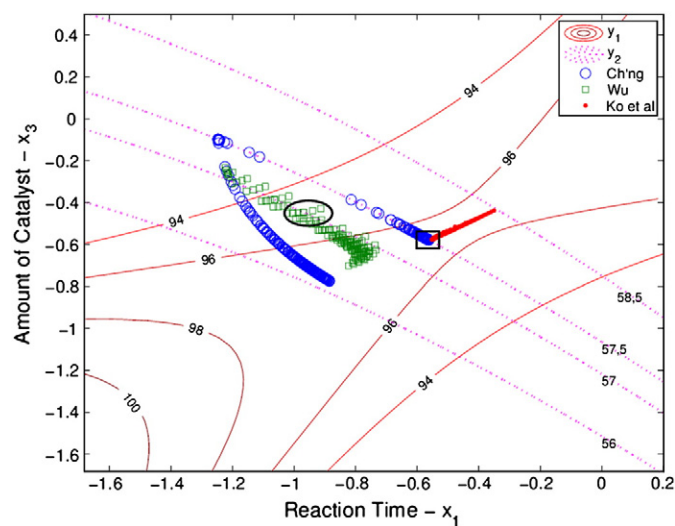


Fig. 8. LQPUR—mean responses.

$$\hat{\mu}_2 = 60.6935 + 3.5894x_1 + 0.3042x_2 + 2.1799x_3 + 0.5301x_1^2 - 0.1719x_2^2 - 0.2055x_3^2 - 0.3381x_1x_2 + 0.0124x_1x_3 + 0.3714x_2x_3$$

$$\hat{\sigma}_2 = 2.0935 - 0.7300x_1^2 - 0.6744x_3^2$$

The specification limits and targets for $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are the same as in Case 2. The response weights and increments are equal to those used in Cases 1 and 2.

Fig. 9 shows that Ko et al.'s method did not yield solutions with Rob values as low as those of the other two desirability-based methods. In fact, Ko et al.'s method yielded solutions, displayed as dots in Fig. 9, with $\hat{\sigma}_1 > 1.5$ and $\hat{\sigma}_2 \geq 1$, whereas the desirability-based methods yielded solutions with $\hat{\sigma}_1 < 1.5$ and $\hat{\sigma}_2 < 1$. At first glance, this seems a serious drawback, although it does not mean that Ko et al.'s method yields compromise solutions with worse response properties than the other two methods. Indeed, QoP and B_{cum} have to be also taken into account in the evaluation of compromise solution.

In terms of B_{cum} values, Fig. 8 shows that Ko et al.'s method can yield solutions with high $\hat{\mu}_1$ values and $\hat{\mu}_2$ values on-target, which are

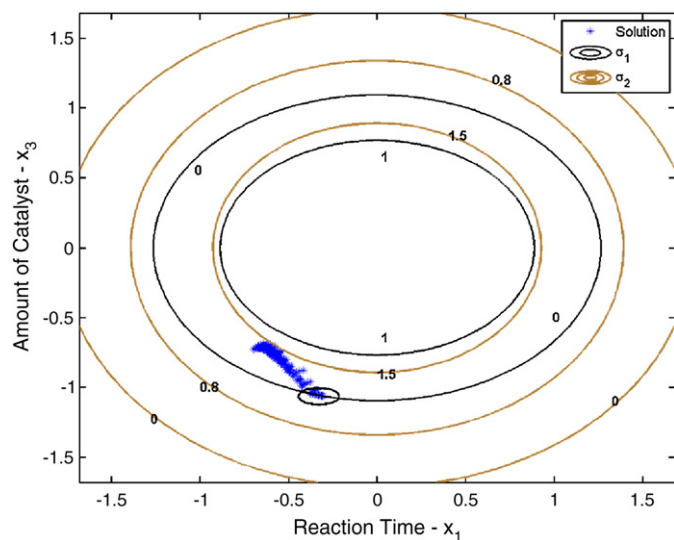


Fig. 7. Wu: HQPUR—standard deviations.

in close agreement with those yielded by Ch'ng et al.'s method that are delimited by a rectangle. These solutions lead to B_{cum} values slightly lower than those yielded by Wu's method, because all their solutions have $\hat{\mu}_2$ off-target, such as Fig. 8 displays.

