Response Surface Methodology: A Retrospective and Literature Survey

RAYMOND H. MYERS
Virginia Tech, Blacksburg, VA 24061

DOUGLAS C. MONTGOMERY
Arizona State University, Tempe, AZ 85287

G. GEOFFREY VINING
Virginia Tech, Blacksburg, VA 24061

CONNIE M. BORROR
Drexel University, Philadelphia, PA 19104

SCOTT M. KOWALSKI
Minitab Inc., State College, PA 16801

Response surface methodology (RSM) is a collection of statistical design and numerical optimization techniques used to optimize processes and product designs. The original work in this area dates from the 1950s and has been widely used, especially in the chemical and process industries. The last 15 years have seen the widespread application of RSM and many new developments. In this review paper we focus on RSM activities since 1989. We discuss current areas of research and mention some areas for future research.

KEY WORDS: Bayesian Designs; Computer Experiments; Generalized Linear Models; Multiple Responses; Robust Parameter Design; Split-Plot Designs; Variance Dispersion Graphs.

Introduction

Previous Review Articles

Box and Wilson (1951) laid the foundations for response surface methodology (RSM) (a list of abbreviations is provided in the Appendix). That paper is important not only because it described what became an entire field of research for the next 50 years but also because it changed dramatically the way that engineers, scientists, and statisticians approached industrial experimentation. Their paper outlined a sequential philosophy of experimentation that encompasses experiments for screening, region seeking (such as steepest ascent), process/product characterization, and process/product optimization. Clearly, RSM includes much more than second-order model fitting and analysis. Indeed, RSM, broadly understood, has become the core of industrial experimentation. Box and Liu (1999) illustrated the application of RSM to the common training example of paper helicopters. Box (1999) provided a retrospective on the origins of RSM. That paper also outlined a more general philosophy of sequential learning, of which RSM is one tool.

Dr. Myers is Professor Emeritus in the Department of statistics.

Dr. Montgomery is a Professor in the Department of Industrial Engineering. He is a Fellow of ASQ. His email address is doug.montgomery@asu.edu.

Dr. Vining is Head of the Department of Statistics. He is a Fellow of ASQ. His email address is vining@vt.edu.

Dr. Borror is an Assistant Professor in the Department of Decision Sciences. She is a Senior Member of ASQ. Her email address is conni@drexel.edu

Dr. Kowalski is a Technical Trainer. He is a Member of ASQ. His email address is skowalski@minitab.com.

Over the past fifty years, there have been three extensive reviews of response surface methodology. The Hill and Hunter (1966) review paper featured an extensive bibliography and presented applications in the chemical and process industries, where the majority of RSM applications were found at that time, which was natural given that the seminal work of Box and Wilson occurred at a major chemical company. The Mead and Pike (1975) review paper in *Biometrics* focused more on the modeling of biological data than on a discussion of RSM as we view it.

The most recent review paper was that of Myers, Khuri, and Carter (1989). They emphasized the changes that had occurred in RSM theory and practice during the 1970s and 1980s. In addition to containing a bibliography of approximately 200 references, the paper discussed applications of RSM in physical science and engineering, food science, social science, and the biological and chemical sciences.

Myers, Khuri, and Carter (1989) emphasized that, at the time, "users [were] sufficiently far behind in the use of RSM tools that the need for proper communication [was] far more pressing." In other words, the authors felt that training and writing that emphasized RSM fundamentals for engineers and scientists in all fields was extremely important. The authors also lamented the need for better software to support RSM. Computer software developers have responded to that need, and today we find that many general-purpose statistics packages have good capability to support RSM and some stand-alone experimental design packages are excellent. These developments have expanded the use of designed experiments over the last decade. Software advances have been particularly encouraging in the area of analysis, including some important development in multiple response optimization.

