5-1-2003

Improved Multiple Comparisons With The Best In Response Surface Methodology

Laura K. Miller
University of North Florida

Ping Sa
University of North Florida, psa@unf.edu

Early Scholars
Improved Multiple Comparisons With
The Best In Response Surface Methodology

Laura K. Miller   Ping Sa
University of North Florida

A method to construct simultaneous confidence intervals about the difference in mean responses at the stationary point and at x for all x within a sphere with radius R_I is proposed. Results of an efficiency study to compare the new method and the existing method by Moore and Sa (1999) are provided.

Key words: Comparison with the best, response surface methodology, bounding algorithm.

Introduction

Response surface methodology uses a polynomial response function to explain and analyze the relationship between a response variable y and several predictor variables $\mathbf{x} = (\mathbf{x}_1, ..., \mathbf{x}_k)'$. Usually the $\mathbf{x}_i$ will be converted to coded variables x_i by $x_i = (\mathbf{x}_i - \mathbf{x}_{i0})/(sc)_i$, where $\mathbf{x}_{i0}$ is a centering constant and $(sc)_i > 0$ is a scaling constant, $i = 1,2,...,k$. The mean response at $\mathbf{x} = (x_1, x_2,..., x_k)'$, $E(y \mid \mathbf{x})$, can be approximated using the quadratic polynomial model with k predictor variables

$$y = \mathbf{b}_0 + \sum_i \mathbf{b}_i x_i + \sum_i \mathbf{b}_{ii} x_i^2 + \sum_{i<j} \mathbf{b}_{ij} x_i x_j + \mathbf{e} \, ,$$

Laura K. Miller received her Master degree in Statistics from the University of North Florida. She was also the recipient of the Most Outstanding Statistics Graduate Student and the Most Outstanding Mathematics Undergraduate Student for University of North Florida in the year 2000 and 1998, respectively. Ping Sa is an Associate Professor of the Department of Mathematics and Statistics at the University of North Florida. She received the Ph.D. in Statistics from the University of South Carolina in 1990. She has published 15 papers. Her recent scholarly activities have involved research in multiple comparisons and quality control. Email address for correspondence regarding this article is psa@unf.edu.

where $\mathbf{b}_0, \mathbf{b}_i, \mathbf{b}_{ij}$ are unknown constants, $i, j = 1,2,...,k$ and random error $\mathbf{e} \sim NID(0, \mathbf{s}^2)$.

The mean response is optimized at the stationary point that may be a minimum, maximum, or a saddle point. After determining the levels of the predictor variables where the mean response is optimized, it is possible that this point is not a reasonable option due to practical considerations, such as expense. In this situation, multiple comparisons can be performed with other points in the region to determine if some other points provide responses that are not significantly different from the optimal point.

This problem will subsequently be referred to as multiple comparisons with the best (MCB) in response surface methodology (RSM). The MCB problem was first approached by Hsu (1984) in design of experiments where he considered the problem of comparing the treatment means under study with the "best" treatment mean. Moore and Sa (1999) first approached the MCB problem in the RSM setting. There has also been other substantial work on related problems within the field of response surface methodology. Sa and Edwards (1993) and Merchant, McCann, and Edwards (1998) investigated the multiple comparisons with the control (MCC) problem.

Sa and Edwards (1993) first addressed the MCC in RSM problem by constructing simultaneous confidence intervals for $\mathbf{d}_C(\mathbf{x}) = E(y \mid \mathbf{x}) - E(y \mid \mathbf{0})$ for all $\mathbf{x}$ within a

pre-specified distance R_I of the origin, such that for all x, $x'x = \sum_{i=1}^{k} x_i^2 \leq R_I^2$, where R_I is the "radius of inference." They showed that for a rotatable design, the bounds of $\boldsymbol{d}_C(\boldsymbol{x}) \in \hat{\boldsymbol{d}}_C(\boldsymbol{x}) \pm (rF_{a,r,n})^{1/2} s(\boldsymbol{x})$ can be improved using a result of Casella and Strawderman (1980) where the Scheffé critical point, $(rF_{a,ru})^{1/2}$ can be replaced by a smaller value c_a depending on $\boldsymbol{a}$, $\boldsymbol{u}$, and the nature of the predictor constraints as summarized by two other constraints, an integer m and a distance $q^2 > 0$.

