A Weighted Loss Function Approach to the Multivariate RPD Problem

Chong Pan, Yi-Zhong Ma
Department of Management Engineering, Nanjing University of Science and Technology, Nanjing, China

Abstract - **As products/processes often possess several quality characteristics, Robust Parameter Design (RPD) problems are likely to be dealt with by considering multiple responses. Several researches have proposed solutions to this problem by using loss function. A weighted loss function approach to make a good trade-off between bias and variation components of the objective function built on these researches is suggested in this paper. In addition, an algorithm of the weight parameter is also proposed, which can be used when the prior information is vague. An example from the literature is given to illustrate the proposed method, and a numerical comparative analysis of the results from different articles is also provided.**

Keywords - **Robust parameter design, loss function, efficiency curve**

I. INTRODUCTION

As a useful method to improve the quality of products/processes, Robust Parameter Design(RPD) has been popularized by the Japanese engineer and statistician G. Taguchi. This method is powerful and effective in helping manufacturers to design their products and processes as well as to solve troublesome quality problems, ultimately leading to higher customer satisfaction and operational performance [1]. The objective of RPD is to select the setting of control factors which provides the best performance and the least sensitivity to noise [2]. The traditional RPD method only focused on a single quality characteristic to optimize the parameter conditions. However, for most products/processes, quality is multidimensional, how to improve several quality characteristics simultaneously based on the methodology of RPD is an important issue to discuss.

In dealing with a product or process with multiple quality characteristics/responses, it is difficult to determine the optimal parameter conditions for all quality characteristics as improving one particular quality characteristic may lead to serious degradation of the other critical quality characteristics [1], especially if there is correlation between characteristics. A common strategy to optimize multiple characteristics is to combine several objective functions into a single objective function. The methodology of loss function raised by Taguchi provides a framework to achieve this and to incorporate the variance-covariance structure of the responses as well as the process economics.

Several methods using the multivariate loss function have been proposed in the literature. Pignatiello [3] derived the expression of expected loss function to minimize the function for the multiple quality characteristics. Wu and Chyu [1] presented an loss function approach to optimizing the correlated multiple quality characteristics with asymmetric loss function by a mathematical programming model. Ko. et al [4] proposed a new form of multivariate loss function, which is composed of the loss of bias, poor robustness and poor quality of predictions. However, these methods were not concerned with that whether the loss of bias and poor robustness should have different weights.

The purpose of this paper is to develop a data-driven weighted loss function method, building on the existing methods. That is to say, we add weights on the terms of the loss function, and adjust the weight parameter, so that the trade-off between the reduction of bias and improvement of robustness can be explicitly incorporated.

This paper is organized as follows: at first, the existing multivariate loss function without weights is introduced. This is followed by representing the weighted loss function and the data-driven method to determine the weight parameter of the objective function. Then the implementation of the method will be illustrated by an example, and the effectiveness of our approach will be shown by comparing the results separately derived by the loss function method, the proportion of conformance method and the proposed method. Finally our conclusion will be given.

II. MULTIVARIATE LOSS FUNCTION APPROACH

Suppose a product/process possesses p quality characteristics, let $\mathbf{y}=[y_1, y_2, ...y_p]'$ be a vector representing the p quality characteristics. Then its target vector is $\boldsymbol{\tau}=[\tau_1, \tau_2, ...\tau_p]'$ (for any LTB(the Larger The Better) characteristic y_i , y_i should be transformed to $\tilde{y}_i = 1/y_i$). The quality characteristics $\mathbf{y}$ are affected by k control factors, which are represented as the vector $\mathbf{x}=[x_1, x_2, ..., x_k]'$, the experimental region of $\mathbf{x}$ is Ω . For a fixed $\mathbf{x}$, the random vector $\mathbf{y}$ has an expectation vector

$\boldsymbol{\mu}(\mathbf{x})$ and a covariance matrix $\boldsymbol{\Sigma}(\mathbf{x})$.

The overall quality of the product/process can be seen in terms of the loss results from the deviation of quality characteristics from their target values. For the case of single response, the loss is often measured by the quadratic loss function $loss(y) = c(y-\tau)^2$, where c is a constant which is determined by the repairing and scraping costs. We assume that given deviation equals to δ and the corresponding amount of quality loss is A, the positive constant $c = A/\delta^2$. The aim of RPD is to find the parameter conditions which minimize the value of risk function $R(x) = E(loss(y)) = c[Var(y) + (\mu-\tau)^2]$.