In addition to the comprehensive review papers, discussions of various aspects of RSM can be found in papers by Draper and Lin (1996), Draper and Pukelsheim (1996), and Draper and Pukelsheim (1998). Mixture experiments are a useful class of response surface problems that we do not fully consider here. A comprehensive presentation of the subject was done by Cornell (2002).

Why A Review Article Now?

It has been 13 years since the last review, and much progress in the area of RSM has been made. Myers (1999) gave a recent overview of some of the current themes in the RSM literature as well as a personal perspective on future research directions. That paper, however, made no thorough attempt to summarize or synthesize the recent literature. It focused on recent literature only to the extent necessary to lay the foundations for some observations on near-future research areas. Myers (1999) helped define some research issues in RSM, but it only scratched the surface in terms of summarizing the recent literature. As a result, now seems to be an appropriate time to provide a more thorough review of the developments in RSM since Myers, Khuri, and Carter (1989).

We offer this review in order to present some historical perspective on an important and vital field of industrial statistics. However, our greatest hope is that this paper will benefit both researchers and practitioners by providing a single resource that summarizes and synthesizes the literature since 1989 and points the reader to original source material for more detail. The organization of the paper is as follows. The next section summarizes the many advances in robust parameter design that have taken place within the RSM framework, including design and modeling strategy. Subsequent sections discuss new developments in response surface designs, including methods for evaluating designs; multiple responses, including both design and optimization issues; RSM and generalized linear models; and other issues concerning applications, including RSM with randomization restrictions, computer experiments, and nonparametric and semi-parametric response surface methods. In the summary section, we present our views on a few key research issues in RSM.

Robust Parameter Design

Background

Taguchi introduced the *robust parameter design* (RPD) problem (see Taguchi (1987, 1991), Taguchi and Wu (1985), and Kackar (1985)). He advocated that experimenters should consider two sets of factors: control factors, x, that are under the complete control of the experimenter both in the experiment and in the process; and noise factors, z, that the experimenter can control in an experiment but cannot control in the process. The noise factors are significant sources of variation in the process. In an RPD problem, the experimenter seeks settings of the control factors that will be robust or insensitive to variability transmitted into the response from the noise factors. Taguchi's solution to this problem consisted of placing the control factors in one design (called the inner array) and the noise factors in a

second design (called the outer array), and then running the set of experiments given by the Cartesian product of these two designs, which produces a crossed array design. The analysis involves summary signal-to-noise ratios computed across the outer array observations. Throughout the 1980s and 1990s many people hotly debated some of the design and analysis aspects of Taguchi's approach; for details and summary see the discussions of Kackar (1985), Box (1985, 1988), Box, Bisgaard, and Fung (1988), Pignatiello and Ramberg (1991), Nair et al. (1992), Box and Jones (1992), Myers and Montgomery (2002), Tsui (1996), and Montgomery (1992, 1999, 2001).

RSM Alternatives to RPD

RSM is well suited to the RPD problem and process robustness studies in general. For examples, see Vining and Myers (1990), Myers (1991), Myers, Khuri, and Vining (1992), Lucas (1989, 1994), Pledger (1996), Khattree (1996), Myers, Kim, and Griffiths (1997), and Montgomery (1999). Two basic approaches evolved during the 1990s.

Vining and Myers (1990) first presented RPD problems as formal constrained optimization problems. Their approach assumed that the basic experiment in the control factors was replicated. The replication could, but not necessarily, be via an outer array. They proposed fitting separate models to the response and to the process variance observed by the replication scheme. They then used the dual response optimization technique of Myers and Carter (1973) to solve this problem. For example, to minimize the process variance while keeping the mean on a target, T, they minimized $\widehat{\mathrm{Var}}(x)$ subject to the constraint $\widehat{y}(x) = T$, where $\widehat{\mathrm{Var}}(x)$ comes from the model for the process variance and $\widehat{y}(x)$ comes from the model for the response. Del Castillo and Montgomery (1993), Copeland and Nelson (1996), Lin and Tu (1995), Kim and Lin (1998), and Fan (2000) presented various extensions for solving the dual response approach.

