Computers & Industrial Engineering 62 (2012) 927–935

Contents lists available at SciVerse ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Responses modeling and optimization criteria impact on the optimization of multiple quality characteristics

Nuno Costa [a,c,*], João Lourenço [a,b], Zulema Lopes Pereira [c]

[a] IPS–Setubal College of Technology, 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

A R T I C L E I N F O

Article history:
Received 8 August 2010
Received in revised form 14 October 2011
Accepted 14 December 2011
Available online 22 December 2011

Keywords:
Desirability
Loss function
OLS
Robustness
SUR
Variance

A B S T R A C T

Responses modeling and optimization criteria impact on the optimization results were investigated. The Ordinary Least Squares and Seemingly Unrelated Regression techniques were illustrated in two examples from the literature and the performance of three optimization criteria evaluated. In contrast to the standard practice, compromise solutions were evaluated in terms of bias and robustness using optimization performance measures. The results of both examples show that responses modeling strongly impacts on the optimization results, while there is no significant difference between criteria performance. The Seemingly Unrelated Regression technique proved to be useful for modeling correlated responses. Otherwise, this technique can lead to results in close agreement to those obtained with models fitted with the OLS technique.

1. Introduction

The variety and quantity of methods available in the literature to deal with optimization of the mean and variance or standard deviation of multiple quality characteristics (responses) are large. Within the framework of the Response Surface Methodology (RSM), which was comprehensibly exposed by Myers, Montgomery, and Anderson-Cook (2009) and has been increasingly used in simulated and real-life problems (Ilzarbe, Álvarez, Viles, & Tanco, 2008), most authors have focused on what to optimize, developing or improving objective functions, in contrast to those who have proposed new optimization algorithms. In general, those methods provide to the user or Decision-Maker (DM) one way to assign weights or priorities to responses. This is a critical issue in multiresponse optimization because preference parameters are dependent on the particular DM and the context of the problem. Thus, methods that require a minimum amount of preference information from the DM are especially appealing and their performance must be evaluated to justify or not their recommendation for solving real-life problems. Another critical issue that has been underestimated is the responses modeling. Most case studies reported in the literature employ the Ordinary Least Squares (OLS) technique and "standard" second order models are fitted to responses. However, for problems where significant responses correlation exists, which are not unusual in practice, the OLS technique is not appropriate. Shaibu and Cho (2009) and Goethals and Cho (2011a) also show that higher-order polynomial models may be more effective in finding better compromise solutions than the commonly-used quadratic model.

The objective of this article is to investigate the impact of the responses modeling and objective functions designed with basis in different approaches on the results of multiresponse optimization problems developed under the RSM framework. In particular, the Seemingly Unrelated Regression (SUR) is presented as alternative to OLS technique and the solutions generated with a compromise programming-based method compared with those of two alternative optimization criteria.

The remainder of the article is organized as follows: Section 2 provides a review on the literature; the proposed optimization criterion is presented in Section 3; Section 4 overviews the SUR and OLS techniques; Section 5 uses two examples from the literature to evaluate the impact of methods and regression techniques on the results; Section 6 discusses the results and Section 7 presents the conclusions.

2. Literature review

Design and conduct experiments are not trivial tasks, especially for non-statistician practitioners who do not have the required

* Corresponding author at: IPS–Setubal College of Technology, Campus do IPS, Estefanilha, 2910-761 Setúbal, Portugal. Tel.: +351 265790000; fax: +351 265721869.

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

0360-8352/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cie.2011.12.015

background in the subject. There are technical issues that need to be carefully managed in addition to the non-technical ones that are thoroughly discussed by Tanco, Viles, Ilzarbe, and Alvarez (2009). For example, choosing an inappropriate experimental design will surely compromise the results or conclusions. To select an appropriate modeling technique and fit models to responses are additional difficulties in multiresponse optimization problems.

Several regression techniques can be used to fit models to responses. Fogliatto and Susan (2000) summarized the assumptions for using four response modeling techniques and give a tutorial on them. What can be stated from their article is the following:

- For uncorrelated responses: OLS is suitable when the responses' variance is homogeneous, while the Generalized Least Squares (GLS) is suitable when the responses' variance is non-homogeneous;
- For correlated responses: Multivariate Regression (MVR) and SUR are suitable when the responses' variance is homogeneous.

Shah, Montgomery, and Carlyle (2004) also reviewed the OLS and SUR techniques and presented the proof of their equivalency in two particular situations, namely, for zero correlations among all model errors and the same model form for each response with the same set of design variables. These authors also showed that the SUR technique can lead to more precise estimates of the regression coefficients than the OLS technique when responses are correlated. The OLS is the most-frequently used technique for estimating the response models in problems developed under the RSM framework (Khuri & Mukhopadhyay, 2011), even when responses are correlated. Moreover, authors rarely fit models of degree higher than two to responses, although diagnostic checks for models adequacy are usually performed using meaningful statistics.

