Modeling Approaches for Dynamic Robust Design Experiment

Suk Joo Bae

Hanyang UniversitySeoul, Korea

ABSTRACT

In general, there are three kinds of methods in analyzing dynamic robust design experiment: loss model approach, response function approach, and response model approach. In this talk, we review the three modeling approaches in terms of several criteria in comparison. This talk also generalizes the response model approach based on a generalized linear model. We develop a generalized two-step optimization procedure to substantially reduce the process variance by dampening the effect of both explicit and hidden noise variables. The proposed method provides more reliable results through iterative modeling of the residuals from the fitted response model. The method is compared with three existing approaches in practical examples.

Keywords: Generalized linear models (GLMs), Response model (RM) approach, Taguchi method, Two-step optimization.

1. INTRODUCTION

The robust design introduced by Taguchi (1987) is a strategic method for improving the performance of the system in the development stage. The main objective of robust design is to reduce the performance variation in products and processes by selecting the setting of easy-to-control factors robust to hard-tocontrol factors (noises). The robust design method has been applied to problems with the static system and the dynamic system. The static system is defined as that for which the desired output of the system has a fixed target value and the dynamic system is that where the target value depends on the input signal controlled by a system operator.

For the dynamic system, an ideal quality is based on an ideal relationship between the signal and the response. a quality loss is caused by the deviation from this ideal relationship. Significant quality improvement can be achieved through using designed experiments to search for an "optimal" design which minimizes the deviation from this function. The derived optimal solutions, however, might be different according to modeling approaches in the dynamic system.

Tsui (1998, 1999, 2000) illustrated those facts thoroughly by comparing the optimization procedure derived from three different approaches: the *response model* (*RM*), *loss model* (*LM*), and *response function model* (*RFM*) approach. He also showed that the RM approach allows greater flexibility to investigate factor effects for dynamic robust design experiment. Up to this point, most response models have been solved under the assumption that the residual error is normally distributed with constant variance in the dynamic system. Ordinary least squares (OLS) estimates are obtained from such an assumption and successfully applied in most applications. However, when the variance is a function of the mean as in the exponential family (*i.e.*, binomial, poisson, and gamma distribution, *etc.*), the inference procedure based on OLS is inaccurate. Introducing a Generalized Linear Model (GLM), the assumption of normality and constant variance for the residual errors is no longer required. In this article, we introduce a GLM to the dynamic robust design experiment for the purpose of reducing the variation caused by the noise variables which are not included explicitly in the experiment.

2. RESEARCH MOTIVATION

Lin and Wen (1994) applied the dynamic robust experimental design to obtain the uniform zinc phosphate coating. Eight control factors and a signal factor (geometric area of low-carbon steel plate) were adopted to examine the uniformity for a phosphating process. The response was the difference in the weight of the phosphate coating before and after stripping. The noise factor was a plated film location where the difference in coating uniformity was present. It cannot be controlled nor observed during a process, therefore, (explicit) noise factors were not given in this experiment. A standard OA18 orthogonal array was used as an experimental design. The objective of this experiment is to determine the best control factor settings that give the uniform plating film thickness, regardless of the unobservable noise variable (plated film location). The ideal relationship between response and signal is assumed to be linear. Figure 1 shows strong linear relationship between signal and response, along with the trend in which response variation is proportional to the signal value.

In the LM approach, the response was first fitted to the signal only for each run of fixed control combinations. Because intercept was turned out to be insignificant, a linear model without intercept was fitted to the data. The effects for the slope and error variance were estimated using ordinary least squares (OLS) at each run. All control factors except one (*A*) have three levels. We decomposed seven factors' (*B* - *H*) effects into orthogonal linear (*l*) and quadratic (*q*) contrast {-1,*0*,*1*} and {1,2,*1*}, respectively. The significant control effects on the slope (β) and log variance ($\log \sigma_L^2$) were identified based on a half-normal probability plots and Lenth's (1989) method. At the 5 % significance level the fitted models were

$$\begin{aligned} \hat{\beta} &= 0.684 + 0.224D_l - 0.122G_l - 0.102H_l \\ &\quad -0.058B_q - 0.062E_q - 0.060H_q, \\ \widehat{\log \sigma_L^2} &= 2.153 - 0.826E_l. \end{aligned}$$

Because explicit noise factors are not given in this experiment, control by noise interactions cannot be estimable. Consequently, the optimization procedure of the LM approach is identical to that of the RFM approach.

