QUALITY ENGINEERING®
Vol. 16, No. 3, pp. 377–385, 2004

Dual-Response Surface Optimization: A Weighted MSE Approach

Rui Ding,[1,*] Dennis K. J. Lin,[1] and Duan Wei[2]

[1]Pennsylvania State University, University Park, Pennsylvania, USA
[2]Taiwan Research Institute, Taipei, Taiwan

ABSTRACT

Dual-response surface methodology is a powerful tool for simultaneously optimizing the mean and the variance of responses in quality engineering. In this article, we suggest a weighted mean squared error (MSE) approach to improve the optimization procedure. In addition, we propose a data-driven approach to determine the weights when the prior information is vague. This is based on the idea of an "efficient curve." Examples are given to illustrate the superiority of the proposed method, as compared with other existing procedures.

Key Words: Efficient curve; Efficient point; Quality engineering.

INTRODUCTION

Response surface methodology is an important tool in modern quality engineering. The basic idea is to fit a model for the response variable and then explore various settings of interest for the explanatory variables. The main emphasis has been on maximizing (or minimizing) the mean value (location) of the response Y. This approach works well under the assumption of the homogeneous variance. However, such an assumption may not be valid in real-life applications. Taguchi (1986) emphasized the need for developing statistical methodology that can simultaneously optimize the mean and the variance of the characteristic being investigated (Phadke, 1989). Such a problem arises in many industrial problems, which require simultaneously achieving a target value and keeping the variance small.

For illustrative purposes, we consider the printing process example given in Box and Draper (1987). The experiment is a 3^3 full-factorial design that has three experimental variables: speed (x_1), pressure (x_2), and distance (x_3) with three replicates at each design combination. The experimental response (y) is a printing machine's ability to apply coloring inks on

*Correspondence: Rui Ding, Ph.D. candidate, 325 Thomas Building, Department of Statistics, Pennsylvania State University, University Park, PA 16802, USA; Fax: 1-814-863-7114; E-mail: rding@stat.psu.edu.

DOI: 10.1081/QEN-120027940

0898-2112 (Print); 1532-4222 (Online)
www.dekker.com

Table 1. The printing process data from Box and Draper (1987).

X_1	X_2	X_3	Y_1	Y_2	Y_3	Mean	Standard deviation
−1	−1	−1	34	10	28	24	12.49
0	−1	−1	115	116	130	120.3	8.39
1	−1	−1	192	186	263	213.7	42.8
−1	0	−1	82	88	88	86	3.46
0	0	−1	44	178	188	136.7	80.41
1	0	−1	322	350	350	340	16.17
−1	1	−1	141	110	86	112.3	27.57
0	1	−1	259	251	259	256.3	4.62
1	1	−1	290	280	245	271.7	23.63
−1	−1	0	81	81	81	81	0
0	−1	0	90	122	93	101.7	17.67
1	−1	0	319	376	376	357	32.91
−1	0	0	180	180	154	171.3	15.01
0	0	0	372	372	372	372	0
1	0	0	541	568	396	501.7	92.5
−1	1	0	288	192	312	264	63.5
0	1	0	432	336	513	427	88.61
1	1	0	713	725	754	730.7	21.08
−1	−1	1	364	99	199	220.7	133.8
0	−1	1	232	221	266	239.7	23.46
1	−1	1	408	415	443	422	18.52
−1	0	1	182	233	182	199	29.45
0	0	1	507	515	434	485.3	44.64
1	0	1	846	535	640	673.7	158.2
−1	1	1	236	126	168	176.7	55.51
0	1	1	660	440	403	501	138.9
1	1	1	878	991	1161	1010	142.5

package labels. The goal here is to determine the "optimal" settings of experimental variables (x_i's) such that the response (y) will be close to a target value T, while keeping the variance small. We will later apply our proposed method on this example and compare it with methods suggested by others. The data is displayed in Table 1.

This article is organized as follows. After reviewing popular dual-response surface optimization methods for this problem, we discuss the weighted average method and introduce the idea of "efficient point" and "efficient curve." It is shown that a good solution for the dual-response problem must be located on the efficient curve. We next propose a method to determine proper weights for such a weighted average method when the information on weights is vague. The printing example above is then used to demonstrate how to apply the proposed method and its advantages over other existing procedures. A general step-by-step procedure is provided for practitioners, using a second example for illustration.