As regards the variance (QoP plus Rob), Table 4 shows that the solutions with the lowest variance values are C13, W13 and K13, respectively. However, these solutions are not the best ones in terms of Cum. The solutions C13 and W13 are even the worst due to higher B_{cum} values.

Looking at Cum values, there is no significant differences among the best solutions of the three methods. Nevertheless, Ch'ng et al.'s method yields two solutions (C23 and C33), which present different D values due to the different weights assigned to responses, with B_{cum} and Cum values slightly better than all the other solutions displayed in Table 4. The solution C23 also presents the lowest D values, in contrast with Ko et al.'s solution with the lowest expected loss (K23), which is not the best in terms of Cum value. The solution K23 is worse than K13 with respect to B_{cum} , QoP, Rob and Cum values, and worse than K33 with respect to B_{cum} , variance (QoP plus Rob) and Cum values.

Wu's method yielded two solutions with similar Cum values, W23 and W33, which have variance slightly higher than W13. These solutions are marginally worse than C23, C33 and K33 with respect to Cum due to higher B_{cum} and QoP values, although they present slightly better Rob values. In particular, W23 gives the lowest Rob value among the methods solutions with the lowest Cum value presented in Table 4, and the highest D value among the solutions yielded by Wu's method, displayed as squares in Figs. 8 and 9. Solutions W23 and W33 are delimited by an ellipse in these figures.

Like Cases 1–2, Wu's method yielded many (327 out of 480) solutions with $x_2 \neq 1.682$ ($x_2 \approx 0$), leading to high Cum values (Cum = 4.756), whereas Ch'ng et al.'s and Ko et al.'s methods yielded 37 out of 424 with $1.51 < \text{Cum} < 2.44$ and 24 out of 480 with $1.55 < \text{Cum} < 1.98$, respectively. These solutions are not displayed in Figs. 8 and 9, because these figures only show solutions with $x_2 = 1.682$.

5. Discussion

To date, there is no effective procedure to guide the analyst in the selection of weights or priorities to responses. The optimization results can vary significantly if the weights change, and very little is usually known about how to choose their values. In practice, it is difficult, if at all possible, to determine beforehand the modifications required in the weights so as to produce a solution of interest, that is,

Table 4
Results – Case 3 (LQPUR).

Solution	Weights	x_i	$\hat{\mu}_1, \hat{\mu}_2$	$\hat{\sigma}_1, \hat{\sigma}_2$	Result	B_{cum}	QoP	Rob	Cum
Ko et al.									
K13	(0.01, 0.01, 0.80)	(−0.5934, 1.516, −0.5450)	93.70, 57.57	1.51, 1.64	$E_{Loss} = 2.738$	1.614	0.047	0.084	1.745
K23	(0.01, 0.04, 0.30)	(−0.4393, 1.614, −0.4708)	95.08, 58.15	1.56, 1.80	$E_{Loss} = 1.069$	1.785	0.050	0.091	1.926
K33	(0.20, 0.01, 0.80)	(−0.5516, 1.682, −0.5772)	96.11, 57.53	1.04, 1.65	$E_{Loss} = 7.119$	1.267	0.063	0.078	1.408
Ch'ng et al.									
C13	(0.15, 0.1, 0.2, 0.55)	(−1.5839, 0.895, −0.6235)	85.97, 55.32	0.00, 0.00	$D = 0.092$	2.275	0.070	0.000	2.345
C23	(0.1, 0.7, 0.1, 0.1)	(−0.6654, 1.682, −0.4980)	95.96, 57.50	1.00, 1.60	$D = 0.031$	1.246	0.065	0.076	1.387
C33	(0.25, 0.35, 0.1, 0.3)	(−0.6273, 1.682, −0.5275)	96.05, 57.50	1.01, 1.62	$D = 0.078$	1.246	0.064	0.077	1.387
Wu									
W13	(0.03, 0.01, 0.7)	(−0.8674, 0.007, −0.4517)	80.06, 56.96	3.22, 1.41	$D^* = 0.001$	3.418	0.014	0.094	3.526
W23	(0.01, 0.01, 0.3)	(−1.0039, 1.682, −0.4509)	95.05, 56.92	0.37, 1.22	$D^* = 0.662$	1.289	0.082	0.053	1.424
W33	(0.01, 0.04, 0.3)	(−0.9131, 1.682, −0.4288)	95.17, 57.17	0.62, 1.36	$D^* = 0.640$	1.284	0.075	0.062	1.421