Because the design used in practice is often not rotatable, Merchant, McCann, and Edwards (1998) introduced a new method which combined the Bonferroni method and the McCann and Edwards (1996) algorithm for two or more predictors that gives much sharper intervals than the Scheffé and also the Sa and Edwards (1993) adaptation of the Casella and Strawderman method. Merchant, McCann, and Edwards' method does not require a rotatable design and allows for one-sided bounds for $\boldsymbol{d}_C(\boldsymbol{x})$. They generated a critical point d with simultaneous confidence bounds for

$$\boldsymbol{d}_C(\boldsymbol{x}) = E(y \mid \boldsymbol{x}) - E(y \mid \boldsymbol{0})$$

for all $\boldsymbol{x}$ within a specified distance of $\boldsymbol{0}$ via a bounding algorithm that requires only a few seconds to a few minutes of computer time.

Closely related and within the field of RSM, Moore and Sa (1999) addressed the MCB problem. They constructed confidence intervals about the difference in mean responses at the stationary point and alternate points over the entire k dimensional hyperplane based on a theory that does not depend on the design of the experiment. To solve the MCB problem, they utilized the delta method to approximate the variance of the estimated difference for

$$\boldsymbol{d}_B(\boldsymbol{x}) = \mathbf{d}_B(\boldsymbol{x}, \mathbf{b})$$
$$= E(y \mid \boldsymbol{x}_0) - E(y \mid \boldsymbol{x})$$
$$= -\boldsymbol{x}'\boldsymbol{b} - \boldsymbol{x}'\boldsymbol{B}\boldsymbol{x} - \frac{1}{4}\boldsymbol{b}'\boldsymbol{B}^{-1}\boldsymbol{b},$$

where $\boldsymbol{d}_B(\boldsymbol{x})$ represents the difference between the mean response at the stationary point $\boldsymbol{x}_0 = -\frac{1}{2}\boldsymbol{B}^{-1}\boldsymbol{b}$ and the mean response at any other point $\boldsymbol{x}$, $\boldsymbol{b} = (\beta_1, \beta_2, ..., \beta_k)'$, and

$$\boldsymbol{B} = \begin{bmatrix} \beta_{11} & \frac{1}{2}\beta_{12} & \cdots & \frac{1}{2}\beta_{1k} \\ & \beta_{22} & \cdots & \frac{1}{2}\beta_{2k} \\ & & \ddots & \vdots \\ sym & & & \beta_{kk} \end{bmatrix}.$$

This confidence interval is useful in determining whether an alternate point could be substituted for the stationary point as an optimizer. Furthermore, it provides how much loss in the mean response can be expected if $\boldsymbol{x}$ is moved away from $\boldsymbol{x}_0$. They investigated both Bonferroni and Scheffé type confidence intervals for the MCB problem. They also investigated Scheffé's F-projection method of constructing conservative confidence intervals. However, the delta method is much less conservative than the F-projection method and of course, much easier to use.

It is the purpose of this article to address the MCB problem in RSM, but instead of considering the entire k-dimensional space, it would be more realistic to restrict the region to provide confidence bounds for $\boldsymbol{d}_B(\boldsymbol{x})$ within a sphere with radius R_I for all $\boldsymbol{x}$ such that $\boldsymbol{x}'\boldsymbol{x} \leq R_I^2$. The method proposed by Merchant, McCann, and Edwards (1998) for the MCC problem should be adaptable to the MCB problem since the requirement for using this method is that the covariance matrix of the estimators must be known.