The second approach places both the control and noise factors in a single design, called a combined array. These designs typically require fewer runs than Taguchi's crossed arrays and simultaneously allow the experimenter to estimate potentially important interactions. Useful references on the combined array and its applications include Welch, Yu, Kang, and Sachs (1990), Lucas (1994), Montgomery (1990-1991), Shoemaker, Tsui, and Wu (1991), Borkowski and Lucas (1997), and Borror and Montgomery (2000).

The combined array approach uses a single response model. We suppose that there are r_1 control factors $\mathbf{x}' = [x_1, x_2, ..., x_{r_1}]$ and r_2 noise factors $\mathbf{z}' = [z_1, z_2, ..., z_{r_2}]$. The response model is usually of the form

$$y(\mathbf{x}, \mathbf{z}) = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} + \mathbf{z}'\boldsymbol{\gamma} + \mathbf{x}'\boldsymbol{\Delta}\mathbf{z} + \varepsilon, \quad (1)$$

where $\boldsymbol{\beta}$ is an $r_1 \times 1$ vector containing the regression coefficients of the control factors, $\mathbf{B}$ is an $r_1 \times r_1$ matrix whose main diagonals are the regression coefficients associated with the pure quadratic effects of the control factors and whose off-diagonals are one-half of the mixed quadratic (interaction) effects of the control factors, $\boldsymbol{\gamma}$ is an $r_2 \times 1$ vector of the regression coefficients for the main effects of the noise factors, and $\boldsymbol{\Delta}$ is a $r_1 \times r_2$ matrix of the control-factor-by-noise-factor interaction effects. Typically, Equation (1) is a full quadratic in the control factors and has all main effects of the noise factors along with all control-factor-by-noise-factor interactions. Often, with Equation (1) it is assumed that ε is NID(0, σ^2) and that the noise factors have been scaled so that they have mean zero and covariance matrix $\mathrm{Var}(\mathbf{z}) = \sigma_z^2\mathbf{V}$, where $\mathbf{V}$ is an $r_2 \times r_2$ symmetric positive definite matrix. Most authors assume that $\mathbf{V} = \mathbf{I}$ so that the noise factors are uncorrelated. There are many scenarios where this assumption is reasonable, such as when the noise factors are difficult-to-control process variables or raw material properties. If, however, the noise factors are environmental variables such as temperature and relative humidity, then they are likely to be correlated. It is also customary to assume that σ_z^2 and the elements of $\mathbf{V}$ are known, based on prior experience with the noise factors.

Myers, Khuri, and Vining (1992) obtained the model for the mean response by taking the conditional expectation of $y(\mathbf{x}, \mathbf{z})$ in Equation (1) with respect to the noise factors $\mathbf{z}$ and the random error, ε, which produces

$$\underset{\mathbf{z},\varepsilon}{E}[y(\mathbf{x}, \mathbf{z})] = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x}. \quad (2)$$

They used the variance operator to obtain the model for the response variance, which is

$$\underset{\mathbf{z},\varepsilon}{\mathrm{Var}}[y(\mathbf{x}, \mathbf{z})] = \sigma_z^2(\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})\mathbf{V}(\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})' + \sigma^2. \quad (3)$$

We note that $\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{x}$ is the vector of partial

derivatives of $y(\mathbf{x}, \mathbf{z})$ with respect to the noise factors $\mathbf{z}$. Thus, $\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{x}$ is the slope of the response surface in the direction of the noise factors. The importance of the slope of the response surface in the general RSM setting has been addressed by several authors, including Murty and Studden (1972), Myers and Lahoda (1975), Hader and Park (1978), Mukerjee and Huda (1985), Park (1987), Draper and Ying (1994), and Ying, Pukelsheim, and Draper (1995a,b). The response variance can also be developed by using the delta method, which is useful for models that are more complex than that in Equation (1). For more information on the delta method see Rice (1995).

