Journal of Modern Applied Statistical Methods

Volume 16 | Issue 1 Article 11

5-1-2017

# Analysis of Robust Parameter Designs

Tak K. Mak
*Concordia University*, tak.mak@concordia.ca

Fassil Nebebe
*Concordia University*, fassil.nebebe@concordia.ca



Recommended Citation
Mak, T. K. & Nebebe, F. (2017). Analysis of robust parameter designs. Journal of Modern Applied Statistical Methods, 16(1), 179-194. doi: 10.22237/jmasm/1493597400

# Analysis of Robust Parameter Designs

**Cover Page Footnote**
This research was supported by NSERC-GRF grant (N01374) and VPRGS seed funding (VS0141) of Concordia University.

Journal of Modern Applied Statistical Methods
May 2017, Vol. 16, No. 1, 179-194.
doi: 10.22237/jmasm/1493597400

ISSN 1538 − 9472

# Analysis of Robust Parameter Designs

**Tak K. Mak**
Concordia University
Montreal, Québec

**Fassil Nebebe**
Concordia University
Montreal, Québec

The analysis of robust parameter design is discussed via a model incorporating mean-variance relationship which, when ignored as in the classical regression approach, can be problematic. The model is also capable of alleviating the difficulties of the regression approach in the search of the minimum variance occurring region.

*Keywords:* Graphical log-linear models, contingency tables, decomposable models, hierarchical log-linear models

## Introduction

As part of their efforts in quality improvement, manufacturers strive to design products that are capable of functioning optimally under a wide range of environmental conditions. Instead of using more expensive parts or components, a more cost-effective means is to look for settings of design factors that would achieve this quality robustness. Specifically, this involves finding design settings that would minimize variance while being on target. In this regard, robust parameter designs have been widely used in the industry to determine the optimal settings of these design factors (Khuri & Mukhopadhyay, 2010; Robinson, Borror, & Myers, 2004). In robust parameter designs, design of experiment techniques are used to obtain data that are subsequently analyzed to explore the relationship between the quality characteristics and the levels of the design factors (Choi & Allen, 2009).

Taguchi advocated the use of crossed array designs and suggested a convenient analysis using signal to noise ratio. However, the limitations of an analysis based on the "signal to noise" ratio have been pointed out by many researchers (Box, 1988; Barreau, Chassagnon, Kobi, & Seibilia, 1999). Various methods of analysis have been proposed in robust parameter designs. Vining and Myers (1990) suggested a dual response surface methodology in which the

*T. K. Mak is a professor in the department of Supply Chain & B. T. M. Email him at: tak.mak@concordia.ca.*

(primary) response surface of standard deviation is minimized subject to a target constraint based on a (secondary) response surface of the mean (see also Chan & Mak, 1995).

Another common approach exploits the possible existence of mean-variance relationship to achieve simpler variance minimization computationally. With this method, the variance is assumed to be a product of a function of the mean and a "Performance Measure Independent of Adjustment" (PerMIA) (Box, 1988, p. 2). The PerMIA is a function of a proper subset of design factors (control factors) and the complement of this subset constitutes the subset of adjustment factors which influence variance only through its presence in the mean. Because of this variance factorization it is possible to minimize variance through the unconstrained minimization of the PerMIA, which is then followed by the searching of the levels of the adjustment variables to attain the desired target value.

Both the dual response surface methodology and the PerMIA approach conduct an analysis based on the sample variance or standard deviation calculated from replicates at each treatment combination. For the crossed array design, the sample variances are calculated from observations in the outer arrays that are crossed with the treatment combinations or inner arrays in the experiment. The sample variances calculated from the outer arrays do not constitute estimates of variances obtained from random samples. In crossed array or combined array designs, the noise factors which occur randomly during the lifetime usage of the product are controlled and have known values in the experiment (Welch, Yu, Kang, & Sacks, 1990; Shoemaker, Tsui, & Wu, 1991; Mak & Nebebe, 2005).

Thus in the design stage, roles of the design and noise factors are indistinguishable. In the analysis of such designs, a regression function is first fitted, from which the variance function is derived with respect to the randomness of the noise factors. Variance minimization can then be conducted based on the inherent variance function from this regression modeling approach. This regression analysis is conceptually simple and exploits the quality characteristic and noise factor relationships to achieve possibly greater efficiency. However, this regression approach has two issues to be properly addressed: First, it does not accommodate the possible dependency of the variance on the mean. Mak and Nebebe (2004) demonstrated with an example that settings determined by the regression approach can yield a variance that is substantially higher than the actual attainable optimal variance. They also proposed a new model that incorporates the mean-variance relation and includes the regression model as a special case. Second, the form of the variance function is not flexible and completely determined by the formulated regression model based on the

interactions between the control and noise variables (O'Donnell & Vining, 1997). It does not permit the formulation of a simple linear relationship between the variance and the design factors and the variance model has to be at least of the second order. Unfortunately, as seen in the simulation studies in this paper, this second order variance model in the regression approach is usually inadequate. The aforementioned issues in the regression approach are addressed in this study, and some practical recommendations are made in the light of the simulation result.