LITERATURE REVIEW

Vining and Myers (1990) point out that the goal of optimizing the mean and the variance simultaneously can be realized via a dual-response surface method. Specifically, suppose the response variable is Y and the controllable experiment variables are $x_1, \ldots, x_k$; Vining and Myers (referred to as VM) first fit second-order polynomial models for the sample mean (ω_μ) and the sample standard deviation (ω_σ) separately.

$$\omega_\mu = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum\sum_{i<j}^{k} \beta_{ij} x_i x_j + \varepsilon_\mu \tag{1}$$

$$\omega_\sigma = \gamma_0 + \sum_{i=1}^{k} \gamma_i x_i + \sum_{i=1}^{k} \gamma_{ii} x_i^2 + \sum\sum_{i<j}^{k} \gamma_{ij} x_i x_j + \varepsilon_\sigma \tag{2}$$

Then the VM procedure optimizes the scheme:

$$\underset{X}{\text{Minimize}} \quad \hat{\omega}_\sigma \qquad \text{(VM)} \quad (3)$$

subject to $\hat{\omega}_\mu = T$, where T is target value.

Note that VM also discusses the cases of "the larger, the better" and "the smaller, the better." Without loss of generality, we focus only on the case of "the target is best." As will be shown later, the extensions to the other two cases are rather straightforward.

Lin and Tu (1995), denoted by LT, show that the VM method above may rule out better conditions due to the restriction that the estimate of the second response is forced to a fixed value. They propose to use the mean squared error (MSE) criterion, namely,

$$\underset{x}{\text{Minimize}} \quad \text{MSE} = (\hat{\omega}_\mu - T)^2 + \hat{\omega}_\sigma^2 \quad \text{(LT)} \qquad (4)$$

Copeland and Nelson (1996), denoted by CN, comment that LT does not specify how far one would be willing to allow ω_μ to deviate from T. Instead, they modify VM by adding restrictions on ω_μ:

$$\begin{aligned} &\underset{x}{\text{Minimize}} \quad \hat{\omega}_\sigma \\ &\text{subject to} \quad (\hat{\omega}_\mu - T)^2 \le \Delta^2 \quad \text{(CN)} \end{aligned} \qquad (5)$$

Kim and Lin (1998), denoted by KL, introduce a fuzzy modeling approach. The general idea is to assign a membership function (or desirability function) $m(.)$ for both $\hat{\omega}_\mu$ and $\hat{\omega}_\sigma$ to measure the "desirability." They suggest the use of an exponential function of the form:

$$m(z) = \begin{cases} \dfrac{e^d - e^{d|z|}}{e^d - 1} & \text{if } d \neq 0 \\ 1 - |z| & \text{if } d = 0 \end{cases} \qquad (6)$$

where d is a chosen parameter.

They formulate the optimization problem as:

$$\begin{aligned} &\underset{x}{\text{Maximize}} \quad \lambda \\ &\text{subject to} \quad m(\hat{\omega}_\mu) \ge \lambda \\ &\qquad\qquad m(\hat{\omega}_\sigma) \ge \lambda \\ &\qquad\qquad x \in \Omega \quad \text{(KL)} \end{aligned} \qquad (7)$$

where Ω may include any restriction conditions as in the CN method.

WEIGHTED AVERAGE METHOD

When optimizing two functions simultaneously, say $f_1(x) > 0$ and $f_2(x) > 0$, we can combine them into a single objective function by taking the convex combination of them, namely,

$$\begin{aligned} &\underset{x}{\text{Minimize}} \; f(X) = \lambda f_1(X) + (1-\lambda) f_2(X) \\ &\text{subject to } X \in \Omega, \end{aligned} \qquad (8)$$

where $\lambda \in [0, 1]$ and Ω is the region of interest.

One can choose different values of weight λ according to the relative importance of two response functions. When λ equals 0 or 1, this is called a marginal optimization. This weighted method is very flexible and can achieve a good balance between marginal optimizations. While the method is computationally simple and often used in practice, a key issue to be addressed, of course, is the choice of the weight λ.

When λ is known, the optimal solution X^* of $f(X)$ can be obtained by single-function optimization. Note that the optimal solution X^* is a function of λ. We can plot the point $(f_1(X^*), f_2(X^*))$ for such a prespecified value of λ, called an efficient point. If we let the weight λ vary from 0 to 1, as demonstrated in Fig. 1, we will end up with a graph of $f_1(X^*)$ vs $f_2(X^*)$. This is usually called an efficient curve or efficient frontier in the optimization literature (see, e.g., Miettinen, 1999). It is the curve of all feasible solutions obtained by different weights.