The *R*-square and adjusted *R*-square have been used to assess the quality of description for models fitted with the OLS technique. In practice, high values for *R*-square and adjusted *R*-square, say higher than 0.9, are desirable. With the SUR technique has been used the system *R*-square (Jitthavech, 2010). The predicted *R*-square and prediction error sum of squares (PRESS), whose value must be as low as possible, have been used to assess the quality of predictions of models fitted with the OLS technique. These and other statistics are reviewed, for example, by Shaibu and Cho (2009) and Goethals and Cho (2012). The models adequacy checking is not limited to these statistics so the reader is referred to Myers et al. (2009) for more details on this issue.

As concerns the optimization criteria, a large variety of alternatives have been put forward in the literature. The desirability-based and loss function-based methods are very popular among practitioners. The desirability-based methods are easy to use, easy to understand, and flexible enough for incorporating the DM's preferences. Moreover, the most popular desirability-based method, the Derringer and Suich method (1980) with the modification introduced by Derringer (1994), is available in several data analysis software packages. However, the analyst needs to specify values to four types of shape parameters (weights) to use the method when the optimization problem includes responses of different types, namely: (1) Nominal-The-Best (NTB) – the value of the estimated response is expected to achieve a particular target value; Smaller-The-Better (STB) – the value of the estimated response is expected to be smaller than an upper bound; Larger-The-Better (LTB) – the value of the estimated response is expected to be larger than a lower bound. This is not a simple task and it impacts on the optimal variable settings. Moreover, the composite function value does not provide a clear interpretation except the principle that either a higher or lower value is preferred, depending on how the composite function is defined. For a tutorial on desirability-based methods the reader is referred to Costa, Lourenço, and Pereira (2011).

Loss function-based methods have also been widely used in practice, namely those that consider the responses' variance level and explore the information on responses correlation. These methods allow the assignment of priorities to individual responses and the results are expressed in monetary units, which are appealing features. However, different response scales, relative variabilities, and relative costs are difficult to take into account in the specification of cost coefficients. This is a critical task that seriously impacts on the final results because inaccurate specification of cost coefficients leads to less favorable compromise solutions (higher expected losses). For an extensive review on loss function-based methods the reader is referred to Murphy, Tsui, and Allen (2005).

Other methods that have been successful illustrated in the literature are those based on Goal Programming (Kazemzadeh, Bashiri, Atkinson, & Noorossana, 2008), Physical Programming (Messac, 2000); Probability-based (Peterson, Miró-Quesada, & Del Castillo, 2009), Lexicographic Weighted Tchebycheff (Shin & Cho, 2009), and Non-parametric models (Besseris, 2009).

Besides the theoretical complexity of some methods and unavailability of the algorithms employed, one premise inherent to those which require preference information from the DM is that he/she is capable of assigning priorities or relative importance to responses in order to find a satisfactory solution. The fundamental significance of the weights or priorities to responses has been studied by Jones (2011), Marler and Arora (2010), Jeong, Kim, and Chang (2005), who also proposed procedures to extract values representing the DM's preferences. However, to the best of our knowledge, a general procedure for multiresponse optimization problems, which ensures that the "best" compromise solution (responses as close as possible to their target values and with minimum variability around them) is achieved, has not been found yet. In general, those preference parameters do not reflect proportionally the relative importance of the responses, and may not be promptly available or easily defined. In practice, they merely locate points in the domain when are varied. Thus, the task of assigning priorities to responses is, in general, performed through a trial-and-error procedure that may be a source of frustration and significant inefficiency, particularly when either the number of responses or control factors is large. Recently, Lee, Kim, and Köksalan (2011) proposed a posterior preference articulation approach that initially finds a set of compromise solutions without preference information from the DM in advance, and then uses an interactive selection procedure based on pairwise comparisons to facilitate the selection of the most satisfactory solution among the set initially generated. However, they conclude that ways of improving the set of generated compromise solutions and efficiency of the selection method require further research. The intricate problem of assigning priorities to responses remains an open research field, and it is known that practitioners prefer methods that are easy to understand and use in addition to effectiveness (Ilzarbe et al., 2008). Thus, a compromise programming-based method (hereafter denoted as the CP method) that requires a minimum amount of subjective information from the DM was selected from the literature and its working abilities evaluated. This method is presented in the next section.

3. Optimization criterion

Compromise Programming is a mathematical programming technique which has proven to be extremely powerful in incorporating and resolving conflicting objectives concurrently for locating efficient solutions in convex and non-convex response