The delta method will be used to approximate the variance of $\hat{\boldsymbol{d}}_B(\boldsymbol{x})$ for the MCB problem. The next section explains the theory and the bounding algorithm used to generate the critical point for the MCB problem. The algorithm is design free, that is, it does not depend on the design of the experiment and should therefore provide consistent results regardless of the design.

Theory Behind the Method

The method proposed by Merchant, McCann, and Edwards (1998) will be adapted to solve the MCB problem. The goal is to generate an improved critical point d with simultaneous upper confidence bounds of the form

$$\hat{d}_B(x) + ds(x) \qquad (1)$$

where

$$\hat{d}_B(x) = \hat{d}_B(x,\hat{b}) = -x'\hat{b} - x'\hat{B}x - \frac{1}{4}\hat{b}'\hat{B}^{-1}\hat{b},$$

such that $\hat{b} = (\hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_k)'$ where the $\hat{\beta}_i$'s are the least square estimators for β_i's and $\hat{B}$ is the matrix such that $\hat{\beta}_{ij}$ is substituted into the matrix B. The estimated standard error of $\hat{d}_B(x)$ is s(x) = s$(l'\Sigma l)^{1/2}$ derived by Moore and Sa (1999) where s^2 is the mean square error which satisfies $\dfrac{us^2}{s^2} \sim c^2(u)$ for integer $u > 0$ and is independent of all $\hat{\beta}_i$'s, $\dot{a}$ is the $(X'X)^{-1}$ matrix with the first row and the first column deleted and l is the vector of partial derivatives of $\delta_B(x,\beta)$ with respect to b such that

$$l = (-\tfrac{1}{2}m_1 - x_1, ..., -\tfrac{1}{2}m_k - x_k, \tfrac{1}{4}m_1^2$$
$$-x_1^2, ..., \tfrac{1}{4}m_k^2 - x_k^2, \tfrac{1}{4}m_1m_2 - x_1x_2, ...,$$
$$\tfrac{1}{4}m_{k-1}m_k - x_{k-1}x_k)',$$

where $m = B^{-1}b$ such that m_i is the *ith* component of m and $x = (x_1, x_2, ..., x_k)'$ is any point satisfying $x'x \leq R_I^2$.

In order to approximate the entire set of interest, we adapt the fine grid of inference G_I, suggested by Merchant, McCann, and Edwards (1998) of individual x_j points for $j = 1, 2, ..., p$ in the region. This grid is constructed by user defined multiples of these x values within a radius R_I radiating from the center $\mathbf{0}$. This matrix is defined as $L: p \times r$ whose j^{th} row is $l'_j = l(x_j)'$

where a simultaneous bound over this finite collection is calculated.

The critical point d must satisfy

$$P\left[\max_{x \in G_I} \frac{\hat{d}_B(x) - d_B(x)}{s(l'\dot{a}l)^{1/2}} \leq d \right] \geq 1 - a . \qquad (2)$$

For each x, $\dfrac{\hat{d}_B(x) - d_B(x)}{s(l'\dot{a}l)^{1/2}} \sim t_u$, where t_u is the univariate-t distribution with u degrees of freedom. Equation (2) can be rewritten as

$$P\left[T_j \leq d, j = 1,2,...,p \right] \geq 1 - a$$

or

$$P\left[T_j > d, j = 1,2,...,p \right] \leq a$$

where $T_1, T_2, ..., T_p$ have a multivariate t distribution (Dunnett & Sobel 1954) with u degrees of freedom and underlying correlation matrix R derived from $s^2 L \sum L'$. The critical point d is then solved by the following equation,

$$\int_0^{1/d} P\{E(t)\} f_T(t)dt = a \text{ for}$$

$$E(t) = \bigcup_{j=1}^{p} (a'_j U > td), \quad (\text{Brown,1984})$$

where f_T is the probability density function of T, a random variable such that rT^2 is distributed as $F(u,r)$; U is a random vector independent of T, distributed uniformly on the r-dimensional sphere; and a'_j are the rows of the full rank matrix $A: p \times r$ such that $R = AA'$.