## Methodology

### Modeling Mean-Variance Relationship

Denote by $y$ the quality characteristic of interest. Let $X_1,\ldots, X_p$ be $p$ design factors influencing $y$. Suppose that there are $q$ noise factors $Z_1,\ldots, Z_q$, the levels of which are controlled in the experimenting stage. Mak and Nebebe (2004) proposed the following model for analyzing robust parameter designs:

$$y = \mu_y\left(\mathbf{x},\boldsymbol{\beta}\right) + \sqrt{\mathrm{V}_\lambda\left(\mu_y\left(\mathbf{x},\boldsymbol{\beta}\right)\right)}\left(\mathrm{h}\left(\mathbf{x}_{(1)},\mathbf{z},\boldsymbol{\theta}\right)+e\right) \tag{1}$$

where $\mathbf{x} = (X_1,\ldots, X_p)'$, $\mathbf{z} = (Z_1,\ldots, Z_q)'$, $e$ is the error term with variance $\sigma_e^2$, $\mu_y(\mathbf{x}, \boldsymbol{\beta})$ is the conditional mean of $y$ given $\mathbf{x}$ (with respect to the distribution of $\mathbf{z}$ and $e$), and $\boldsymbol{\beta}$ is the vector of regression parameters. Furthermore, $\mathbf{x}_{(1)} \subseteq \mathbf{x}$ is a subset of "control factors", and $\mathrm{V}_\lambda(\mu_y)$ is a scalar function with parameter $\lambda$ specifying the dependence of the variance on the mean. It is required that $\mathrm{V}_\lambda(\mu_y) \equiv 1$ for a certain $\lambda$, say $\lambda = 0$. Because $\mathrm{E}(y) = \mu_y(\mathbf{x}, \boldsymbol{\beta})$, it follows that $\mathrm{E}(\mathrm{h}(\mathbf{x}_{(1)}, \mathbf{z}, \boldsymbol{\theta})) = 0$, where the expectation $\mathrm{E}(\cdot)$ is taken with respect to the distribution of $\mathbf{z}$. It follows from (1) that

$$\mathrm{Var}\left(y\right) = \mathrm{V}_\lambda\left(\mu_y\right)\left\{\mathrm{Var}\left(\mathrm{h}\left(\mathbf{x}_{(1)},\mathbf{z},\boldsymbol{\theta}\right)\right)+\sigma_e^2\right\}$$

where Var is the variance operator taken with respect to the distribution of $\mathbf{z}$ and $e$, $\mu_y$ is written in place of $\mu_y(\mathbf{x}, \boldsymbol{\beta})$ when there is no possibility of ambiguity. Because $\mathrm{V}_\lambda(\mu_y) \equiv 1$ when $\lambda = 0$, Mak and Nebebe's model includes the situation where there is no mean-variance relation as a special case. It is clear that the PerMIA is given by $\mathrm{Var}(\mathrm{h}(\mathbf{x}_{(1)},\mathbf{z},\boldsymbol{\theta}))+\sigma_e^2$, and $\mathbf{x}_{(2)} = \mathbf{x} \setminus \mathbf{x}_{(1)}$ is the vector of adjustment factors that influence the variance through its presence in the mean. Thus to

minimize variance around a target, one can choose levels of $\mathbf{x}_{(1)}$ to minimize the PerMIA and then adjust the levels of $\mathbf{x}_{(2)}$ to attain the desired target mean. With the use of PerMIA, only unconstrained minimization is needed and a change in the target value requires only readjustment of the levels of the adjustment factors (Box, 1988). In the next section, we give an algorithm for computing iteratively estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ given $\lambda$.

Box (1988) proposed a transformation approach for designs with replicates which can be easily extended to crossed array or combined array designs. Specifically, it is assumed that there exists a transformation $T_\lambda$ such that $T_\lambda(y) = m(\mathbf{x}, \boldsymbol{\beta}) + h(\mathbf{x}_{(1)}, \mathbf{z}, \boldsymbol{\theta}) + e$, eliminating the dependency of the variance on the mean on the transformed scale. Thus $y = T_\lambda^{-1}\left\{m\left(\mathbf{x},\boldsymbol{\beta}\right)+h\left(\mathbf{x}_{(1)},\mathbf{z},\boldsymbol{\theta}\right)+e\right\}$ and to terms of the linear order, we have approximately,

$$y = T_\lambda^{-1}\left(m\left(\mathbf{x},\boldsymbol{\beta}\right)\right)+\left.\frac{d\left(T_\lambda^{-1}\left(u\right)\right)}{du}\right|_{u=m\left(\mathbf{x},\boldsymbol{\beta}\right)}\left\{h\left(\mathbf{x}_{(1)},\mathbf{z},\boldsymbol{\theta}\right)+e\right\}$$

which is in the form of (1). Thus we have seen that Mak and Nebebe's model is approximately equivalent to Box's (1988) transformation approach. The analysis conducted on the transformed scale using Box's approach has to be followed by an "aim-off" analysis in order that variance be minimized on the original metric.