so as to know which response(s) will change and which is the direction and magnitude of that change.

To code the optimization methods usually used in the RSM framework, Matlab® is an alternative that avoids testing various weights to each response using the usual trial-and-error procedure, which may be a source of frustration and significant inefficiency, particularly when either the number of responses or control factors is large. Nevertheless, the range and increment for the weights assigned to responses are required. Their specification impacts on computational time, so those values must be specified taking into account this inconvenience. The analyst must be also aware that if they are not properly defined, satisfactory compromise solutions will be lost.

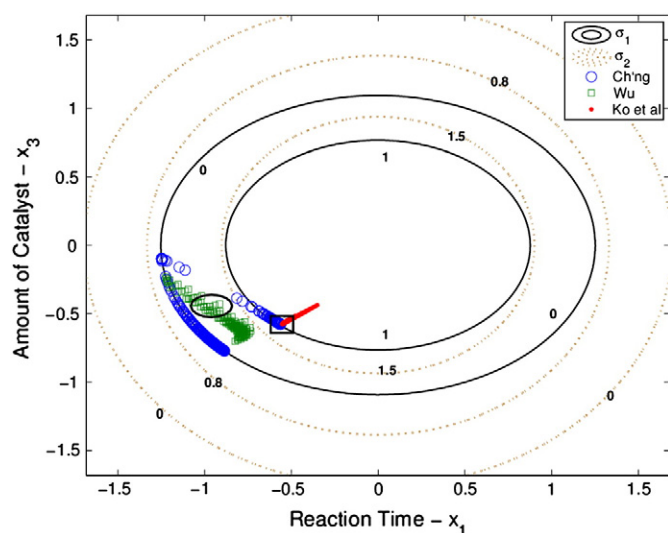
Regarding the results of Cases 1–3, there are no significant differences among Ch'ng et al.'s method and the others. All the methods yielded competitive compromise solutions for the three operating conditions tested here. Although the spread out of solutions may indicate that a particular method is more flexible in the sense that can yield solutions with either lower bias or variance values by changing the weights assigned to responses, the results show that a significant improvement in one property may degrade seriously other (s). This fact points out the intrinsic objective to multiresponse optimization problems: to identify a compromise solution among the generated responses. The methods evaluated here can do it, but Ch'ng et al.'s method may take some advantage over the others due to its major feature. This method requires a number of weights just equal to the number of responses, and to use it, the user only has to decide about the magnitude of the weights increment, because the sum of

the weights assigned to responses must be equal to one. In practice, less than 1 min is enough for running the routine in Matlab® for the problems usually developed in the RSM framework and illustrated in this article. However, it is not guaranteed that the compromise solution with the lowest composite function value leads to responses with desired properties, namely low bias, high quality of predictions and high robustness if the weight increments are not appropriate. The same occurs whenever loss function-based methods are used and the loss coefficients are not properly specified. To specify loss coefficients taking into account different scales, relative variabilities, and relative costs is difficult [41,43]. In practice, they may be not readily available or easily defined. Here, we tested loss coefficients around those used in Refs. [41] and [43]. Therefore, one cannot guarantee that the solutions presented here are the best ones that their methods can yield.