In practice, the experimenter fits the response model using data from a designed experiment. This results in

$$\widehat{y}(\mathbf{x}, \mathbf{z}) = \widehat{\beta}_0 + \mathbf{x}'\widehat{\boldsymbol{\beta}} + \mathbf{x}'\widehat{\mathbf{B}}\mathbf{x} + \mathbf{z}'\widehat{\boldsymbol{\gamma}} + \mathbf{x}'\widehat{\boldsymbol{\Delta}}\mathbf{z}. \quad (4)$$

The parameter estimates in this model are used to obtain fitted versions of Equations (2) and (3),

$$\underset{\mathbf{z},\varepsilon}{\widehat{E}}[y(\mathbf{x}, \mathbf{z})] = \widehat{\beta}_0 + \mathbf{x}'\widehat{\boldsymbol{\beta}} + \mathbf{x}'\widehat{\mathbf{B}}\mathbf{x} \quad (5)$$

and

$$\widehat{\underset{\mathbf{z},\varepsilon}{\mathrm{Var}}}[y(\mathbf{x}, \mathbf{z})] = \widehat{\sigma}_z^2(\widehat{\boldsymbol{\gamma}}' + \mathbf{x}'\widehat{\boldsymbol{\Delta}})\mathbf{V}(\widehat{\boldsymbol{\gamma}}' + \mathbf{x}'\widehat{\boldsymbol{\Delta}})' + \widehat{\sigma}^2, \quad (6)$$

where $\widehat{\sigma}^2$ is the residual mean square from the fitted response model. Using Equations (5) and (6) the standard RPD problems can be formulated. One can then apply an appropriate constrained optimization procedure to obtain recommended settings.

References on the interactions between the noise and control factors include papers by Myers, Khuri, and Vining (1992), Shoemaker, Tsui, and Wu (1991), and Shoemaker and Tsui (1993). Bingham and Li (2002) proposed a criterion to maximize a design's ability to estimate models with at least one control-by-noise interaction. Focusing on the first-order case, they constructed optimal designs and tabulated some designs to facilitate practical use. Borror, Montgomery, and Myers (2002) developed scaled prediction error variance models for both the mean and the slope of the response along with combined arrays. Vuchkov and Boyadjieva (1992) presented mean and variance models for the situation where the random error is present in both product and noise factors. Brenneman and Myers (2003) consider robust design problems where the noise variables are categorical, an important consideration in designing robust processes where the categorical noise variable might indicate the supplier, for example. Engel and Huele (1996) considered second-order response surfaces in the factors. Steiner and Hamada (1997) extended the idea of robust design to mixture experiments. They considered two cases. In the first case, there are both mixture and process variables and some of the process variables are noise variables. The second case deals with error or noise in formulation of the mixture components. Goldfarb, Borror, and Montgomery (2003) also consider robust design for mixture-process situations, using a more general model than in Steiner and Hamada (1997).

Other Aspects of Mean and Variance Modeling

Several recent articles addressed other variations of the RPD problem and different approaches to modeling the mean and variance. Vining and Bohn (1998) presented a nonparametric approach to dual response modeling. Kim and Lin (1998) illustrated dual response surface optimization using fuzzy modeling. Kuhn, Carter, and Myers (2000) modeled the mean and variance for time-to-event data. Vining and Schaub (1996) and Chipman (1998) discussed Bayesian approaches. Some authors have considered the RPD problem for situations where the response is a relationship such as a curve or profile instead of a single quantity. Taguchi referred to this as a "dynamic" characteristic, although the terminology "signal-response system" has become more standard. Good discussions and examples of this problem can be found in papers by Lunani, Nair, and Wasserman (1997), Miller and Wu (1996), and Miller (2002).

There are some experimental situations where non-homogenuous variance, not necessarily due to the noise factors, is a problem. Nonconstant variance may result from different sources, which leads to location and dispersion effects. A particular variable is considered a dispersion effect if as the level of that variable is changed the *variance* in the response also changes.