### Determining $\lambda$ and the identification of $\mathbf{x}_{(1)}$

From (1), if

$$y_* = \frac{y-\mu_y}{\sqrt{V_\lambda\left(\mu_y\right)}} = h\left(\mathbf{x}_{(1)},\mathbf{z},\boldsymbol{\theta}\right)+e \tag{2}$$

then there is an ordinary regression model with homogeneous errors. However, if the $\lambda$ used on the left is different from the true value of $\lambda$, say $\lambda_0$, then

$$y_* = \frac{y-\mu_y}{\sqrt{V_\lambda\left(\mu_y\right)}} = \frac{\sqrt{V_{\lambda_0}\left(\mu_y\right)}}{\sqrt{V_\lambda\left(\mu_y\right)}}h\left(\mathbf{x}_{(1)},\mathbf{z},\boldsymbol{\theta}\right)+\frac{\sqrt{V_{\lambda_0}\left(\mu_y\right)}}{\sqrt{V_\lambda\left(\mu_y\right)}}e$$

**Table 1.** Simulated data and mean, variance calculations

| | | | | $z_1$ | -1 | 1 | 1 | -1 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $z_2$ | -1 | 1 | -1 | 1 | True | | Sample | |
| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $z_3$ | -1 | -1 | 1 | 1 | mean | var | mean | var |
| -1 | 0 | 0 | 0 | | 22.2 | 25.1 | 22.5 | 28.0 | 26.2 | 2.5 | 24.5 | 5.5 |
| -1 | 1 | 1 | 1 | | 26.0 | 28.4 | 25.6 | 26.9 | 29.8 | 4.3 | 26.7 | 1.1 |
| -1 | -1 | -1 | -1 | | 19.1 | 20.1 | 17.2 | 24.9 | 20.8 | 1.0 | 20.3 | 8.1 |
| 0 | -1 | 0 | 1 | | 9.0 | 10.5 | 10.5 | 10.1 | 9.8 | 0.7 | 10.0 | 0.4 |
| 0 | 0 | 1 | -1 | | 24.5 | 36.9 | 40.0 | 22.9 | 30.2 | 62.7 | 31.1 | 56.0 |
| 0 | 1 | -1 | 0 | | 20.4 | 31.0 | 28.7 | 21.1 | 26.0 | 34.7 | 25.3 | 21.6 |
| 1 | -1 | 1 | 0 | | 25.8 | 26.9 | 24.7 | 28.7 | 28.2 | 3.4 | 26.6 | 2.2 |
| 1 | 0 | -1 | 1 | | 19.5 | 18.9 | 19.5 | 20.3 | 19.5 | 0.8 | 19.5 | 0.2 |
| 1 | 1 | 0 | -1 | | 23.6 | 25.0 | 22.6 | 28.5 | 26.1 | 2.5 | 24.9 | 4.9 |

Because $\mu_y = E(y) = \mu_y(\mathbf{x}, \boldsymbol{\beta})$ is a function of possibly all the design factors, the adjustment factors would also appear to have some influences on the variable $y_*$. Thus the relationship between $y_*$ and the adjustment variables will be zero only when the true $\lambda$ is used in (2) so that the explained variation by each of these adjustment variables should attain its minimum at a value around the true $\lambda$. This fact could be exploited to determine the value of $\lambda$ approximately. To illustrate this, consider a set of simulated data from a crossed array design involving four design factors and three noise factors. The inner and outer arrays are, respectively, L9 and L4 arrays so that the experiment consists of (9)(4) = 36 experimental runs crossed between the inner and outer arrays. For each experimental run, the quality characteristic y is simulated using the model

$$y = \mu_y + \mu_y^2\left(\left(0.0024 + 0.003u_2\left(X_1\right)\right)Z_1 + 0.001Z_2 - 0.002Z_3 + e\right)$$

where $u_2(X)$ is the quadratic orthogonal polynomial $2 - 3X$, $u_1(X) = X$ is the linear orthogonal polynomial, and $\sigma_e = 0.003$. The levels of the design and noise factors in the design and the simulated data are given in Table 1, along with the true mean $\mu_y$ of $y$ for each inner array used to simulate the data.

The mean of $y$ for each of the nine inner arrays is simply estimated by the mean of the $y$ values from the corresponding crossed outer array. $V_\lambda\left(\mu_\lambda\right) = \mu_y^\lambda$ and the true $\lambda$ is equal to 4. Note the h function can be approximated by a quadratic function in $\mathbf{z}$ but the second order terms vanish since $E(h(\mathbf{x}_{(1)}, \mathbf{z}, \boldsymbol{\theta}) = 0$. If we also retain only up to the quadratic effects of the $X_i$ on the variance, then the h function is approximately a linear combination