In the RSM framework few authors have addressed the evaluation of response's properties to the extent it deserves. In general they focus on the objective function value used to select the compromise solution. However, that value depends on the weights assigned to responses and do not provide information on responses properties to the analyst. What is known is that a higher or a lower value is desirable, depending on the optimization scheme used, which does not guarantee that desired responses properties are achieved. Such as what Costa et al. [40] showed, these function values are not appropriate to identify a solution with desired response's properties. For this purpose they proposed optimization measures that allow assessing each one of those properties separately. Thus, whenever solutions more favorable than others in terms of bias, quality of predictions or robustness exist, the analyst can explore the method's solutions putting emphasis on the property(ies) of interest and selecting solutions feasible in practice.

The measures proposed in Ref. [40] were used in this article and the results of Cases 1–3 show that solutions with B_{cum} , QoP and Rob values similar or slightly better than those obtained through the objective functions can be achieved. To select the compromise solutions the analyst may take as reference the Cum value (cumulative value of B_{cum} , QoP and Rob). This measure is more useful than the loss and desirability composite functions, because they depend on the weights assigned to responses and give different results for similar solutions. For example: C11 is similar to C41, however their desirability values (D values) are very different; V31 and V51 have similar response properties, however they have different expected loss values; P12 has response properties worse than P32, however its expected loss is lower; K23 and V33 have similar response properties, however they have different expected loss values.

These results provide evidence that the optimization measures are effective, but this does not mean that they are the panacea to achieve optimal solutions. They serve to assess response's properties but not

**Fig. 9.** LQPUR—standard deviations.

for generate compromise solutions. Moreover, the analyst must be aware that points in non-convex response surfaces cannot be captured by weighted sums like those of the objective functions tested here [45], even if the proposed measures are used.

6. Conclusions

This article provides a review on desirability-based methods and evaluates, under adverse variance conditions, two of the less sophisticated ones, namely the methods proposed by Ch'ng et al. [27] and Wu [28]. To assess the properties of the compromise solutions at variables setting with respect to bias, quality of predictions and robustness, optimization measures were used. These measures can be used along with any method of practitioner's interest, and have proved to be useful for selecting a compromise solution taking into account the responses properties of the analyst or decision-maker preference.

The results show that none of the methods outperforms the others. Nevertheless, Ch'ng et al.'s method presents appealing features and along with the presented measures allows the analyst to identify compromise solutions with response properties similar or slightly better than those ones yielded by other methods used in practice. This gives confidence to suggest this approach for solving real life problems developed in the RSM framework as alternative to more sophisticated methods or approaches. This does not mean that the problem of choosing a method for optimizing multiple mean and standard deviation responses is resolved. A large number of methods, including parametric and non-parametric methods, from those requiring little information to the methods that require extensive information on each objective, have been reported in the literature. To compare their performance in adverse variance conditions and assess their ability to capture points in non-convex response surfaces are issues that deserve further research.

Acknowledgment

The authors are grateful to reviewers for their comments and suggestions. The Fundação para a Ciência e Tecnologia (FCT) through the PIDDAC Program funds (INESC-ID multiannual funding) supported the second author.