Pignatiello proposed an extension of the risk function for the case of multi-response optimization in [3]:

$$R(x) = E(L(\mathbf{y})) = trace(\mathbf{C}\boldsymbol{\Sigma}(\mathbf{x})) + (\boldsymbol{\mu}(\mathbf{x}) - \boldsymbol{\tau})'\mathbf{C}(\boldsymbol{\mu}(\mathbf{x}) - \boldsymbol{\tau}) \tag{1}$$

Where $\mathbf{C}$ is a $p \times p$ dimensional symmetric matrix, the diagonal elements c_{ii} of $\mathbf{C}$ represent the loss coefficient of quality characteristic y_i, and the off-diagonal elements c_{ij} represent the correlated loss coefficient of quality characteristic y_i and y_j.

He also suggested a response modeling strategy based on Response Surface Methodology (RSM) for the purpose of relating the expected values of the responses with the parameter conditions.

We commonly build up the regression models as (2)-(4) for the elements of $\boldsymbol{\mu}(\mathbf{x})$ and $\boldsymbol{\Sigma}(\mathbf{x})$:

$$\mu_i = \tilde{\mathbf{x}}^{(m_1)}\boldsymbol{\alpha}_i, i = 1, 2, ..., p \tag{2}$$

$$\log(\sigma_i^2) = \tilde{\mathbf{x}}^{(m_2)}\boldsymbol{\beta}_i, i = 1, 2, ..., p \tag{3}$$

$$\tanh^{-1}(\rho_{ij}) = \tilde{\mathbf{x}}^{(m_3)}\boldsymbol{\gamma}_{ij}, 1 \le i < j \le p \tag{4}$$

Where $\tanh^{-1}(\rho) = 1/2 \times \log[(1+\rho)/(1-\rho)]$, $\tilde{\mathbf{x}}^{(m_\xi)}$ $(\xi = 1, 2, 3)$ is a row vector of m_ξ covariates which denotes the factors (e.g., 1, x_1, $x_1 x_2$) that have significant effect on the response.

In order to get the least square estimators of the coefficient vectors $\hat{\boldsymbol{\alpha}}_i$, $\hat{\boldsymbol{\beta}}_i$ and $\hat{\boldsymbol{\gamma}}_{ij}$, we calculate the point estimators of μ_i, σ_i^2 and ρ_{ij} as follows:

$$\hat{\mu}_i(\mathbf{x}_h) = \overline{y_i(\mathbf{x}_h)} = \frac{\sum_{m=1}^{n} y_{im}(\mathbf{x}_h)}{n} \tag{5}$$

$$\hat{\sigma}_i^2(\mathbf{x}_h) = \frac{\sum_{m=1}^{n}\left(y_{im}(\mathbf{x}_h) - \hat{\mu}_i(\mathbf{x}_h)\right)^2}{n-1} \tag{6}$$

$$\hat{\rho}_{ij}(\mathbf{x}_h) = \frac{\sum_{m=1}^{n}\sum_{l=1}^{n}(y_{im}(\mathbf{x}_h) - \hat{\mu}_i(\mathbf{x}_h))(y_{il}(\mathbf{x}_h) - \hat{\mu}_i(\mathbf{x}_h))}{\sqrt{\sum_{l=1}^{n}(y_{im}(\mathbf{x}_h) - \hat{\mu}_i(\mathbf{x}_h))^2 \sum_{l=1}^{n}(y_{il}(\mathbf{x}_h) - \hat{\mu}_i(\mathbf{x}_h))^2}} \tag{7}$$

$1 \le i < j \le p; h = 1, 2, ..., v$

Where v denotes how many runs the experiment takes, n denotes how many replications of each run.