Note that any optimal solution, X_0, for simultaneously minimizing $f_1(X)$ and $f_2(X)$, must fall on the efficient curve. To see this, first, X_0 cannot be above the efficient curve. Consider any point of $(f_1(X^*), f_2(X^*))$ represented by point A in Fig. 1. Points B and C on the efficient curve will provide two better solutions. For example, point B gives the identical value of $f_2(X^*)$ but a smaller $f_1(X^*)$, and point C gives the identical $f_1(X^*)$ but a smaller $f_2(X^*)$. Second, X_0 cannot be below the efficient curve. Suppose it is below the efficient curve. Then by definition, there exists a $0 < \lambda^* < 1$ and X^*, where X^* is the solution of Eq. (8), and we have $f_1(X_0) = f_1(X^*)$ and $f_2(X_0) < f_2(X^*)$. This implies $\lambda^* f_1(X_0) + (1-\lambda^*) \times f_2(X_0) < \lambda^* f_1(X^*) + (1-\lambda^*) f_2(X^*)$, which is a contradiction to the fact that X^* is the solution of Eq. (8). So an optimal solution must fall on the efficient curve.

All the solutions to the above scheme with different weights are considered to be equally good in some sense. It is usually up to the practitioners to decide which value for λ is the most appropriate. In some situations, the practitioners may have a very clear idea of the suitable values for λ (based upon the past experience, for example). More likely, clear

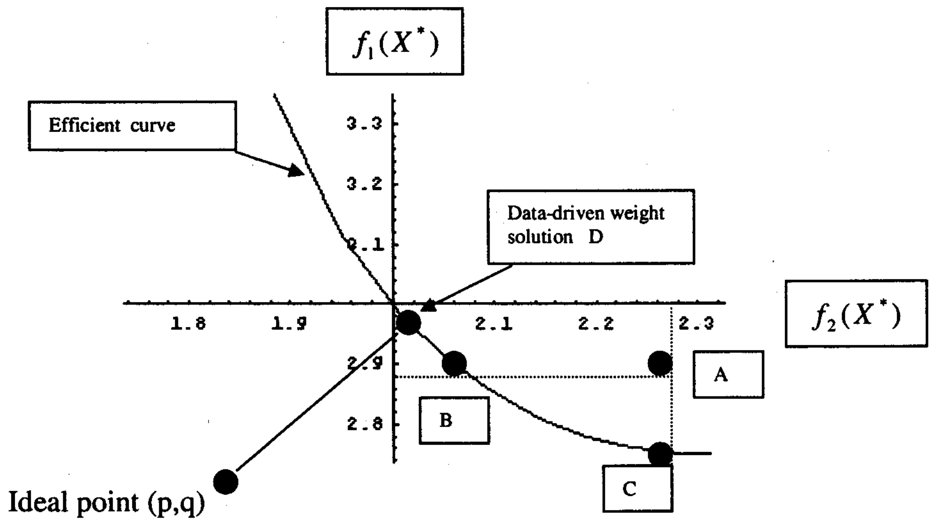

Figure 1. The illustration plot of $f_2(X^*)$ vs. $f_1(X^*)$.

information about λ is not available and the practitioners may have to make a wild guess or simply place an equal weight using $\lambda = 0.5$ regardless. Here, we provide a data-driven weighted average (DDWA) approach, as follows. The basic idea here is motivated by the concept of generalized distance in Khuri and Colon (1981).

Suppose we have fitted dual responses $f_1(x)$ and $f_2(x)$ with achieved marginal optimizations (p, q), that is, $\min_x f_1(X) = p$ and $\min_x f_2(X) = q$. The ideal point (p, q) represents the best optimization possible. This is typically not obtainable, however. For any specific feasible X^*, we will usually have either $f_1(X^*) > p$ or $f_2(X^*) > q$.

A data-driven weight is defined as the point on the efficient curve that is closest to the ideal point (p, q). As illustrated in Fig. 1, point D is the point on the efficient curve, that is closest to the ideal point (p, q), and thus is the solution from the proposed method. The solution achieved by this data-driven weighted average method has two advantages: (1) among all feasible solution points on the efficient curve, it is closest to the ideal point; and (2) it achieves a balance in the sense that it is determined by both marginal optimizations.

In summary, if one has prior knowledge or strong preference for the weight λ, then the optimal solution X^* can be obtained by solving Eq. (8) directly. Otherwise, all solutions on the efficient curve are considered to be good. Among them, the data-driven weight as described here seems to be a natural choice. This is particularly important when a specific optimal solution is required.