References

- [1] N. Costa, A. Pires, C. Ribeiro, Guidelines to help practitioners of design of experiments, *Total Qual. Manage. Mag.* 18 (2006) 386–399.
- [2] R. Myers, D. Montgomery, C. Anderson-Cook, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, third ed. Wiley, New Jersey, 2008.
- [3] M. Tanco, E. Viles, L. Ilzarbe, M. Alvarez, Implementation of design of experiments projects in industry, *Appl. Stochastic Models Data Anal.* 25 (2009) 478–505.
- [4] M. Govindaluri, B. Cho, Robust design modeling and optimization with correlated quality characteristics using a multicriteria decision framework, *Int. J. Adv. Manuf. Technol.* 32 (2007) 423–433.
- [5] K. Xu, D. Lin, L. Tang, M. Xie, Multiresponse systems optimization using a goal attainment approach, *IIE Trans.* 36 (2004) 433–445.
- [6] J. Kovach, B. Cho, Development of a multidisciplinary-multiresponse robust design optimization model, *Eng. Optim.* 40 (2008) 805–819.
- [7] E. Sivertsen, F. Bjerke, T. Almøy, V. Segtnan, T. Næs, Multivariate optimization by visual inspection, *Chemom. Intell. Lab. Sys.* 85 (2007) 110–118.
- [8] A. Messac, From dubious construction of objective functions to the application of physical programming, *AIAA J.* 38 (2000) 155–163.
- [9] J. Kovach, B. Cho, J. Antony, Development of an experiment-based robust design paradigm for multiple quality characteristics using physical programming, *Int. J. Adv. Manuf. Technol.* 35 (2008) 1100–1112.
- [10] J. Peterson, G. Miró-Quesada, E. Del Castillo, A Bayesian reliability approach to multiple response optimization with seemingly unrelated regression models, *Qual. Technol. Quant. Manage.* 6 (2009) 353–369.
- [11] L.-I. Tong, C.-C. Chen, C.-H. Wang, Optimization of multi-response processes using the VIKOR method, *Int. J. Adv. Manuf. Technol.* 31 (2007) 1049–1057.
- [12] P. Surajit, G. Kumar, Multi-response optimization using multiple regression-based weighted signal-to-noise ratio, *Qual. Eng.* 22 (2010) 336–350.
- [13] L.-I. Tong, K.-L. Hsieh, A novel means of applying neural networks to optimize the multiresponse problem, *Qual. Eng.* 13 (2000) 11–18.
- [14] R. Noorossana, S. Tajbaksh, A. Saghaei, An artificial neural network approach to multiple-response optimization, *Int. J. Adv. Manuf. Technol.* 40 (2009) 1227–1238.
- [15] M. Ortiz, L. Sarabia, A. Herrero, M. Sánchez, Vectorial optimization as a methodological alternative to desirability function, *Chemom. Intell. Lab. Sys.* 83 (2006) 157–168.
- [16] G. Derringer, R. Suich, Simultaneous optimization of several response variables, *J. Qual. Technol.* 12 (1980) 214–218.
- [17] G. Derringer, A balancing act: optimizing a product's properties, *Qual. Prog.* (June 1994) 51–58.
- [18] W. Dewé, R. Marini, P. Chiap, Ph. Hubert, J. Crommen, B. Boulanger, Development of response models for optimizing HPLC methods, *Chemom. Intell. Lab. Sys.* 74 (2004) 263–268.
- [19] P. Ebrahimi, R. Mohammad, Optimization of the separation of coumarins in mixed micellar liquid chromatography using Derringer's desirability function, *J. Chemom.* 21 (2007) 35–42.
- [20] G. Pages, C. Delaurent, R. Phan-Tan-Luu, M. Sergent, Different chemometric approaches to optimize the assay of St. John's Wort active ingredients, *Chemom. Intell. Lab. Sys.* 86 (2007) 159–167.
- [21] O. Akalin, K. Akay, B. Sennaroglu, M. Tez, Optimization of chemical admixture for concrete on mortar performance tests using mixture experiments, *Chemom. Intell. Lab. Sys.* 104 (2010) 233–242.