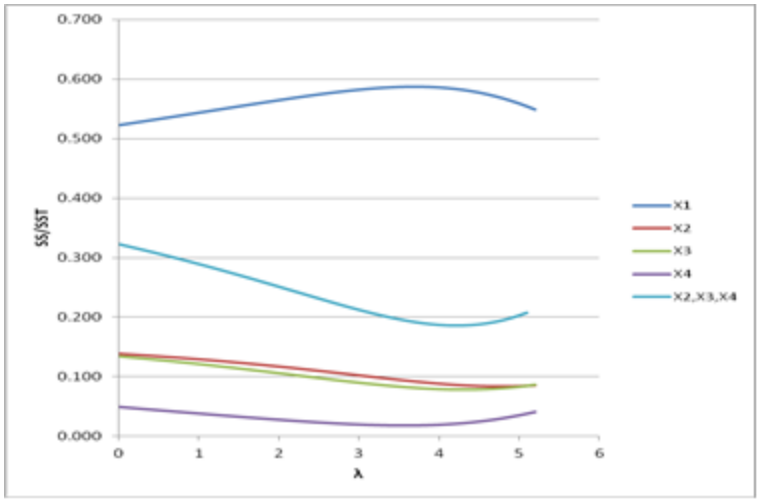


**Figure 1.** The proportion of sum of squares plot

$$h(\mathbf{x}_{(1)}, \mathbf{z}, \boldsymbol{\theta}) = \sum_{j=1}^{3} \theta_j Z_j + \sum_{i=1}^{4} \sum_{j=1}^{3} \left[ \theta_{ij1} u_1(X_i) Z_j + \theta_{ij2} u_2(X_i) Z_j \right]$$

of the terms. $Z_j$, $u_1(X_i)Z_j$, $u_2(X_i)Z_j$, $i = 1,\ldots, 4$, $j = 1, 2, 3$. Thus, even if the functional form of the h function is not exactly known, the suggested method can still be used to determine $\lambda$, though the value is only approximately unbiased. Also, as with any estimation procedure, the suggested method will not yield the true value of $\lambda$ due to the randomness of the error term $e$. In the present example, the linear combination of effects described above is used as the h function in fitting the regression model (2) in order to determine $\lambda$. For any given $\lambda$, the linear regression model (2) can therefore also be fitted to the data yielding the total sum of squares SS($X_i$) corresponding to each design factor $X_i$ (i.e. total of the sum of squares for $u_1(X_i)Z_j$, $u_2(X_i)Z_j$, $j = 1, 2, 3$), for $i = 1,\ldots, 4$. A graphical plot is presented in Figure 1 of P($X_i$) = SS($X_i$)/SST against λ for $i = 1,\ldots, 4$.

It is clear that $X_1$ is the only control variable affecting variance and $X_2$, $X_3$, and $X_4$ are adjustment variables. Furthermore, the value of P($X_i$) is smallest when $\lambda$ is equal to 4.7, 4.3, and 3.6 for $X_2$, $X_3$, and $X_4$, respectively. The proportion of sum of squares

$$\mathrm{P} = \sum_{i=1,2,3,4} \frac{\mathrm{SS}(X_i)}{\mathrm{SST}}$$

corresponding to all the adjustment variables $X_2$, $X_3$, and $X_4$ is also plotted against $\lambda$ in Figure 1. The value of P attains a minimum at $\lambda = 4.2$, which indicates that, collectively, the observed relationship between $y_*$ and the variables $X_2$, $X_3$, and $X_4$ are lowest when $\lambda$ is close to 4. Thus in the present example, the suggested method determines quite accurately the true value of $\lambda$.

## Estimation of Parameters

Consider the iterative estimation of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ for given $\lambda$. Suppose there are $n$ experimental runs in the experiment. Let $X_{1i},\ldots, X_{pi}$ and $Z_{1i},\ldots, Z_{qi}$ be the levels of the design and noise factors, respectively, in the $i^{\text{th}}$ run. Let $y_i$ be the observed quality characteristic and $\mu_{yi} = E(y_i) = \mu_y(\mathbf{x}_i, \boldsymbol{\beta})$, where now $\mathbf{x}_i = (X_{1i},\ldots, X_{pi})'$. The computing of the estimates calls an external algorithm denoted, say, by WLS($\mathbf{y}$, $\mathbf{X}$, r(·), $\mathbf{w}$, $\mathbf{p}$). The input arguments $\mathbf{y}$, $\mathbf{X}$, r, and $\mathbf{w}$ of WLS represent, respectively, the array of values of the dependent variable, the design matrix, the regression function, and the array of weights. The array $\mathbf{p}$ holds the output weighted least squares estimates of the regression parameters. The algorithm for computing the estimates $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ is given below.

Step 0. Initialize and save the starting values of $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ in the arrays $\mathbf{b}_0$ and $\mathbf{f}_0$.

Step 1. For $i = 1$ to $n$
- i. Read the values of $\mathbf{x}_i$, $\mathbf{x}_{(1)i}$, $\mathbf{z}_i$ into the $i^{\text{th}}$ row of two-dimensional arrays **XA**, **XA1**, **ZA**
- ii. Next
- iii. Read the values of $y$ into a one-dimensional array $\mathbf{y}$

Step 2. For $i = 1$ to $n$
- i. Let $m = \mu(\mathbf{x}_i, \mathbf{b}_0)$. Here $\mathbf{x}_i$ is from the $i^{\text{th}}$ row of **XA**
- ii. Let $ys(i) = \left[ y(i) - m \right] \Big/ \sqrt{V_\lambda(m)}$
- iii. Let $w(i) = 1$
- iv. Read $\mathbf{x}_{(1)i}$, $\mathbf{z}_i$ into the $i^{\text{th}}$ row of a two-dimensional array **XZA**
- v. Next