With the RM approach, all terms at the 5% significant level were included in the following OLS model:

$$\begin{aligned} \hat{y} &= (0.684 + 0.049B_l + 0.075C_l + 0.224D_l \\ &\quad +0.053E_l - 0.122G_l - 0.102H_l - 0.058B_q \\ &\quad +0.014D_q - 0.062E_q - 0.031F_q + 0.036G_q \\ &\quad -0.060H_q)M, \qquad (1) \end{aligned}$$

with $R^2 = 0.993$. We should first minimize the process variance on the basis of the fitted model (1). However, because noise factors are not included explicitly in this experiment variance effects cannot be identified using the traditional RM approach.

3. GENERALIZED LINEAR MODEL (GLM) - RESPONSE MODEL (RM) APPROACH

3.1. Generalized Two-Step Optimization Procedure

Suppose that the true response model is linear in some functions of *p* control factors (*C*), *q* explicit noise factors (*N*) and signal factor *M* of *K* levels, and also linear in *n* known functions of the signal factor ($f_1(M),\ldots,f_n(M)$). All these factors and functions interact with each other. When there are *I* control runs, *J* noise runs, and *L* replications in each one of the signal levels, the response model for $i = 1,\ldots,I$, $j = 1,\ldots,J$, $k = 1,\ldots,K$, and $l = 1,\ldots,L$, is

$$\begin{aligned} Y_{ijkl} &= \beta_0(C_i) + e_0(C_i, N_j) + [\beta_1(C_i) \\ &\quad + e_1(C_i, N_j)] f_1(M_k) + \cdots + [\beta_n(C_i) \\ &\quad + e_n(C_i, N_j)] f_n(M_k) + \varepsilon_{ijkl}, \qquad (2) \end{aligned}$$

where $\beta(\mathbf{C}_i) = (\beta_0(\mathbf{C}_i),\ldots,\beta_n(\mathbf{C}_i))$ represents the functional relationship between the response and *i-th* control factors' combination $\mathbf{C}_i$, $e(\mathbf{C}_i,\mathbf{N}_j) = (e_0(\mathbf{C}_i,\mathbf{N}_j),\ldots,e_n(\mathbf{C}_i,\mathbf{N}_j))$ denotes the relationship between the response and *j-th* noise factors' combination $\mathbf{N}_j$ only or control-by-noise interactions. It follows that $\mathbf{e}(\mathbf{C},\mathbf{N})$ are *iid* $N(\mathbf{0},\sigma_e^2(\mathbf{C}))$, where $\sigma_e^2(\mathbf{C}) = Var\,[\mathbf{e}(\mathbf{C},\mathbf{N})]$. The errors ($\varepsilon_{ijkl}$'s) are *iid* $N(\mathbf{0},\sigma_\varepsilon^2(\mathbf{C},\mathbf{f}(M)))$, where $\sigma_\varepsilon^2(\mathbf{C},\mathbf{f}(M)) = Var[\mathbf{e}(\mathbf{C},\mathbf{f}(M))]$. We assume that ***e*** and *ε* are independent. $\sigma_e^2(\mathbf{C})$ represents the variation caused by the explicit noise factor *N* and $\sigma_\varepsilon^2(\mathbf{C},\mathbf{f}(M))$ is the variation caused by a hidden noise variable. The conditional mean and variance of the response *Y* given ***C***, *M* are, respectively