- [22] V.-M. Taavitsainen, A. Lehtovaara, M. Lähteenmäki, Response surfaces, desirabilities and rational functions in optimizing sugar production, *J. Chemom.* 24 (2010) 505–513.
- [23] G. Hahn, N. Doganaksoy, The future of industrial statistics: a panel discussion, *Technometrics* 50 (2008) 103–126.
- [24] E. Del Castillo, D. Montgomery, D. McCarville, Modified desirability functions for multiple response optimization, *J. Qual. Technol.* 28 (1996) 337–345.
- [25] F.-C. Wu, M. Hamada, *Experiments: Planning, Analysis, and Parameter Design Optimization*, Wiley-Interscience, New York, 2000.
- [26] K.-J. Kim, D. Lin, Simultaneous optimization of multiple responses by maximizing exponential desirability functions, *Appl. Stat. C* 49 (2000) 311–325.
- [27] C. Ch'ng, S. Quah, H. Low, A new approach for multiple-response optimization, *Qual. Eng.* 17 (2005) 621–626.
- [28] F.-C. Wu, Optimization of correlated multiple quality characteristics using desirability function, *Qual. Eng.* 17 (2005) 119–126.
- [29] E. Harrington, The desirability functions, *Ind. Qual. Control* 12 (1965) 494–498.
- [30] J. Kros, C. Mastrangelo, Comparing multi-response design methods with mixed responses, *Qual. Reliab. Eng. Int.* 20 (2004) 527–539.
- [31] F.-C. Wu, C. Chyu, Optimization of robust design for multiple quality characteristics, *Int. J. Prod. Res.* 42 (2004) 337–354.
- [32] C. Ribardo, T. Allen, An alternative desirability function for achieving 'six sigma', *Qual. Reliab. Eng. Int.* 19 (2003) 227–240.
- [33] F. Ortiz, J. Simpson, J. Pignatiello, A. Heredia-Langner, A genetic algorithm approach to multiple-response optimization, *J. Qual. Technol.* 36 (2004) 432–450.
- [34] S. Pasandideh, S. Niaki, Multi-response simulation optimization using genetic algorithm within desirability function framework, *Appl. Math. Comput.* 175 (2006) 366–382.
- [35] M.-S. Lee, K.-J. Kim, Expected desirability function: consideration of both location and dispersion effects in desirability function approach, *Qual. Technol. Quant. Manage.* 4 (2007) 365–377.
- [36] P. Das, S. Sengupta, Composite desirability index in cases of negative and zero desirability, *J. Manage. Res.* 10 (2010) 25–38.
- [37] Z. He, J. Wang, J. Oh, S. Park, Robust optimization for multiple responses using response surface methodology, *Appl. Stochastic Models Bus. Ind.* 26 (2010) 157–171.
- [38] F. Fogliatto, L. Albin, Variance of predicted response as an optimization criterion in multiresponse experiments, *Qual. Eng.* 12 (2000) 523–533.
- [39] H. Shah, D. Montgomery, W. Matthew, Response surface modeling and optimization in multiresponse experiments using seemingly unrelated regressions, *Qual. Eng.* 16 (2004) 387–397.
- [40] N. Costa, Z. Pereira, M. Tanco, Optimization measures for assessing compromise solutions in multiresponse problems, *World Congress on Engineering – ICMEEM'10*, London, 2010.
- [41] G. Vining, A compromise approach to multiresponse optimization, *J. Qual. Technol.* 30 (1998) 309–313.
- [42] R. Myers, D. Montgomery, *Response Surface Methodology: Process and Product Improvement with Designed Experiments*, second ed. John Wiley & Sons, New York, 1995.
- [43] Y.-H. Ko, K.-J. Kim, C.-H. Jun, A new loss function-based method for multiresponse optimization, *J. Qual. Technol.* 37 (2005) 50–59.
- [44] J. Pignatiello, Strategies for robust multiresponse, *IIE Trans.* 25 (1993) 5–15.
- [45] A. Messac, G. Sundararaj, R. Tappeta, J. Renaud, Ability of objective functions to generate points on non-convex pareto frontiers, *AIAA J.* 38 (2000) 1084–1091.