Step 3. Call WLS($\mathbf{ys}$, **XZA**, h(·), $\mathbf{w}$, $\mathbf{f}_1$). Here the regression function h is $h(\mathbf{x}_{(1)i}, \mathbf{z}_i, \boldsymbol{\theta})$

Step 4. For $i = 1$ to $n$
- i. Let $m = \mu(\mathbf{x}_i, \mathbf{b}_0)$. Here $\mathbf{x}_i$ is from the $i^{\text{th}}$ row of **XA**

ii. Let $u(i) = y(i) - \sqrt{V_\lambda(m)}\, h\left(\mathbf{x}_{(1)}, \mathbf{z}_i, \mathbf{f}_1\right)$. Here $\mathbf{x}_{(1)}$, $\mathbf{z}_i$ are from the $i^{th}$ row of **XA1**, **ZA**

iii. Let $w(i) = 1/V_\lambda(m)$

iv. Next

Step 5. Call WLS(**u**, **XA**, μ(·), **w**, $\mathbf{b}_1$). Here the regression function μ is μ($\mathbf{x}_i$, **β**)

Step 6. If $\mathbf{b}_1$ and $\mathbf{f}_1$ differ from respectively $\mathbf{b}_0$ and $\mathbf{f}_0$ by less than certain prescribed small values

i. Then

- Stop the program

ii. Else

- Let $\mathbf{b}_0 = \mathbf{b}_1$
- Let $\mathbf{f}_0 = \mathbf{f}_1$
- Go to Step 2

iii. End if

Let $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ be the final values of $\mathbf{b}_1$ and $\mathbf{f}_1$ obtained from the iterative procedure. The estimate $\left(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}}\right)$ is a solution of the system of equations in **β** and **θ**:

$$\sum \left( \frac{y_i - \mu_y\left(\mathbf{x}_i, \boldsymbol{\beta}\right)}{\sqrt{V_\lambda\left(\mu_y\left(\mathbf{x}_i, \boldsymbol{\beta}\right)\right)}} - h\left(\mathbf{x}_{(1)i}, \mathbf{z}_i, \boldsymbol{\theta}\right) \right) \frac{\partial h\left(\mathbf{x}_{(1)i}, \mathbf{z}_i, \boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}} = 0,$$

$$\sum \frac{1}{V_\lambda\left(\mu_y\left(\mathbf{x}_i, \boldsymbol{\beta}\right)\right)} \left( y_i - \sqrt{V_\lambda\left(\mu_y\left(\mathbf{x}_i, \boldsymbol{\beta}\right)\right)}\, h\left(\mathbf{x}_{(1)i}, \mathbf{z}_i, \boldsymbol{\theta}\right) - \mu_y\left(\mathbf{x}_i, \boldsymbol{\beta}\right) \right) \frac{\partial \mu_y\left(\mathbf{x}_i, \boldsymbol{\beta}\right)}{\partial \boldsymbol{\beta}} = 0$$

Because the left sides of this system of equations have zero expected values, the estimator $\left(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}}\right)$ is consistent for (**β**, **θ**).

Another method of estimation is the familiar maximum likelihood approach in which the error term is assumed to be normally distributed. The maximum likelihood estimators do not have a closed form and must be obtained numerically. One possibility is to use a generic search optimization algorithm, such as the simplex search, that does not require first or second order derivatives. It is however quite well-known in the regression literature that the estimation of the mean could be more adversely affected when the variance function is mis-specified. In this regard, the dual response surface methodology in which separate

regression models are fitted with the sample means and standard deviations calculated from the replicates (or observations from the outer arrays) may be more robust in the estimation of the mean to the misspecification of the variance component. However, for experiments with few experimental runs and a large number of factors, a quadratic response surface for the variance may not be feasible. The estimation procedure proposed in this paper may also be adversely affected by the mis-specification of the h function as it also appears in the second estimation equation for estimating the **β**.

## Results: Extensions of the Classical Regression Approach

To simplify the notation and the discussion of the classical regression (CR) approach and its possible ramification, assume the variance is functionally not dependent on the mean. In the presence of a variance-mean relationship, the analyses proposed here can be readily generalized to incorporate such a relationship using the method above. In the classical regression approach, the mean function $\mu_y(\mathbf{x}, \boldsymbol{\beta})$ is usually assumed to be quadratic in form (without cross-product terms between factors in some designs). The dependence of the variance on the control factors is typically introduced into the model by incorporating in the function $h(\mathbf{x}_{(1)}, \mathbf{z}, \boldsymbol{\theta})$ cross-product terms between the terms in the quadratic mean model and the noise factors. For example, for the orthogonal inner array L9 used in Vandenbrande's (1998) experiment, the mean function involves the saturated model:

$$\mu_y\left(\mathbf{x},\boldsymbol{\beta}\right)=\mu+\beta_{11}u_1\left(X_1\right)+\beta_{12}u_2\left(X_1\right)+\ldots+\beta_{41}u_1\left(X_4\right)+\beta_{42}u_2\left(X_4\right)$$

and

$$\begin{aligned}\mathrm{h}\left(\mathbf{x}_{(1)},\mathbf{z},\boldsymbol{\theta}\right)&=\sum_{j=1}^{3}\theta_j Z_j+\sum_{i=1}^{4}\sum_{j=1}^{3}\left[\theta_{ij1}u_1\left(X_i\right)z_j+\theta_{ij2}u_2\left(X_i\right)Z_j\right]\\&=\sum_{j=1}^{3}\left[\theta_j+\sum_{i=1}^{4}\theta_{ij1}u_1\left(X_i\right)+\theta_{ij2}u_2\left(X_i\right)\right]Z_j\end{aligned}$$