After finishing the regression modeling and fitting, we substitute the elements μ_i, σ_i^2 and ρ_{ij} of $\boldsymbol{\mu}(\mathbf{x})$ and $\boldsymbol{\Sigma}(\mathbf{x})$ in (1) with the estimated models:

$$\hat{\mu}_i = \tilde{\mathbf{x}}^{(m_1)}\hat{\boldsymbol{\alpha}}_i, i = 1, 2, ..., p \tag{8}$$

$$\hat{\sigma}_i^2 = \exp(\tilde{\mathbf{x}}^{(m_2)}\hat{\boldsymbol{\beta}}_i), i = 1, 2, ..., p \tag{9}$$

$$\hat{\rho}_{ij} = \frac{2\exp(\tilde{\mathbf{x}}^{(m_3)}\hat{\boldsymbol{\gamma}}_{ij}) - 1}{2\exp(\tilde{\mathbf{x}}^{(m_3)}\hat{\boldsymbol{\gamma}}_{ij}) + 1}, 1 \le i < j \le p \tag{10}$$

And get:

$$\hat{R}(\mathbf{x}) = trace(C\hat{\boldsymbol{\Sigma}}(\mathbf{x})) + (\hat{\boldsymbol{\mu}}(\mathbf{x}) - \boldsymbol{\tau})' C(\hat{\boldsymbol{\mu}}(\mathbf{x}) - \boldsymbol{\tau}) \tag{11}$$

The two terms of (11) can be referred to as $\hat{L}_r$ 和 $\hat{L}_b$. L_r represents the expected loss of poor robustness which dues to the variation of the responses, and L_b represents the expected loss of bias which dues to the deviation of mean from the targets. Then (1) can be simplified as:

$$E(L) = L_r + L_b \tag{12}$$

Where $L_r = trace(\mathbf{C}\boldsymbol{\Sigma}(\mathbf{x}))$ and $L_b = (\boldsymbol{\mu}(\mathbf{x}) - \boldsymbol{\tau})'\mathbf{C}(\boldsymbol{\mu}(\mathbf{x}) - \mathbf{T})$.

III. WEIGHTED MULTIVARIATE LOSS FUNCTION BASED ON EXPERIMENTAL DATA

The existing loss function method mentioned above does not consider the trade-off between bias and variation. In different circumstances, decision makers will have different preference for the closeness of the mean to the target and the improvement of robustness. In case of univariate response optimization, Box and Jones [5] introduced the following general class of loss function:

$$E(L) = c[\phi Var(y) + (1-\phi)(E(y) - \tau)^2] \tag{13}$$

Where ϕ is the weight parameter, and $0 \le \phi \le 1$. So as an extension, we defined a new weighted multivariate loss function as follow:

$$E(L) = c[\phi L_r + (1-\phi)L_b] \tag{14}$$

Where $L_r = trace(\mathbf{C}\boldsymbol{\Sigma}(\mathbf{x})) \ge 0$ and $L_b = (\boldsymbol{\mu}(\mathbf{x}) - \boldsymbol{\tau})'\mathbf{C}(\boldsymbol{\mu}(\mathbf{x}) - \boldsymbol{\tau}) \ge 0$.

In case of the precondition that the decision makers have a very clear idea of the suitable value of ϕ, the value of ϕ decided by them is often fairly appropriate. However, the information about ϕ is usually unavailable. Here we provide a data-driven weighting approach, which is also mentioned in [6].

Given a specified value of ϕ, the optimal setting of the control factor $\mathbf{x}^*$ can be obtained by nonlinear programming optimization:

$$\begin{aligned} &\underset{\mathbf{x}}{\text{Min}}\ z = \phi \times trace(\mathbf{C}\hat{\boldsymbol{\Sigma}}(\mathbf{x})) \\ &+ (1-\phi)(\hat{\boldsymbol{\mu}}(\mathbf{x}) - \mathbf{T})'\mathbf{C}(\hat{\boldsymbol{\mu}}(\mathbf{x}) - \mathbf{T}) \\ &\text{st} : \mathbf{x} \in \Omega \end{aligned} \tag{15}$$

Then we can consider $\mathbf{x}^*$ as a function of ϕ. We can plot the point $(\hat{L}_b(\mathbf{x}^*), \hat{L}_r(\mathbf{x}^*))$ which called an efficiency point for every fixed value of ϕ. If we let the value of ϕ vary from 0 to 1, we will get a curve of $\hat{L}_r(\mathbf{x}^*)$ versus $\hat{L}_b(\mathbf{x}^*)$. This is usually called an efficiency curve.

When $\phi = 0$, we can get a marginal optimization of $\hat{L}_b$, suppose it equal to θ. When $\phi = 1$, we get the marginal optimization of $\hat{L}_r$ that equals to ρ. Then the point $D(\theta, \rho)$ is called ideal point, it represents the best optimization which is possible. However, this is typically unobtainable. What we could do is to find the point *O* on the efficiency curve which is closest to the ideal point *D*. We can obtain the corresponding value of ϕ, and the optimal solution of the control factors sequentially.