FORMULATION OF THE DATA-DRIVEN WEIGHTED MSE METHOD

Specifically, when applying the above idea to the dual-response problem, it is desirable that the mean be close to the target value T and the variance be small. A natural choice is to let $f_1(x) = (\hat{\omega}_\mu(X) - T)^2$ and $f_2(x) = \hat{\omega}_\sigma^2(X)$. The optimization scheme is as follows (denoting WMSE as the weighted MSE):

$$\begin{aligned} &\underset{x}{\text{Minimize}}\ \text{WMSE} = \lambda(\hat{\omega}_\mu(x) - T)^2 + (1-\lambda)\hat{\omega}_\sigma(x)^2 \\ &\text{subject to } X \in \Omega \\ &\text{where } \lambda \in [0, 1] \end{aligned} \tag{9}$$

Note that the LT method is a special case of Eq. (9) by taking $\lambda = 0.5$. In general, the solution of X for minimizing $\lambda(\hat{\omega}_\mu - T)^2 + (1-\lambda)\hat{\omega}_\sigma^2$ may not be unique, although these Xs will result in one single point in Fig. 1.

For the simplicity of the presentation, we define $p = \hat{\omega}_\mu(X_1^*)$, where X_1^* minimizes $(\hat{\omega}_\mu(X) - T)^2$ and $q = \hat{\omega}_\sigma(X_2^*)$, where X_2^* minimizes $\hat{\omega}_\sigma^2(X)$. The efficient curve that represents the best simultaneous optimizations possible is plotted. We then find the data-driven weight such that the distance between the efficient curve and the idea point (p, q) is minimized. The corresponding solution X^* of Eq. (9), when λ takes the value of such a weight, is the optimal solution of our proposed method.

For completeness sake, the special cases when $\lambda = 0$ or 1 deserve special mention. For example, when $\lambda = 1.0$, WMSE equals to $(\hat{\omega}_\mu(X) - T)^2$. The solution of X is not unique. Each of them may result in a different value of $\hat{\omega}_\sigma(X^*)$ and thus produce more than one point in Fig. 1. Note that if solutions for $(\hat{\omega}_\mu(X) - T)^2 = 0$ (ideal point) exist, then among all solutions, the solution X^* with minimal $\hat{\omega}_\sigma(X^*)$ will be the solution of VM.

PRINTING PROCESS EXAMPLE REVISITED

We now return to the printing process example in Table 1. Vining and Myers (1990) fit a quadratic model for the mean and standard deviation of the variable within a cubic region as follows, where $-1 \leq x_i \leq 1$ $(i = 1, 2, 3)$ and the target value for ω_μ is $T = 500$:

$$\begin{aligned}\hat{\omega}_\mu &= 327.6 + 177.0x_1 + 109.4x_2 + 131.5x_3 \\ &\quad + 32.0x_1^2 - 22.4x_2^2 - 29.1x_3^2 + 66.0x_1x_2 \\ &\quad + 75.5x_1x_3 + 43.6x_2x_3 \\ \hat{\omega}_\sigma &= 34.9 + 11.5x_1 + 15.3x_2 + 29.2x_3 + 4.2x_1^2 \\ &\quad - 1.3x_2^2 + 16.8x_3^2 \\ &\quad + 7.7x_1x_2 + 5.1x_1x_3 + 14.1x_2x_3 \qquad (10)\end{aligned}$$

Assuming the models have been correctly identified, we first solve marginal optimizations by taking $\lambda = 0$ and 1, i.e., solving min $(\omega_\mu - T)^2$ and min ω_σ^2 separately. Let p and q be the obtained marginal optimization values of ω_μ and ω_σ, respectively. We have $p = \hat{\omega}_\mu(X^*) = 500.000$ when $\lambda = 1.0$ and $q = \hat{\omega}_\sigma(X^*) = 14.758$ when $\lambda = 0$. Thus, $(p,q) = (500, 14.758)$ is the ideal point. We next solve the optimal solution X^* in Eq. (9) for various λ ranging from 0 to 1. We can use any nonlinear optimization software to find X^*. One possible software is the add-on package for Mathematica, called Multiplier-Method (Culioli and Skudlarek, 2000). Using these X^*s, we can plot the efficient curve $\hat{\omega}_\mu(X^*)$ vs. $\hat{\omega}_\sigma(X^*)$. We then find the point on the curve closest to point (p,q), i.e., a point among these $(\hat{\omega}_\mu(X^*), \hat{\omega}_\sigma(X^*))$ that minimizes $[\hat{\omega}_\mu(X^*) - p]^2 + [\hat{\omega}_\sigma(X^*) - q]^2 = [\hat{\omega}_\mu(X^*) - 500]^2 + [\hat{\omega}_\sigma(X^*) - 14.758]^2$. The corresponding weight is the data-driven weight, and the corresponding solution X^* is the optimal design solution for this dual-response problem.