Consequently,

$$\mathrm{Var}(y) = \sigma_e^2 + \sum_{j=1}^{3}\left[\theta_j + \sum_{i=1}^{4}\theta_{ij1}u_1(X_i) + \theta_{ij2}u_2(X_i)\right]^2 \sigma_{z_j}^2$$

where, without loss of generality, the variances of $Z_1$, $Z_2$, and $Z_3$ can be taken to be unity. Thus under the classical regression model the fourth order of a control factor may be involved. If this interaction model is approximately valid and the variance function can be reasonably approximated by a quadratic function, then $\theta_{ij2} = 0$ for all $i$ and $j$ and the following reduced model can be considered:

$$\mathrm{h}\left(\mathbf{x}_{(1)}, \mathbf{z}, \boldsymbol{\theta}\right) = \sum_{j=1}^{3}\theta_j Z_j + \sum_{i=1}^{4}\sum_{j=1}^{3}\theta_{ij1}u_1(X_i)Z_j$$

The variance function for this reduced model becomes:

$$\mathrm{Var}(y) = \sigma_e^2 + \sum_{j=1}^{3}\left(\theta_j + \sum_{i=1}^{4}\theta_{ij1}X_i\right)^2$$

because $u_1(x) = x$ and $\sigma_{z_j}^2 = 1$. It is clear that a major drawback of the interaction model is that the variance can never assume the form of a linear function of the control factors and is therefore not appropriate for designs used in the steepest descent stage for locating the region containing the optimal variance solutions.

Proposed here is a generalization of the classical model to attain greater efficiency and flexibility in handling a wider range of applications. Consider the following functional form of the function h:

$$\mathrm{h}\left(\mathbf{x}_{(1)}, \mathbf{z}, \boldsymbol{\theta}\right) = \sum_{j}\mathrm{sign}\left(\psi_j\right)|\psi_j|^{\gamma} Z_j \tag{3}$$

indexed by $\gamma$, where $\psi_j$ is a function of $\mathbf{x}_{(1)}$ with a vector parameter $\boldsymbol{\theta}_{(j)}$. In most applications, we could choose $\psi_j$ to be a quadratic function of $\mathbf{x}_{(1)}$. When $\gamma = 1$ and $\psi_j$ is a quadratic function, the model clearly becomes the classical regression model. The case $\gamma = 0.5$ is also of special interest since it yields a linear or quadratic approximation to the variance function depending on whether the functional forms chosen for $\psi_j$ are linear or quadratic.

For the analysis of crossed array designs, the dual response surface method (Vining & Myers, 1990) is a serious competitor to the classical regression

approach. In the dual response surface approach, a quadratic response surface is additionally postulated with the sample variance (or standard deviation) as the response variable. It is argued that the calculation of the sample variance for each inner array using the observations from the crossed outer array is not entirely appropriate, because they do not constitute a random sample. If

$$y = \mu_y(\mathbf{x}, \boldsymbol{\beta}) + \mathrm{h}\left(\mathbf{x}_{(1)}, \mathbf{z}, \boldsymbol{\theta}\right) + e$$

for any functional form of h, it is not difficult to show that an unbiased estimator of the variance for each inner array is indeed given by

$$a^{-1}\sum(y - \bar{y})^2 - a^{-1}\sigma_\varepsilon^2$$

where $a$ is the dimension of the outer arrays (number of observations in the outer array). Note that the divisor in the first term is $a$ not $a - 1$. Thus if $\sigma_\varepsilon$ is small, an approximately unbiased estimator of the variance for each inner array is given by the simple estimator $v_i = a^{-1}\sum(y - \bar{y})^2$.

## Simulation

Simulation studies were conducted to compare the performances of the following four variance estimators discussed above:

1. *Response surface method (RSM)*. The value $v_i$ is used as the response value of the $i^{\text{th}}$ inner array. A quadratic model is then fitted to this variance response surface.
2. *The classical regression approach (CR)*.
3. *The classical regression approach leading to a quadratic variance function (CRQV)*. This model includes only the cross-product terms between the linear effects of the design factors and the noise factors in the function h, which gives rise to a quadratic variance function as explained in the previous section.
4. *The generalized regression approach (GR)* with $\gamma = 0.5$ in (3).