To sum up the above, the operational steps are:

(1) Apply the original experimental data to calculate the point estimators of μ_i, σ_i^2 and ρ_{ij};

(2) Fit the regression models for the elements of $\boldsymbol{\mu}(\mathbf{x})$ and $\boldsymbol{\Sigma}(\mathbf{x})$ by least square method;

(3) Set up the optimization model with weight parameter ϕ for the problem.

(4) Plot the efficiency curve and the ideal point.

(5) Find the solution point on efficiency curve, which is closest to the ideal point.

(6) Calculate the corresponding value of ϕ and the optimal solution of the control factors.

IV. EXAMPLE

This case has been studied in [1, 7]. The experiment described a PECVD (plasma-enhanced chemical enhanced vapor deposition) process in the fabrication of integrated circuit. In this experiment, there are two quality characteristics to be concerned: y_1 denotes the deposition thickness (NTB type) with a target value of $1000\dot{A}$ and specifications of $1000 \pm 50\dot{A}$; y_2 denotes the refractive index (NTB type) with a target value of 2 and specifications of 2 ± 0.1. There are eight control factors and their levels are listed in Table Ⅰ. It is suggested that readers can refer to [1] or [7] for the experimental arrangement and experimental data.

We employed Minitab, to obtain the coefficient vectors of the polynomial regressions and the sample determination coefficients R^2 as showed in Table Ⅱ.

TABLE I
TYPE SIZES FOR CAMERA-READY PAPERS

Factor	Level(-1)	Level(0)	Level(1)
1 Cleaning method	No	Yes[a]	
2 The chamber temperature	100° C^a	200° C	300° C
3 Number of runs after cleaning	1st	2nd[a]	3rd
4 The flow rate of SiH_4	[a]6%	7%[a]	8%
5 The flow rate of N_2	30%	35%[a]	40%
6 The chamber pressure	160*mtorr*	190*mtorr*[a]	220*mtorr*
7 R.F.power	30*watt*	35*watt*[a]	40*watt*
8 Deposition time	11:5*min*	2:5*min*[a]	13:5*min*

Note: [a] denotes the initial level

TABLE II
COEFFICIENT VECTORS AND DETERMINATION COEFFICIENTS

$\tilde{\mathbf{x}}^{(15)}$	$\hat{\boldsymbol{\alpha}}_1$	$\hat{\boldsymbol{\alpha}}_2$	$\hat{\boldsymbol{\beta}}_1$	$\hat{\boldsymbol{\beta}}_2$	$\hat{\boldsymbol{\gamma}}$
1	2.9205	0.3230	3.7969	-2.7295	-0.3633
x_1	-0.0069	0.0455	0.0105	-0.9715	0.0652
x_2	0.0091	-0.0628	0.8839	0.8042	-0.0382
x_3	-0.0088	-0.0098	-0.0549	0.2593	-0.1571
x_4	0.0673	0.0609	-0.6326	-0.6613	-0.0364
x_5	-0.0503	0.0340	-0.3647	-0.2372	0.2403
x_6	0.0860	0.0231	0.3826	0.6669	0.1868
x_7	0.0726	-0.0338	0.4976	-0.0622	-0.2614
x_8	0.0474	-0.0200	0.0130	0.1421	-0.5095
x_3x_4	-0.0247	0.0376	-0.0971	0.4464	-0.5810
x_3x_5	-0.0785	0.0454	-0.2650	-1.4385	0.1504
x_3x_8	0.1543	-0.0851	1.4269	1.1917	0.1699
x_4x_5	-0.0223	-0.0144	0.1146	0.4275	-0.5600
x_4x_8	0.0028	-0.0166	0.3635	0.5043	-0.0333
x_5x_8	-0.0382	0.0542	-0.8106	-0.6534	-0.0681
R^2	0.948	0.983	0.898	0.938	0.966