Figure 2(a) illustrates the efficient curve and the ideal optimization point (p, q). The solutions at $\lambda = 0$ and 1.0 (when minimizing only one response) are not unique and thus produce more than one point. For example, in the case of $\lambda = 1.0$, there are several X^*s that minimize $(\hat{\omega}_\mu(X) - T)^2$, resulting in different values of $\hat{\omega}_\sigma(X^*)$ and thus several points of $(\hat{\omega}_\mu(X^*), \hat{\omega}_\sigma(X^*))$. The points shown in Fig. 2(a) are the ones chosen by Mathematica. Note that VM performs a further step optimization of ω_σ among all such X^*s. However, any point near the end of the efficient curve can be a better solution than that of VM.

The near-straight line for the efficient curve is due to the fact that all the optimal X^*s are on the boundaries of $x_1 = 1$ in this example. This is not necessarily the case in general. The efficient point closest to (p, q) is $(\hat{\omega}_\mu(X^*), \hat{\omega}_\sigma(X^*) = (496.473, 44.671)$. This is obtained at $X^* = (1.000, 0.089, -0.255)$ with the weight $\lambda = 0.6$, as shown in Fig. 2(b). Note that Fig. 2(a) is not plotted in the same scale for both responses. So the line between our solution and idea point (p, q) does not look perpendicular to the efficient curve, although, in fact, it is. The results of several methods are clustered together in Fig. 2(a). Figure 2(b) is the magnified local graph, so one can better visualize the differences. Figure 2(b) shows that the LT and KL methods result in solutions as a special case of the proposed method with $\lambda = 0.5$ and 0.58, respectively. While LT is always a special case of the proposed method as previously discussed, whether KL is always the case deserves further investigation.

A direct comparison of the various numerical solutions is in general not possible, because each alternative method seeks to optimize a different criterion. As previously discussed, however, any solution far away from the efficient curve is not recommended. Figure 2(b) shows that better solutions than those obtained by CN and VM methods are the solutions with $\lambda = 0.52$ and 0.99 (or any weight close to but not equal to 1), respectively. While the solution of CN is quite close to the efficient curve, the solution of VM is not, as indicated below. Compared to the VM solution $(\hat{\omega}_\mu(X^*), \hat{\omega}_\sigma(X^*)) = (500, 51.90)$, the solution on the efficient curve $(\hat{\omega}_\mu(X^*), \hat{\omega}_\sigma(X^*)) = (499.265, 45.01)$ does not have its predicted mean equal to 500 (T) exactly, but its predicted standard deviation is much smaller (45.01, as compared to 51.90 in VM). Furthermore, we can easily add any constraints to Ω in Eq. (9) to address the concern of Copeland and Nelson (1996), i.e., how far one would be willing to allow ω_μ to deviate from T, if so desired.

Method	$\hat{\omega}_\mu(X^*)$	$\hat{\omega}_\sigma(X^*)$
CN ($\rho = 1$, $\Delta = 5$)	495.020	44.727
Weighted MSE with $\lambda = 0.52$*	495.088	44.510
VM	500.000	51.900
Weighted MSE with $\lambda = 0.99$*	499.265	45.010

