Chemometrics and Intelligent Laboratory Systems 138 (2014) 171–177

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

A comparative study of multiresponse optimization criteria working ability

Nuno Costa [a,b,*], João Lourenço [a,c,1]

[a] Escola Superior de Tecnologia de Setúbal, Campus do IPS, Estefanilha, 2910-761 Setúbal, Portugal
[b] UNIDEMI, Faculdade de Ciências e Tecnologia-Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
[c] INESC-ID, Rua Alves Redol 9, 1000-029 Lisboa, Portugal

ARTICLE INFO

Article history:
Received 28 January 2014
Received in revised form 31 May 2014
Accepted 2 August 2014
Available online 9 August 2014

Keywords:
Compromise
Desirability
Multiobjective
Optimal
Pareto
RSM

ABSTRACT

There is not a unique optimal solution for problems with multiple responses. Some solutions lead to operation conditions that are more hazardous, more costly or more difficult to implement and control. Therefore, it is useful for the analyst or decision-maker to use a criterion that can capture solutions evenly distributed along the so-called Pareto frontier. To provide information about criteria's working ability to depict Pareto frontiers, four optimization criteria built on different approaches were evaluated. Results show differences in criteria's performance. In particular, a consistent performance of a global criterion and limitations of the widely used desirability-based criterion is stated.

1. Introduction

Simultaneous optimization of multiple objectives (quality characteristics or responses) is a usual problem in all science fields and often involves incommensurate and conflicting responses that must be in some sense optimized simultaneously, because their separate analysis may result in incompatible solutions [1]. From visual inspection of contour plots [2] to refined optimization methods that have been used in Chemometrics [3–6], the variety and quantity of optimization methods are large.

Numerous case studies recently published (to cite only four manuscripts see Ref. [7–10]) illustrate optimization method's usefulness, though the optimization problem in a multiresponse setting is not as well defined as in the single response case. A relevant difference, which has not been enough valued by researchers and practitioners, is that no solution is the best for problems with multiple responses. Typically, multiresponse optimization (MO) problems have many optimal solutions that impact differently on process or product. Some of them may lead to operation conditions that are more hazardous, more costly or more difficult to implement and control. Therefore, it is useful for the analyst or decision-maker (*DM*) to use a method or criterion that can capture solutions evenly distributed along the so-called Pareto frontier. If the criterion fails to capture this set of nondominated solutions (solutions where any improvement in one response cannot be done without degrading the value of, at least, another response), the *DM* may have denied the possibility of finding a more favorable compromise solution. However, optimization criteria ability to depict Pareto Frontiers has been rarely evaluated in the Response Surface Methodology (RSM) framework, which makes the practitioner's task difficult in choosing an effective criterion to solve MO problems. In addition, some criteria display many flaws as regards the efficient solution of real world problems, and the final choice of solution is often not the best possible choice, although it appears to be the optimal solution from among those considered by the algorithm applied [6]. This article contributes for reducing this gap in the literature and help researchers and practitioners in making more informed decisions when they need to select an optimization criterion. Four criteria, including the widely popular desirability-based criterion proposed by Derringer and Suich [11], which has been used in many science fields, namely in Chemometrics and related areas (see, as instance, Ref. [12–19]), are evaluated using three examples from the literature.

The remainder of the article is organized as follows: Next section provides a literature overview; then selected optimization criteria are reviewed; Sections 4 and 5 include the examples and results discussion; conclusions are presented in Section 6.

* Corresponding author. Tel.: + 351 265790000.
E-mail addresses: nuno.costa@estsetubal.ips.pt (N. Costa), joao.lourenco@estsetubal.ips.pt (J. Lourenço).
[1] Tel.: +351 265790000.

http://dx.doi.org/10.1016/j.chemolab.2014.08.004
0169-7439/

2. Literature overview

The statistical design and analysis of experiments, namely the RSM, which is comprehensively exposed in Ref. [20,21], have been extensively used in analytical chemistry [22], electroanalytical chemistry [23], and in chemistry in general [24] as a tool for improving and validating a formulation, a product, a process, or an analytical method.

Regarding the optimization of conflicting responses, which has been a widely researched subject and is the focus of this manuscript, a strategy often used to solve these problems consists of aggregating second order models fitted to responses into a single function followed by its optimization. To employ this reduction strategy, a great quantity and variety of composite functions or optimization criteria are available in the literature, though the most popular are built on desirability and loss function approaches.

An extensive review on desirability-based criteria is presented in Ref. [1]. From the popular desirability criterion proposed by Derringer and Suich [11], later modified by Derringer [25], to less known proposal of Das and Sengupta [26] who modified Gatza's desirability function to accommodate customer perceptions that take positive and negative values, twelve methods were reviewed. Murphy et al. [27] provided an extensive review on loss function-based criteria and summarized the relative merits of twelve multivariate loss-based and desirability-based criteria; Ko et al. [28] combined the strengths of two popular criteria, namely the Pignatiello's [29] and Vining's [30] criteria. Integration of a double-exponential desirability function with a loss function is illustrated in Ref. [31]. Relationships of desirability, loss, and utility approaches developed in the field of Operations Research/Management Science for Multiple Criteria Decision Making are highlighted in Ref. [32].