Finally, the probability $P\{E(t)\}$ for the MCB problem can be calculated using the same bounding algorithm proposed by Merchant, McCann, and Edwards (1998) for the MCC problem. This bounding algorithm is a combination of Bonferroni method and the McCann and Edwards algorithm (1996) and is for upper bound only. If a lower bound is required

over this region, that is $\hat{\boldsymbol{d}}_B(\boldsymbol{x}) - ds(\boldsymbol{x})$, it can be computed by constructing the upper bound for $-\boldsymbol{d}_B(\boldsymbol{x}) = E(y\mid \boldsymbol{x}) - E(y\mid \mathbf{x}_0)$ because an upper bound for $-\boldsymbol{d}_B(\boldsymbol{x})$ is equivalent to a lower bound for $\boldsymbol{d}_B(\boldsymbol{x})$.

The critical points for the MCB problem were computed using a Fortran program and routines from the IMSL Fortran Numerical Libraries (1997). These included calling the routines DLINRG to calculate the inverse of a matrix, DLFTDS to compute the Cholesky factorization of a matrix, DFDF to evaluate the F distribution function, and DQDAGS to perform the numerical integration. The Fortran program is available from the first author.

Examples and comparisons

Box and Draper (1987) give an example from an investigation by Derringer and Suich (1980) in which RSM is used to analyze the effects of $\boldsymbol{x}_1 =$ hydrated silica level in phr (parts per hundred) and $\boldsymbol{x}_2 =$ silane coupling agent level in phr on the elongation at break of a tire tread compound. One of the goals was to maximize $y =$ elongation at break. Convert $\boldsymbol{x}_1$ to the coded variable $x_1 = (\boldsymbol{x}_1 - 1.2)/0.5$ and $\boldsymbol{x}_2$ to the coded variable $x_2 = (\boldsymbol{x}_2 - 50)/10$. The design points $\boldsymbol{x}_i$ and the responses y_i are listed in Table 1.

Table 1. Experimental Results: Elongation at break y of a tire tread compound Versus $x_1 =$(phr silica $-$ 1.2)/0.5 and $x_2 =$ (phr silane $-$ 50)/10 (Source: Derringer 1980)

Run	x_1	x_2	y
1	-1	-1	900
2	1	-1	860
3	-1	1	800
4	1	1	2294
5	-1	-1	490
6	1	-1	1289
7	-1	1	1270
8	1	1	1090
9	-1.633	0	770
10	1.633	0	1690
11	0	-1.633	700
12	0	1.633	1540
13	0	0	2184
14	0	0	1784
15	0	0	1300
16	0	0	1300
17	0	0	1145
18	0	0	1090
19	0	0	1260
20	0	0	1344

The estimated polynomial response function is

$$\hat{y} = 1412.892 + 268.151x_1 + 246.503x_2 - 97.794x_1^2 - 139.044\,x_2^2 + 69.375x_1x_2 \ .$$

The vector $\hat{b} = (268.151, 246.503)'$, the matrix $\hat{B} = \begin{Bmatrix} -97.794 & 34.688 \\ 34.688 & -139.044 \end{Bmatrix}$, and the matrix

$$\dot{a} = \begin{bmatrix} .075 & 0 & 0 & 0 & 0 \\ 0 & .075 & 0 & 0 & 0 \\ 0 & 0 & .075 & .005 & 0 \\ 0 & 0 & .005 & .075 & 0 \\ 0 & 0 & 0 & 0 & .125 \end{bmatrix}$$

are calculated.