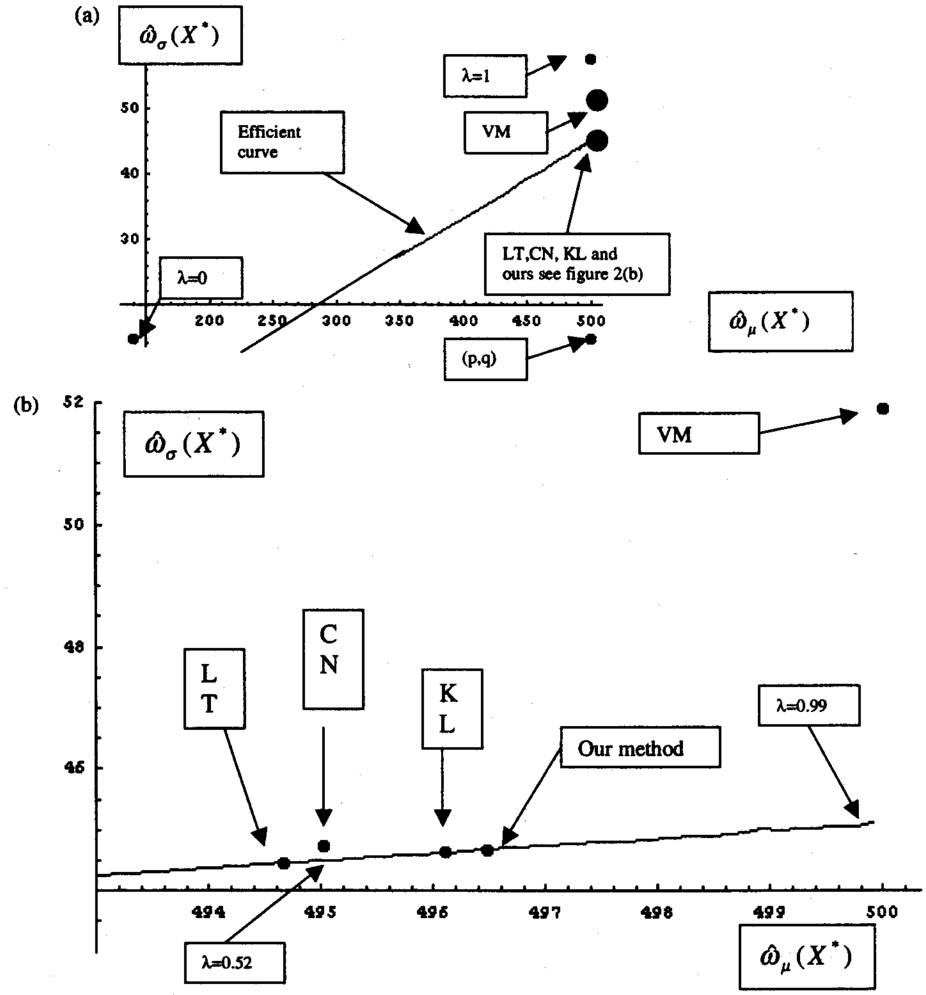

Figure 2. (a) The overall plot of $\hat{\omega}_\mu(X^*)$ vs. $\hat{\omega}_\sigma(X^*)$ in example 1. (b) The local plot of the Fig. 2(a) with $\lambda > 0.4$.

The comparison results with other approaches (VM, LT, CN, and KL) are summarized in Table 2, including the optimal solutions of X^* and the resulting $\hat{\omega}_\mu(X^*)$ and $\hat{\omega}_\sigma(X^*)$. The parameters of KL and CN methods are chosen to make a fair comparison with the proposed model. All numerical results are taken directly from the original articles. In Table 2, the corresponding weights λ of their results (or the weights of a more efficient point on the curve) are also listed in the last column.

GENERAL PROCEDURES FOR SOLVING PRATICAL DUAL-RESPONSE PROBLEMS

Specifically, the following steps are proposed for solving dual-response surface optimization problems:

1. Develop the experimental design, conduct the experiments, and collect data.
2. Fit response surfaces for the mean (ω_μ) and standard deviation (ω_σ) responses separately.
3. Plot the efficient curve using the solutions of Eq. (9) and determine the ideal optimization point (p, q).
4. Choose a solution from the efficient curve according to a prior weight λ or select a data-driven weight with the solution closest to the ideal optimization (p, q).

We next demonstrate the step-by-step procedure by an example with constraints.

A SECOND EXAMPLE

This example is from Luner (1994) and also studied by Kim and Lin (1998). Three variables, x_1 (arm length), x_2 (stop angle), and x_3 (pivot height),

Table 2. Comparison with other methods.

Method	Optimal Setting X^*	$\hat{\omega}_\mu(X^*)$	$\hat{\omega}_\sigma(X^*)$	λ
VM	(0.620, 0.230, 0.100)	500.000	51.900	0.99[a]
LT	(1.000, 0.074, −0.252)	494.659	44.463	0.50
KL ($d_\mu = -4.39$, $d_\sigma = 0$)	(1.000, 0.086, −0.254)	496.111	44.632	0.58
CN ($\rho = 1$, $\Delta = 5$)	(0.975, 0.056, −0.214)	495.020	44.727	0.52[a]
Proposed method	(1.000, 0.089, −0.255)	496.473	44.671	0.60

[a]The solution is not on the efficient curve. But by choosing an efficient point with such a weight, we can achieve the identical (or near identical) $\hat{\omega}_\mu(X^*)$ with a smaller $\hat{\omega}_\sigma(X^*)$.

are under consideration to predict the distance to the point where a projectile landed from the base of the Roman-style catapult. The experiment is a central composite design with three replicates.