Again, the design studied in Mak and Nebebe (2004) and Vandenbrande (1998) were used. The means and variances used to simulate the data are given in Table 2. The true model used to generate the *y* observations is:

**Table 2.** True means and variance of the model used to simulate the data

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | True mean | True variance |
|---|---|---|---|---|---|
| -1 | 0 | 0 | 0 | 41.15 | 18 |
| -1 | 1 | 1 | 1 | 44.80 | 18 |
| -1 | -1 | -1 | -1 | 35.83 | 18 |
| 0 | -1 | 0 | 1 | 24.83 | 7 |
| 0 | 0 | 1 | -1 | 45.18 | 7 |
| 0 | 1 | -1 | 0 | 41.03 | 7 |
| 1 | -1 | 1 | 0 | 43.20 | 8 |
| 1 | 0 | -1 | 1 | 34.47 | 8 |
| 1 | 1 | 0 | -1 | 41.05 | 8 |

$$y = \mu_y\left(\mathbf{x},\boldsymbol{\beta}\right) + \mathrm{h}\left(\mathbf{x}_{(1)},\mathbf{z},\boldsymbol{\theta}\right) + e$$

where h is given by (3) with $\gamma = 0.5$. Here, for the simplicity of comparisons, the factor $x_1$ is the only control factor appearing in the functions $\psi_j$:

$$\psi_1 = -5 + 3X_1 + 2u_2\left(X_1\right),\ \psi_2 = 1 - X_1,\ \psi_3 = -1 + X_1$$

The standard deviation of the normally distributed error term is 2. Two hundred samples were simulated from the true model in the Monte Carlo studies. For each simulated sample, the four methods RSM, CR, CRQV, and GR are each used to fit a variance function. Table 3 summarizes the results for estimating the variances of *y* for $X_1$ = -1.5, -1.0, -0.5, 0, 0.5, 1.0, 1.5. In reporting the simulation results, we calculate both an estimate's relative bias (RB, defined as (mean of variance estimate – true variance)/true variance) and the coefficient of variation (CV, defined as SD of variance estimate/Mean of variance estimate).

It is seen that in most cases, the CR and the CRQV approaches can be heavily biased (with RB greater than 15%) even for $X_1$ within the boundary of the experimental region. The bias is particularly severe if the two approaches are used for extrapolating variances ($X_1$ = -1.5 and 1.5). The two approaches have in general about the same CV in estimating variances. These CV of variance estimates are also comparable to those of the GR approaches (with the exception

of the case $X_1 = 1.5$ where the GR approach has a considerably smaller CV) which in general has smaller biases. The RSM approach has about the same biases as GR and far smaller biases than CRQV in most cases. These observations suggest that the classical approach can be very inadequate even if the true model is reasonably approximated by a quadratic function.

**Table 3.** Expected values and standard deviations of variance estimates obtained by simulations

| | | RSM | | GR | | CRQV | | CR | |
|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | True variance | Mean RB | SD CV | Mean RB | SD CV | Mean RB | SD CV | Mean RB | SD CV |
| -1.5 | 28 | 26.08 | 8.74 | 26.88 | 6.25 | 19.25 | 4.90 | 41.47 | 13.02 |
| | | -0.069 | 0.335 | -0.040 | 0.233 | -0.313 | 0.254 | 0.481 | 0.314 |
| -1.0 | 18 | 16.69 | 3.94 | 16.86 | 3.49 | 15.00 | 3.23 | 17.51 | 3.64 |
| | | -0.073 | 0.236 | -0.064 | 0.207 | -0.167 | 0.215 | -0.027 | 0.208 |
| -0.5 | 11 | 10.25 | 2.80 | 10.15 | 2.32 | 11.80 | 2.51 | 8.55 | 2.16 |
| | | -0.068 | 0.273 | -0.077 | 0.228 | 0.073 | 0.213 | -0.223 | 0.253 |
| 0.0 | 7 | 5.96 | 2.10 | 5.99 | 1.59 | 8.77 | 1.96 | 5.33 | 1.45 |
| | | -0.148 | 0.351 | -0.145 | 0.265 | 0.253 | 0.224 | -0.239 | 0.271 |
| 0.5 | 6 | 4.84 | 1.93 | 5.02 | 1.61 | 6.88 | 1.81 | 4.63 | 1.36 |
| | | -0.194 | 0.398 | -0.164 | 0.321 | 0.146 | 0.263 | -0.229 | 0.294 |
| 1.0 | 8 | 7.03 | 3.28 | 7.19 | 3.03 | 5.97 | 2.29 | 7.09 | 2.90 |
| | | -0.122 | 0.466 | -0.101 | 0.421 | -0.253 | 0.383 | -0.114 | 0.409 |
| 1.5 | 15 | 12.14 | 6.91 | 13.57 | 5.27 | 5.86 | 3.05 | 16.63 | 8.61 |
| | | -0.190 | 0.569 | -0.096 | 0.388 | -0.609 | 0.521 | 0.109 | 0.518 |