The estimated stationary point for this surface where elongation at break (y) is maximized is $\hat{x}_0 = (1.849, 1.348)$ yielding an estimated response of 1826.91. Figure 1 gives the estimated surface plot.
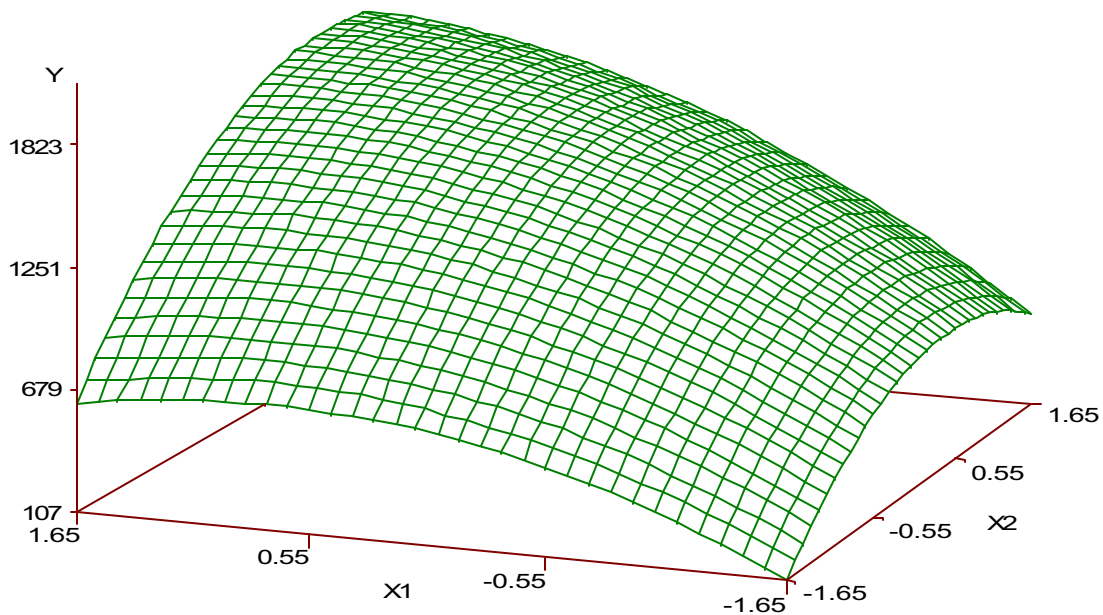

Figure 1. Estimated Surface Plot for the Tire Tread Compound Example.

As one can see, the stationary point is out of the experimental region and it may not be a reasonable option due to practical considerations or expense. Therefore, multiple comparisons can be performed with other points in a region to determine if any other point within the region of operability will produce a response that is not significantly different from the point that maximizes elongation at break (y). Since the optimal point was a maximum, this suggests that lower bounds for $d_B(x) = E(y \mid x_0) - E(y \mid x)$ are more important than upper bounds.