Several authors have addressed the issue of estimating dispersion effects. Bergman and Hynen (1997) discussed identifying dispersion effects in unreplicated two-level fractional factorial experiments. Blomkvist, Hynen, and Bergman (1997) extended the method proposed by Bergman and Hynen (1997) to multilevel experiments. Ferrer and Romero (1993, 1995) used transformed residuals to investigate the process variance, while Engel (1992), Ghosh and Duh (1992), and Rosenbaum (1994, 1996) analyzed summary measures over the levels of the

noise variables. Steinberg and Bursztyn (1994, 1998) discussed direct modeling of the location and dispersion effects. Steinberg and Bursztyn (1994, 1998) and Berube and Nair (1998) concluded that if noise factors are present, then one should explicitly model their effects in order to effectively estimate dispersion effects. Analysis of summary measures (such as signal-to-noise ratios) in these situations tends not to reveal dispersion effects as readily as direct modeling of the noise factors. Wolfinger and Tobias (1998) recommended simultaneous estimation of location, dispersion, and random effects. Mays and Myers (1996) presented a Bayesian approach for design and analysis of factorial experiments when dispersion effects are present. McGrath and Lin (2001a) demonstrated the need for estimation methods that can efficiently separate location effects from dispersion effects. They presented methods to determine the type of effect, location, or dispersion for when active effects are discovered. The authors also pointed out that if the mean model is incorrectly specified, then using residuals to identify dispersion effects can give misleading results. McGrath and Lin (2001b) developed a method for testing for multiple dispersion effects in unreplicated designs. Brenneman and Nair (2001) reviewed and evaluated many of the methods available for identifying dispersion effects from unreplicated designs. They noted that all methods suffer from large biases. They proposed iterative methods for model selection and estimation of dispersion effects.

Response Surface Designs

Since Box and Wilson (1951), substantial progress has been made in the area of response surface designs for both first- and second-order models. Many of the areas of research through the late 1980s fall into the following categories:

i) development of standard designs and selection of design parameters;

ii) development of design criteria and design properties;

iii) the practical implementation of design optimality or "computer generated designs;"

iv) and design robustness.

Consequently, there are today many design tools that allow the practitioner to construct response surface experiments for both screening and optimization. Many software packages generate the standard designs, including regular factorial and fractional factorial plans, Plackett-Burman, central composite, small composite, Box-Behnken, and hybrid designs. The application of computer generated designs based primarily on the D criterion has grown substantially. Myers, Khuri, and Carter (1989) discussed much of the interesting research that had appeared on design robustness. However, the appeal made by Box and Draper (1975) for the use of design robustness rather than design optimality has been relatively ignored by software designers. Using design optimality with a single criterion (perhaps D-optimality) is the antithesis of design robustness, but, because the design construction methods are straightforward to implement, the commercial products available continue to emphasize this approach.

Importance of Design Robustness, Sequential Designs, and Bayesian Design

The practitioner's easy access to an "optimal" response surface design often results in the use of design optimality when standard designs may be more robust. The importance of design robustness cannot be overemphasized. Indeed, the expansion of RSM into relatively new areas makes design robustness even more important today. The areas in which design robustness must play a role include (a) model misspecification, (b) outliers, and (c) errors in control. It is becoming clear that RSM is moving into nonlinear modeling, specifically generalized linear models. In this area, good designs, and certainly optimal experimental designs, depend on knowledge of the parameters of the model. So, obviously, robustness to initial parameter guesses becomes an important issue.

Bayesian design has and will continue to play a vital theoretical role in response surface design. Its use may be necessary when there are uncertainties in model parameters and model selection. For the latter case, DuMochel and Jones (1994) laid a foundation on which other researchers can build. They assumed that there are two types of model terms, *certain* terms and *potential* terms, and placed a prior distribution on the potential terms. They then constructed a Bayesian $D-$optimal design that maximizes the determinant of the posterior information matrix. For the case of uncertainties in model parameters, Chanoler and Verdinelli (1995) reviewed the Bayesian approach to design. Andere-Rendon, Montgomery, and Rollier (1997) used this approach to construct Bayesian $D-$optimal designs for mixture experiments. Lin, Myers, and Ye (2000) utilized a

two-stage approach to Bayesian design where the prior information is updated at the completion of the first stage.