Other contributions introduced in the last decade include the mean squared error [33], weighted signal-to-noise ratio [34], PCA-based grey relational analysis [35], weighted principal component [36], capability index [37,38], patient rule induction [39,40], design envelopment analysis [41,42], compromise programming [43,44], goal programming [45–49], physical programming [50,51], Bayesian probability [52], augmented and lexicographic weighted Tchebycheff formulations [53,54], and modified ε-constraint method [55–57]. Relationships and differences among commonly used criteria are highlighted by Ardakani and Wulff [58], who also identified the issues faced by the *DM* to solve MO problems. This list is not exhaustive. Many other researchers have contributed to the growing wealth of knowledge in the field. However, little attention is paid to the criteria's ability to depict Pareto frontiers. Exceptions are the works reported in Ref. [53,59,60], though these studies only evaluated a small number of criteria and disregard the most popular one for those who have solved multiresponse optimization problems in the RSM framework; the Derringer and Suich's criterion.

3. Optimization criteria

In the next subsections four criteria built on different approaches are reviewed: two desirability-based criteria, namely the popular geometric mean (DGM) introduced by Derringer [25] and the arithmetic mean (DAM) introduced by Ch'ng et al. [61]; a global criterion-based (GC) criterion, and a lexicographic weighted Thchebichev (LWT) criterion. Popularity, ease of implementation, and performance were the major guidelines to select them.

3.1. DGM criterion

Derringer [25] modified the Derringer and Suich's method aggregating individual desirability functions into a weighted geometric mean defined as

$$D = \left((d_1)^{\omega_1}(d_2)^{\omega_2}\ldots\left(d_p\right)^{\omega_p}\right)^{\frac{1}{\sum \omega_i}} \tag{1}$$

where d_i is the individual desirability function of the i-th response ($i = 1, \ldots, p$), and ω_i are user-specified parameters to assign 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).

Derringer and Suich [11] proposed one-sided desirability transformations for Larger-The-Best (LTB) response type (the estimated response value is expected to be larger than a lower bound L; $\hat{y}>L$) as

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

where r is a user-specified parameter ($r > 0$), $\hat{y}$ represents the estimated response model, and U is the upper bound, such that $d = 1$ for $\hat{y}\geq U$ and $d = 0$ for $\hat{y}\leq L$. For Smaller-The-Best (STB) response type (the estimated response value is expected to be smaller than the upper bound U; $\hat{y}<U$) as

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

such that $d = 1$ for $\hat{y}\leq L$, and $d = 0$ for $\hat{y}\geq U$.

For two-sided transformations, which arise when the value of the estimated response is expected to achieve a particular target value (T), called Nominal-The-Best (NTB) response type, the individual desirability functions are as follows:

$$d = \begin{cases} \left(\frac{\hat{y}-L}{T-L}\right)^s & , \quad L\leq\hat{y}\leq T \\ \left(\frac{\hat{y}-U}{T-U}\right)^t & , \quad T\leq\hat{y}\leq U \\ 0 & , \quad \text{otherwise} \end{cases} \tag{4}$$

where s and t are user-specified parameters (s, t > 0), $d = 1$ for $\hat{y} = T$, and $d = 0$ for $\hat{y}<L$ or $\hat{y}>U$. Specification limits denoted by U and L are usually available for product or process quality control.

3.2. DAM criterion

Ch'ng et al. [61] proposed to minimize an arithmetic mean (D)

$$D = \left(\sum_{i=1}^{p}\omega_i|d_i-d_i(T_i)|\right)/p \tag{5}$$

where $d_i(T_i)$ is the value of the i-th individual desirability function for $\hat{y}_i$ at target value, ω_i represents the priority (weight or importance) assigned to $\hat{y}_i$, p is the number of responses, and $\sum_{i=1}^{p}\omega_i = 1$. The individual desirability functions are defined as

$$d = \frac{2\hat{y}-(U+L)}{U-L}+1 = \frac{2}{U-L}\hat{y}+\frac{-2L}{U-L} = m\hat{y}+c \tag{6}$$

with $0 \leq d \leq 2$.

3.3. GC criterion

Costa and Pereira [62] proposed to minimize an arithmetic function defined as

$$\sum_{i=1}^{p}\left(\frac{|\hat{y}_i-T_i|}{U_i-L_i}\right)^{\omega_i} \tag{7}$$

where ω_i is the user-specified parameters (shape or power factors, $\omega_i > 0$). In this criterion, for STB-type response $T = L$ and for LTB-type response $T = U$.