Simultaneous 90% lower confidence bounds are constructed for $d_B(x)$ for all x whose values are on the grid defined by multiples of .2 with a radius of $R_I = \sqrt{2}$ radiating from the center of the region of interest. Figure 2 shows the contours for the estimated difference and the simultaneous lower confidence bounds $L(x) = \hat{d}_B(x) - ds(x)$ for generated $d^2 = 6.871607$ by the method detailed in the previous section.
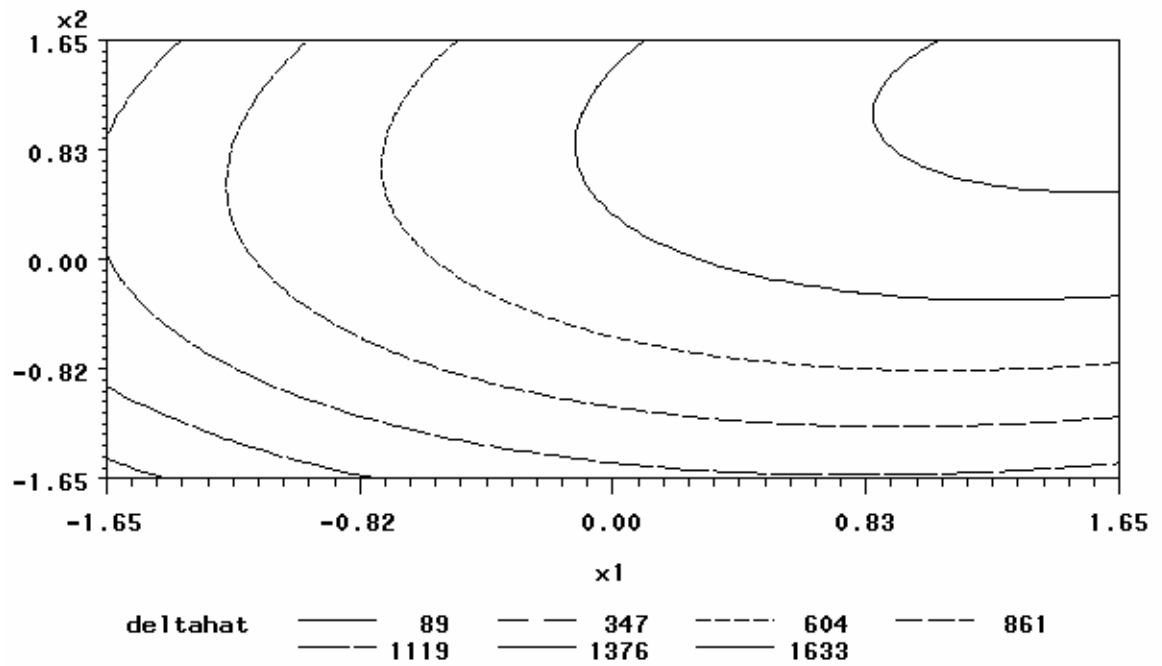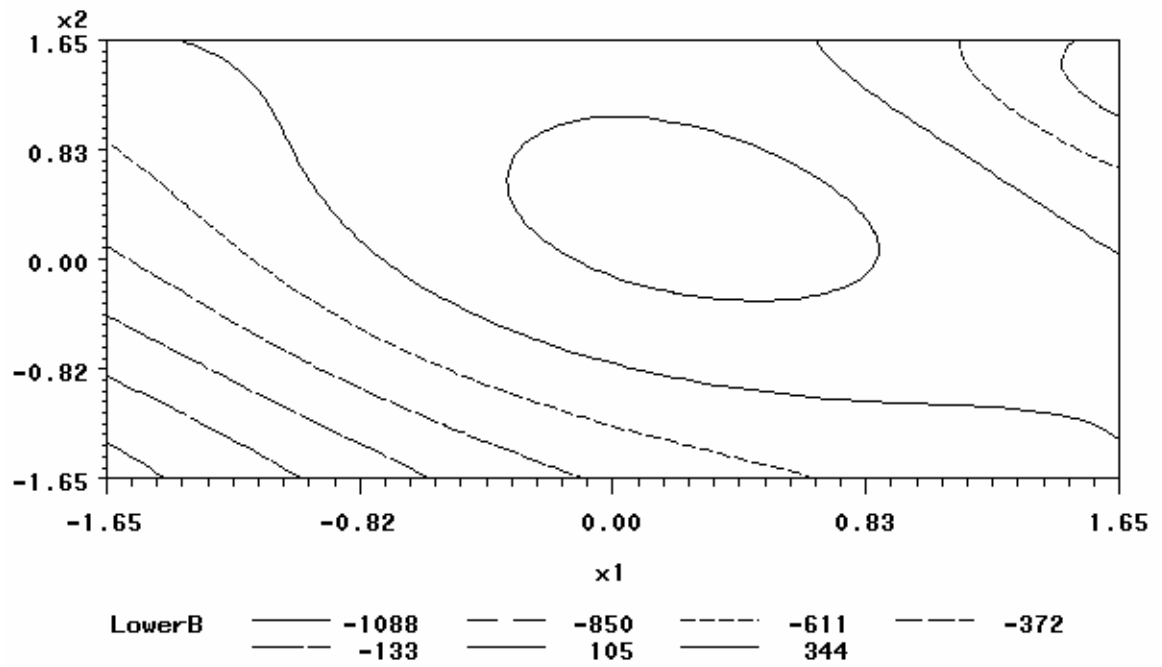
Figure 2. Contour Plots