$$\begin{aligned} E(Y) &= \mu(C,M) = \beta(C)' f_1(M), \qquad (3) \\ Var(Y) &= Var[e(C,N)' f_1(M)] + Var(\epsilon) \\ &= f_1(M)' \sigma_e^2(C) f_1(M) + \sigma_\epsilon^2(C,M) \quad (4) \end{aligned}$$

where $\mathbf{f}_1(M) = (1, f_1(M),\ldots,f_n(M))'$. As *Var*(*Y*) depends on the signal factor *M*, we integrate over *M* to determine optimal settings of the control factors. Then the variance of the response *Y* is

$$\sigma_Y^2(C) = \int_{M_L}^{M_H} [f_1(M)'\sigma_e^2(C) f_1(M) + \sigma_\epsilon^2(C,M)]\,dM, \qquad (5)$$

where M_L, M_H denote the low and high limit of the signal, respectively. Suppose the control factors can be divided into three disjointed groups, *i.e.*, $\mathbf{C} = (\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3)$, and $\sigma_Y^2(\mathbf{C})$ is a function of C_2 and C_3. Under the constraint that the mean function must be adjusted to *t*(*M*), we minimize the average loss $R(\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3)$. That is,

$$\underset{\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3}{Min}\ R(\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3)$$

subject to

$$\mu(\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3,M) = t(M)$$

for any *M*.

A generalized two-step optimization procedure is introduced to minimize the loss function $R(\mathbf{C}_1,\mathbf{C}_2,\mathbf{C}_3)$. The procedure is as follows:

Step 1. (a) Find C_2^* that minimize $\sigma_e^2(C_2)$

(b) Find C_3^* that minimize $\sigma_\epsilon^2(C_2^*,C_3)$.

Step 2. Find C_1^* so that $\mu(C_1^*,C_2^*,C_3^*,M) = t(M)$ for any M.

3.2. Modeling with Generalized Linear Models (GLMs)

The additional variation caused by a hidden noise variable can be reduced by solving the dynamic system after modeling the errors (ε_{ijkl}) via control, explicit noise, and signal factors simultaneously. An error from the response model (2) is

$$\varepsilon_{ijkl} = Y_{ijkl} - \psi(C_i, N_j, f(M_k)),$$

where ψ is an unknown true response function. The distribution of squared error is $\varepsilon^2_{ijkl} \sim \sigma^2_{ijk}\chi^2_{l-1}$. Since ε_{ijkl} is not an observed residual, the error is approximated by the residual (e_{ijkl}) derived from the fitted response model. The squared residual model can be represented as

$$g(e^2_{ijkl}) = x(C_i, N_j, M_k)'\gamma,$$

where $\mathbf{x}(\mathbf{C}_i, \mathbf{N}_j, M_k)$ represents a predictor vector consisting of the control, explicit noise, and signal factors, γ is a coefficient vector, and g denotes a link function for the response e^2_{ijkl}. Note that the element of first column in x is 1. We assume that higher orders of signal functions are negligible, thus only the linear term is included in the residual modeling. Based on the fact that squared residuals are approximately gamma distributed and the variance must remain positive, we consider a gamma-multiplicative model with

$$e^2_{ijkl} = exp[x(C_i, N_j, M_k)'\gamma]. \qquad (6)$$

The maximum likelihood estimate (MLE), $\hat{\gamma}$ is derived from the log-gamma link function for the mean of the squared residuals. The actual implementation of maximum likelihood results in an algorithm based on iteratively reweighted least squares (IRLS) (see McCullagh and Nelder (1989)). Thus, MLE $\hat{\gamma}$ is obtained by minimizing

$$\sum_i \sum_j \sum_k \sum_l \frac{[e^2_{ijkl} - \mu(x'_{ijk}\gamma)]^2}{h[\mu(x'_{ijk}\gamma)]},$$

where $\mu(x'_{ijk}\gamma) = E[e^2_{ijkl}|\{C, N, M\}]$, and $h[\mu(x'_{ijk}\gamma)] = Var[e^2_{ijkl}|\{C, N, M\}]$.