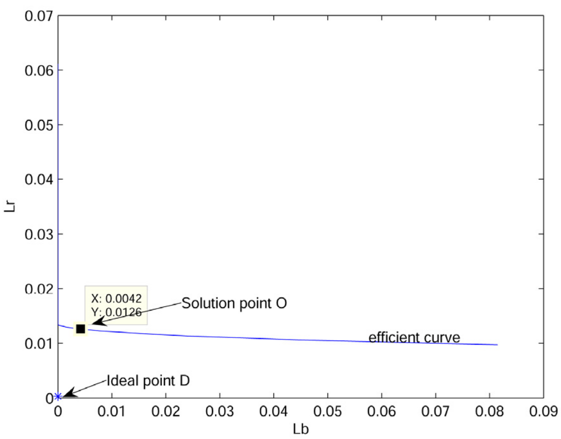

Fig. 1. Efficiency curve and the solution point

For the purpose of illustrating that the proposed procedure can manage to consider both the process economics and the correlation structure, and to keep the consistency of the data, according to some conclusions of [1], we assume the loss coefficients: c_{11}= 0.0000148, c_{22}= 5.42611; and the correlated loss coefficients c_{12} = c_{21} = 0.00294.

Then we can plot the efficiency curve and obtain the ideal point $D(0, 0.0003)$, as showed in Fig. 1. Finally, we

TABLE III
COMPARISON

	Loss function method	proportion of conformance method	Weighted loss function method
$\hat{\mu}_1$	1007.14	1561.30	984.78
$\hat{\mu}_2$	2.08	1.50	1.99
$\hat{\sigma}_1^2$	2535.7	46153	765.54
$\hat{\sigma}_2^2$	0.001	0.002	0.0001
$\hat{L}_b$	0.0360	4.2848	0.0042
$\hat{L}_r$	0.0382	0.6787	0.0126
$\hat{E}(L)$	0.0742	4.9635	0.0168

find the solution point $O(0.0042, 0.0126)$ on the efficiency curve, which is the closest to the point D. The corresponding value of ϕ is 0.9.

Solved by Matlab, we get the solution setting of control factor: $\mathbf{x}^* = [-1, -1, 1, -0.7439, 1, -0.3931, 0, 1]'$. The optimal solutions generated by the loss function method in [1] and the proportion of conformance method in [7] are $[0, -1, -1, -1, 1, -1, 1, 0]'$ and $[0, -1, 1, -1, 0, 0, 0, 1]'$. Employing the three sets of data, we separately calculate the comparison results of proposed model and two other methods rose in [1, 7] are summarized in Table Ⅲ.

Concluding from Table 3, the overall results indicate that our method is more reasonable than the methods raised in [1, 7], although the value of $\hat{\mu}_1$ obtained by the method of [1] is closer to the target than our result.

V. CONCLUSION

We propose a weighted loss function method in this paper to solve the multivariate RPD problem. In case that no priori experience can clarify the relative importance between bias and variation, we further suggest a data-driven approach to determine the weight parameter. Our method allows for the explicit consideration of the trade-off between meeting two targets: no bias and maintaining robustness.

Through an example, we show the effectiveness of the proposed method that it can simultaneously optimize correlated multiple quality characteristics, and the flexibility which makes it possible to obtain good balance of bias and variation.

REFERENCES

[1] F.-C. W and C.-C. Chyu, "Optimization of robust design for multiple quality characteristics," *International Journal of Production Research*, vol. 42, no. 2, pp. 337-354, 2004.

[2] K.-L. Tsui, "Robust design optimization for multiple characteristic problems," *International Journal of Production Research*, vol. 37, no. 2, pp. 433-445, 1999.

[3] J.J. Pignatiello, "Strategies for robust multiresponse quality engineering," *IIE Trans*, vol.25, no. 3, pp.5-15, 1993.

[4] Y. Kod, K. Kim, and C. Jun, "A new loss function-based method for multiresponse optimization," *Journal of Quality Technology*, vol. 37, no. 1, pp. 50-59, 2005.

[5] G. Box, S. Jones, "Designs for minimizing the effects of environmental variables," Technical report, Center for Quality and Productivity Improvement, University of Wisconsin, 1990.

[6] R. Ding, D.K.J. Lin, D. Wei, "Dual-response surface optimization: a weighted MSE approach," *Quality engineering*, vol. 16, no. 3, pp. 377-385, 2004.

[7] C. Chiao and M. Hamada, "Analyzing experiments with correlated multiple responses," *Journal of Quality Technology*, vol. 33, no. 4, pp. 451-465, 2001.