Here are the steps:

1. The experiment is conducted and the result is displayed in Kim and Lin (1998).
2. Fit the second-order regression models for the mean response and standard deviation (see Kim and Lin (1998).

$$\hat{\omega}_\mu = 84.88 + 15.29x_1 + 0.24x_2 + 18.80x_3 - 0.52x_1^2 - 11.80x_2^2 + 0.39x_3^2 + 0.22x_1x_2 + 3.60x_1x_3 - 4.42x_2x_3$$

and

$$\hat{\omega}_\sigma = 4.53 + 1.84x_1 + 4.28x_2 + 3.73x_3 + 1.16x_1^2 + 4.40x_2^2 + 0.94x_3^2 + 1.20x_1x_2 + 0.73x_1x_3 + 3.49x_2x_3$$

3. The problem requires that the target value for the mean be 80 with constraints $79 \le \hat{\omega}_\mu \le 81$ and $\hat{\omega}_\sigma \le 3.5$. Equation (9) now becomes

$$\underset{x}{\text{Minimize}}\ \text{WMSE} = \lambda(\hat{\omega}_\mu - T)^2 + (1-\lambda)\hat{\omega}_\sigma^2$$
$$\text{subject to } 79 \le \hat{\omega}_\mu \le 81$$
$$\hat{\omega}_\sigma \le 3.5$$
$$-1 \le x_i \le 1, i = 1, 2, 3, \text{ where } \lambda \in [0,1] \text{ and } T = 80 \quad (11)$$

We first run marginal optimizations, i.e., $\min(\omega_\mu - T)^2$ and $\min\ \omega_\sigma^2$ separately and find $p = \hat{\omega}_\mu(X^*) = 80$ at X^*=(−0.0074, −0.0451, −0.2507), and $q = \hat{\omega}_\sigma(X^*) =$ 3.04301 at X^*= (0.0644, −0.2806, −0.2908). Letting λ evenly increase from 0 to 1 at a step of 0.01, we can solve for the optimal solution X^* in Eq. (11) for each λ with Mathematica. Using these X^*s, we plot the efficient curve $((\hat{\omega}_\mu(X^*), \hat{\omega}_\sigma(X^*))$ in the graph of $\hat{\omega}_\mu$vs.$\hat{\omega}_\sigma$. Note that the constraints in Eq. (11) result in a bended line, as shown in Fig. 3.
4. The optimal solution suggested by our method is the efficient point closest to (p, q), which turns out to be $\hat{\omega}_\mu(X^*) = 79.9813$ and $\hat{\omega}_\sigma(X^*) = 3.1490$ with the data-driven weight $\lambda = 0.95$ and the optimal setting X^*= (0.1290, −0.2848, −0.2856).

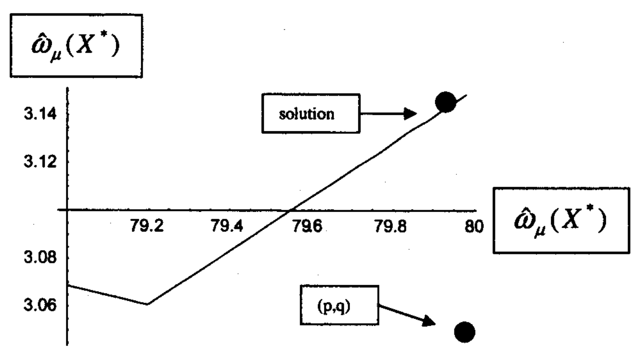

Figure 3. The plot of $\hat{\omega}_\mu(X^*)$ vs. $\hat{\omega}_\sigma(X^*)$ in example 2.

CONCLUSION

Various methods have been proposed in the literature to obtain an optimal setting for dual-response problems. This is typically done by minimizing the cost. The cost (or loss in terms of out-of-specification limits) is typically a function of (1) the deviance of mean to the target (bias) and (2) variance. The concept that any optimal solution must fall on the efficient curve is proposed here. As previously discussed, some existing methods may end up with their final solution away from the efficient curve. This is not desirable at all.

If the relative importance between bias and variance is known, either from the past experience or as a priori, the value of λ can be determined accordingly. On the other hand, if the relative importance between bias and variance is not so clear, here we propose the idea of a data-driven weight, following the generalized distance idea of Khuri and Colon (1981) in a two-dimensional case. The proposed method is easy to implement and interpret. Compared to other existing methods, it can be adjusted to achieve a good balance between bias and variance by selecting different weights. We have shown that other existing methods are either special cases of the proposed method with a specific prior weight or inferior in the sense that a better solution exists.