In Figure 2 (above), the contour plots are for the estimated difference $\hat{d}_B(x)$ (top) and the simultaneous 90% lower confidence bounds (bottom) for the Tire Tread Compound Example with generated $d^2 = 6.871607$. The points that lie inside the negative contour lines indicate possible alternate points that will produce responses that are not significantly different from the point that maximizes elongation at break (y). The region inside the negative contour lines indicates possible region that will produce responses that are not significantly different from the point that maximizes elongation at break (y).

The squared critical constant $d^2 = 6.871607$ compares very favorably to that of the Scheffé method, $rF_a(r,u) = (\sqrt{5F_{.10}(5,14)})^2 = 11.534702$. Therefore, an experimenter using the Scheffé method would have to increase the experiment size by a factor of approximately $(\sqrt{5F_{.10}(5,14)})^2/d^2 = 11.534702 \, / \, 6.871607 = 1.6786$, in other words, by 67.86% in order to achieve a precision of estimation (interval width) equal to what would be obtained using the adapted method's critical constant.

Next, three different designs will be used for an example using the bounding algorithm to generate improved critical points for the MCB problem where the sample-size savings will be compared to the Scheffé and Bonferroni critical points.

Khuri and Cornell (1987) provide an example in which they use RSM to investigate the effects of the amounts of two fertilizers, x_1 and x_2, on the yield of peanuts measured in pounds per acre. For the purpose of the efficiency study, the estimated parameters from this example will be treated as the true parameters of an underlying model. The true quadratic response function is given by

$$y = 13.85 - .90x_1 + .56x_2$$
$$-1.94x_1^2 - .78x_2^2 - .57x_1x_2 + e \, .$$

The vector $b = (.90, .56)'$, the matrix

$$B = \begin{bmatrix} -1.94 & -.285 \\ -.285 & -.78 \end{bmatrix}, \text{ and the matrix}$$

$$\dot{a} = \begin{bmatrix} .125 & 0 & 0 & 0 & 0 \\ 0 & .125 & 0 & 0 & 0 \\ 0 & 0 & .144 & .019 & 0 \\ 0 & 0 & .019 & .144 & 0 \\ 0 & 0 & 0 & 0 & .25 \end{bmatrix} \text{ are found.}$$

The stationary point for this surface is $x_0 = (.189, .290)$ yielding a response of 13.676. Assume that this option is not a reasonable option, multiple comparisons are performed to determine if alternate points can substitute for the stationary point in terms of maximizing peanut yield. Therefore, the critical point d is required to perform these comparisons.

Three central composite designs were chosen. The three designs are a rotatable central composite design with uniform precision, a rotatable central composite design without uniform precision, and a central composite design with one centerpoint. These designs will be referred to as Design 1, 2, and 3 respectively.

Table 2 (following References section) shows the critical points that were generated for one and two replications of the three different designs using multiples of .2, $R_I = 1$, and $\sqrt{2}$ for this example using Merchant, McCann, and Edwards' (1998) method in order to compare the critical values and the approximate sample-size savings for each design. Because the Bonferroni method is conservative due to the large number of comparisons, only the approximate sample-size savings vs the Scheffé method were calculated.

Considerable improvement (between 34% and 47%) over the Scheffé adaptation for all three designs is possible using the new method by choosing the radius of inference $R_I = 1$. For $R_I = \sqrt{2}$ (which is near the limits of the experimental region for Designs 1 and 2), the sample-size savings are 26% to 33% over the Scheffé method. Also, as expected, the increased sample sizes produced by replicating the designs resulted in smaller critical values.