After approximating $\hat{\sigma}^2_{ijk}$ with the squared residuals using a gamma-multiplicative model, the weighted least squares (WLS) method is used to improve the response model. We can optimize the dynamic system based on the effect estimates resulting from this iterative procedure. Hereafter, our procedure for optimizing the dynamic system will be called the *GLM-RM approach*.

4. EXAMPLES REVISITED

Zinc phosphate coating data were re-analyzed following to the iterative procedure in Section 4. First, the residuals from the fitted response model (1) were calculated, then the squared residuals were regressed on the control, signal, and control-signal interactions using a gamma-log link function. The final response and GLM variance model with the 5% significant level were, respectively

$$\begin{aligned} \hat{y} &= (0.676 + 0.044B_l + 0.073C_l + 0.208D_l \\ &\quad +0.060E_l - 0.119G_l - 0.113H_l - 0.057B_q \\ &\quad -0.060E_q - 0.032F_q + 0.037G_q - 0.048H_q)M, \\ \hat{\sigma}^2_\varepsilon &= exp[1.464 + 0.918B_l - 0.913E_q + 0.019M \\ &\quad +0.005E_qM]. \end{aligned} \qquad (7)$$

At last, we applied the generalized two-step optimization procedure to the response and variance model estimated from the GLM-RM approach. First, the variance function (7) should be integrated over M

$$\begin{aligned} \hat{\sigma}^2_Y(C) &= \int_{M_L}^{M_H} \exp[1.464 + 0.918B_l - 0.913E_q \\ &\quad +0.019M + 0.005E_qM]dM \\ &= \left(\frac{44.70\exp[0.087E_q] - 1.46\exp[-0.813E_q]}{0.019 + 0.005E_q}\right) \\ &\quad \times \exp[1.464 + 0.918B_l]. \end{aligned}$$

The optimization results from the GLM-RM approach are quite different from the LM approach. Factor E was fitted at level 1 in the LM approach, while $E = 0$ was selected to minimize the process variance in the GLM-RM approach.

5. CONCLUDING REMARKS

For the dynamic robust design problem, the RM approach allows greater flexibility to study factor effects and results in experimental cost savings. This article illustrates how we extend the response model approach more general cases: there are no explicit noise factors and there are no significant control-by-noise interactions. It was also noted that GLM modeling of the residuals from the response model reveals the potential for reducing the variance caused by hidden noise variables in the RM approach. In conclusion, we suggested the GLM-RM approach and generalized two-step optimization procedure as the tool for providing more reliable results.

REFERENCE

[2] Joseph, V. R., and Wu, C. F. J. (2002), "Robust Parameter Design of Multiple-Target Systems", Technometrics,44, 338-346.

[3] Lenth, R. V. (1989), "Quick and Easy Analysis of Unreplicated Experiments", Technometrics, 31, 469-473.

[4] Lin, S. M. and Wen, T. C. (1994), "Experimental Strategy: Application of Taguchi's Quality Engineering Method to Zinc Phosphate Coating Uniformity", Plating and Surface Finishing, 59-64.

[5] McCullagh, P., and Nelder, J. A. (1989), Generalized Linear Model, Chapman and Hall, New York.

[6] Taguchi, G. (1987), System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost, White Plains, NY: Unipub/Kraus International.

[7] Tsui, K. L. (1998), "Alternatives of Taguchi's Approach for Dynamic Robust Design Problems", International Journal of Reliability, Quality and Safety Engineering, 5, 115–131.

[8] Tsui, K. L. (1999), "Response Model Analysis of Dynamic Robust Design Experiments", IIE Transactions, 31, 115–131.

[9] Tsui, K. L. (2000), "Modeling and Analysis of Dynamic Robust Design Experiments", Frontiers in SQC, 31, 115–131.