The simulation studies shed some light on the performance of the different methods in practice. Of the four approaches, the RSM is the only one that does not rely on the knowledge of the functional form of the function h. In fact, it does not even model variance involving the noise variables **Z** controlled in the experiment. It simply approximates the variance function directly with a linear or quadratic function of the design variables. Consequently, it does not suffer from the same potential model misspecification (of the variance) experienced by the other methods. This is consistent with the simulation results as the bias of RSM is seen to be generally smaller than those based on the regression approach. However, since the fitting relies on the sample standard deviations based on repeated observations, which in general have greater sampling variability, this robustness is achieved at the expense of an inflated variance of variance estimates – of the four approaches, it has substantially higher CV. For the regression approaches (GR, CR, CRQV), the variance parameters in $\boldsymbol{\theta}$ are more efficiently

estimated as the regression coefficients of a mean regression model and therefore, as observed in the simulation studies, have smaller variances than RSM. Thus in practice, the regression approaches may be preferred, but caution must be taken to ensure the validity of the model, especially the functional form of the function h. The quadratic variance function of the CRQV approach is actually in the form of the square of a linear function and therefore does not have the same effectiveness in approximating h as a general quadratic function. In this regard, the extension suggested in the previous section, provides a more flexible and effective means of approximating the true h function, as demonstrated in the simulation studies where GR has considerably smaller biases than CR and CRQV in most cases.

## Conclusion

Mak and Nebebe (2004) demonstrated the importance of the incorporation of the mean-variance relationship, if it exists, in analyzing crossed or combined array designs. They also proposed a model generalizing the traditional method of analysis. In this paper, we proposed a simple method of determining an appropriate mean-variance relation to be used in the model. An estimation procedure is also proposed for the model. With a numerical example, the advantages of Mak and Nebebe's model is demonstrated in terms of variance minimization. In terms of robustness of mean estimation to mis-specification of the variance function, the dual response surface methodology is also appealing, though it has other limitations. It might also be interesting to modify the proposed estimation by modifying the second estimating equation so that the estimation of the regression parameter is still consistent but less adversely affected by model misspecification.

The model proposed by Mak and Nebebe (2004) assumed an error term with homogeneous variances. In analyzing combined array designs, Engel and Huele (1996) considered a model in which the error terms have heterogeneous error variances which are functions of some of the design factors (they however assume $V_\lambda(\mu_y) \equiv 1$). This generalization may also be incorporated in Mak and Nebebe's model and the iterative estimation method suggested will then have to be modified accordingly, using traditional methods of regression analysis with heterogeneous variances. However, when the noise factors have already accounted for the majority of the unconditional variance of the quality characteristic so that the error term is in general small, this modification may not yield substantial practical differences.

## Acknowledgements

The authors are grateful to the referee for the helpful comments. This research was supported by NSERC-GRF grant (N01374) and VPRGS seed funding (VS0141) of Concordia University.

## References

Barreau, A., Chassagnon, R., Kobi, A., & Seibilia, B. (1999). Taguchi's parameter design: An improved alternative approach. *53rd Annual Quality Congress Proceedings* (pp. 400-404). Milwaukee, WI: American Society for Quality.

Box, G. E. P. (1988). Signal-to-noise ratios, performance criteria, and transformation. *Technometrics, 30*(1), 1-31. doi: 10.2307/1270311

Chan, L. K., & Mak, T. K. (1995). A regression approach for discovering small variation around a target. *Applied Statistics, 44*(3), 369-377. doi: 10.2307/2986043

Choi, H. J., & Allen, J. K. (2009). A metamodeling approach for uncertainty analysis of nondeterministic systems. *Journal of Mechanical Design, 131*(4), 041008. doi: 10.1115/1.3087565

Engel, J., & Huele, A. F. (1996). A generalized linear modeling approach to robust design. *Technometrics, 39*(4), 365-373. doi: 10.2307/1271307

Khuri, A. I., & Mukhopadhyay, S. (2010). Response surface methodology. *Wiley Interdisciplinary Reviews: Computational Statistics, 2*(2), 128-149. doi: 10.1002/wics.73

Mak, T. K., & Nebebe, F. (2004). Modeling problems in parameter designs. *Journal of Statistical Research, 38*, 155-166.

Mak, T. K., & Nebebe, F. (2005). Estimation of process variances in robust parameter designs. *Journal of Modern Applied Statistical Methods, 4*(2), 394-401. Retrieved from http://digitalcommons.wayne.edu/jmasm/vol4/iss2/5/

O'Donnell, E. M., & Vining, G. G. (1997). Mean squared error of prediction approach to the analysis of a combined array. *Journal of Applied Statistics, 24*(6), 733-746. doi: 10.1080/02664769723468

Robinson, T. J., Borror, C. M., & Myers, M. H. (2004). Robust parameter design: A review. *Quality and Reliability Engineering International, 20*(1), 81-101. doi: 10.1002/qre.602

Shoemaker, A. C., Tsui, K. L., & Wu, C. F. (1991). Economical experimentation methods for robust design. *Technometrics, 33*(4), 415-427. doi: 10.2307/1269414

Vandenbrande, W. (1998). SPC in paint application: Mission impossible. *52nd Annual Quality Congress Proceedings* (pp. 708-715). Milwaukee, WI: American Society for Quality.

Vining, G. G., & Myers, R. H. (1990). Combining Taguchi and response surface philosophies – A dual response approach. *Journal of Quality Technology, 22*(1), 38-45.

Welch, W. J., Yu, T. K., Kang, S. M., & Sacks, J. (1990). Computer experiments for quality control by parameter design. *Journal of Quality Technology, 22*(1), 15-22.