Conclusion

In conclusion, this article has addressed the problem of multiple comparisons with the best in RSM via simultaneous confidence bounds for $d_B(x) = E(y \mid x_0) - E(y \mid x)$ for all x such that $x'x \leq R_I^2$. The method proposed by Merchant, McCann, and Edwards (1998) for the MCC problem has been adapted to the MCB problem. It has provided confidence bounds for an example for two predictors where the critical values compare favorably to the Bonferroni and Scheffé methods as shown by Table 2 (following page).

This will also hold true for problems containing more than two predictor variables. For the example provided, this method has been shown to provide approximate sample-size savings of at least 25% for three different central composite designs. In fact, based on the theory behind the bounding algorithm, the Merchant, McCann, and Edwards' method for the MCB problem will always outperform the Scheffé and Bonferroni methods (Merchant, McCann, and Edwards, 1998).

References

Box, G. E. P. & Draper, N. R. (1987), *Empirical model-building and response surfaces*, New York: John Wiley.

Brown, L. D. (1984), A note on the Tukey-Kramer procedure for pairwise comparisons of correlated means, in *Design of experiments: ranking and selection*, eds. T. J. Santner & A. C. Tamhane, New York: Marcel Dekker, pp.1-6.

Casella, G., & Strawderman, W. E. (1980), Confidence bands for linear regression with restricted predictor variables, *Journal of the American Statistical Association*, *75,* 862-868.

Derringer, G. C. & Suich, R. (1980), Simultaneous optimization of several response variables, *Journal of Quality Technology*, *12,* 214-219.

Dunnett, C. W., & Sobel, M. (1954), A bivariate generalization of student's *t* distribution, with tables for certain special cases, *Biometrika, 31,* 153-169.

Hsu, J. C. (1984), Constrained simultaneous confidence intervals for multiple comparisons with the best, *Annals of Statistics, 12,* 1136-1144.

International Mathematical & Statistical Libraries, Inc. (1997), *Fortran routines for mathematical applications*, Visual Numerics, Inc.

Khuri, A. I., & Cornell, J. A. (1987), *Response surfaces*, New York: Marcel Dekker.

McCann, M., & Edwards, D. (1996), A path-length inequality for the multivariate *t* distribution, *Journal of the American Statistical Association, 91,* 211-216.

Merchant, A., McCann, M., & Edwards, D. (1998), Improved multiple comparisons with a control in response surface analysis, *Technometrics, 40,* 297-303.

Moore, L. J., & Sa, P. (1999), Comparisons with the best in response surface methodology, *Statistics and Probability Letters, 44,* 189-194.

Sa, P. & Edwards, D. (1993), Multiple comparisons with a control in response surface methodology, *Technometrics, 35,* 436-445.

Table 2. Generated critical points using the improved method (one-sided bounds) and Scheffé and Bonferroni critical points with grid spacing = .2 for each design and approximate Sample-Size Savings of the New Method Versus the Scheffé method.

Design	R_I	reps	u	Improved Critical Point d	Scheffé Critical Point	Bonferroni Critical Point	Sample-Size Savings vs Scheffé
1	1	1	7	3.233020	3.794733	4.605120	37.78%
		2	20	2.803443	3.286335	3.460804	37.42%
	$\sqrt{2}$	1	7	3.368990	3.794733	5.207830	26.87%
		2	20	2.918108	3.286335	3.756539	26.83%
2	1	1	3	4.423153	5.152669	9.505157	35.71%
		2	12	2.980882	3.456877	3.813342	34.49%
	$\sqrt{2}$	1	3	4.576950	5.152669	12.008948	26.74%
		2	12	3.079915	3.456877	4.195280	25.98%
3	1	1	3	4.255092	5.152669	9.505157	46.64%
		2	12	2.875548	3.456877	3.813342	44.52%
	$\sqrt{2}$	1	3	4.482340	5.152669	12.008948	32.15%
		2	12	3.018781	3.456877	4.195280	31.13%