Further extensions of the proposed method are considered below; while Euclidean distance is used here, we can use other distance measures in general. Any meaningful distance measure can be and should be used in practice. It is also worthwhile to note that we used the response surfaces derived from designs with replicates in both of the two examples. The proposed method can also be applied in situations in which the response functions were estimated based on inner-outer array of combined array designs.

We have focused on the case of "the target value is the best" in the article so far. For the case of "the larger the better," we can modify the formulation of Eq. (9) to:

$$\begin{aligned} &\underset{x}{\text{Minimize}}\ \text{WMSE} = -\lambda\hat{\omega}_\mu(x)^2 + (1-\lambda)\hat{\omega}_\sigma(x)^2 \\ &\text{subject to } X \in \Omega \\ &\text{where } \lambda \in [0, 1] \end{aligned} \tag{12}$$

For the case of "the smaller the better," since the mean of the response in application is usually positive, this case will be equivalent to letting $T = 0$ in Eq. (9). We have:

$$\begin{aligned} &\underset{x}{\text{Minimize}}\ \text{WMSE} = \lambda\hat{\omega}_\mu(x)^2 + (1-\lambda)\hat{\omega}_\sigma(x)^2 \\ &\text{subject to } X \in \Omega \\ &\text{where } \lambda \in [0, 1] \end{aligned} \tag{13}$$

All computations in this article can be easily performed by Mathematica (Culioli and Skudlarek, 2000). Computer codes for all the computations involved are available from the authors.

ACKNOWLEDGMENTS

We thank Dr. Necip Doganaksoy for his constructive comments and helpful suggestions, which led to a significant improvement of this article. Dennis Lin is partially supported by the National Security Agency, via grant MDA904-02-1-0054.

ABOUT THE AUTHORS

Rui Ding is a Ph.D. candidate at the Department of Statistics, the Pennsylvania State University, University Park. He received his B.S. in applied mathematics from Tsinghua University (P.R. China) in 1998. His research interests are response surface methodology, experiment design, and empirical financial analysis.

Dennis K. J. Lin is a professor of management science and statistics at the Pennsylvania State University, University Park. He received a Ph.D. in statistics from the University of Wisconsin at Madison. He is an elected Fellow of the American Statistical Association, an elected member of the International Statistical Institute, and a senior member of the American Society for Quality. He has published over 100 articles in the areas of design of experiments, response surface methodology, reliability, control charts, data mining, and statistical inference.

Duan Wei is the director of the Decision Science Research Center in National Sun Yat-Sen University (Taiwan), the Convener of Finance Monetary Division in National Policy Foundation, and currently serves as the president of the Chung-Hua Data Mining Society. He received a Ph.D. in mathematics from the University of South Carolina. He is a fellow of the International Statistical Institute and has published more than 60 research articles. His research areas include public finance, quality management, financial policy, data mining, and leadership.

REFERENCES

Box, G. E. P., Draper, N. R. (1987). *Empirical Model Building and Response Surfaces*. New York, NY: John Wiley & Sons.

Copeland, K., Nelson, P. (1996). Dual response optimization via direct function minimization. *Journal of Quality Technology* 28:331–336.

Culioli, J. C., Skudlarek, J. P. (2000). Multiplier Method for Mathematica at http://library.wolfram.com/mathgroup.

Khuri, A., Colon, M. (1981). Simultaneous optimization of multiple responses represented by polynomial regression functions. *Technometrics*. 23:363–375.

Kim, K., Lin, D. K. J. (1998). Dual response surface optimization: a fuzzy modeling approach. *Journal of Quality Technology* 30:1–10.

Lin, D. K. J., Tu, W. (1995). Dual response surface optimization. *Journal of Quality Technology* 27:34–39.

Luner, J. (1994). Achieving continuous improvement with the dual response approach: a demonstration of the roman catapult. *Quality Engineering* 6:691–705.

Miettinen, K. (1999). *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers.

Phadke, M. (1989). *Quality Engineering Using Robust Design*. London: Prentice-Hall.

Taguchi, G. (1986). *Introduction to Quality Engineering: Designing Quality into Products and Processes*. White Plains, NY: Kraus International Publications.

Vining, G. G., Myers, R. H. (1990). Combining Taguchi and response surface philosophies: a dual response approach. *Journal of Quality Technology* 22:38–45.