The origins of Bayesian design are in nonlinear modeling in biological problems (almost exclusively for the case of a single independent variable). However, its use in response surface applications in both chemical and biomedical applications has become important. To a large extent, many aspects of modern RSM call for the construction of designs based on information that may not be known, thus the emphasis on design robustness and Bayesian design. For examples, see Jia and Myers (2001), Sitter and Torsney (1995), and Atkinson and Haines (1996). Vining and Schaub (1996) used Bayesian based procedures to determine appropriate response designs for dual response modeling situations, such as those described in the second section of this paper.

Another very important and very effective approach uses sequential design. Such an approach dates back to Box and Wilson (1951), who suggested that the central composite design be deployed sequentially, with the first stage being a 2-level factorial or fractional factorial design and the axial points forming a second stage. The axial points are used if curvature is found in the system by a lack-of-fit test. While there are few formal sequential or multiple stage designs, this approach should be considered by statistics researchers. Any experienced professional in the field of design has certainly experienced "in retrospect" views that are critical of a chosen design strategy. Box (1999) clearly espoused such a view. The use of multiple stage designs often results in more efficient experiments. Further advances in robust design are credited to Borkowski and Valeroso (2001), who studied standard second order designs with regard to their efficiency in accommodating various posterior subset models.

Designs For Qualitative and Quantitative Variables

Often, in practice, response surface studies involve both qualitative and quantitative factors. Several commercial software packages do distinguish between qualitative and quantitative factors, and thus D-optimal designs for these situations can be found. Qualitative factors must be included in the model and the design, and model interpretation, prediction, and optimization must include consideration of the role of qualitative factors.

Choice of an appropriate design requires one to consider carefully how the qualitative factors interact with the model terms that involve the quantitative factors. These decisions have a potentially dramatic effect on the ability to adequately represent the response surface in the quantitative factors at different combinations of levels of the qualitative factors. At times it is difficult to anticipate the nature of these interactions, but including all possible interactions may substantially increase the size of the design. For example, in the case of two quantitative variables x_1 and x_2 and a qualitative factor at two levels represented by a categorical variable z, possible model terms are x_1, x_2, x_1^2, x_2^2, x_1x_2, zx_1, zx_2, zx_1^2, zx_2^2, and zx_1x_2. If the number of quantitative and qualitative variables is large, then the number of model terms that must be estimated can become quite large. Obviously, an appropriate design is to *cross* two designs, one of which is appropriate for the quantitative variables and one of which is appropriate for the qualitative variables. However, this design will often be too large in many cases. Draper and John (1988) discussed this problem and showed examples of more efficient designs. Myers and Montgomery (2002) presented illustrations with a numerical example of a complete RSM analysis. Atkinson and Donev (1992) presented algorithms for constructing near D-optimal designs for experiments involving both qualitative and quantitative factors. Aggarwal and Bansal (1998), Aggarwal, Gupta, and Bansal (2000), and Wu and Ding (1998) illustrated other approaches for the construction of response surface designs for both qualitative and quantitative factors.

Other New Designs

Recent progress in the development of small first-order designs has been made with so-called supersaturated designs, or designs that employ fewer data points than number of factors. Here, of course, the purpose of the experiment is purely one of variable screening. In practice, the designs are used in conjunction with exploratory information for assessing first order effects. Generally, the designs were developed much like Plackett-Burman designs but with a view toward attractive projection properties. Supersaturated designs have a long history beginning with Satterthwaite (1959) and Booth and Cox (1962). More recent work includes papers by Lin (1993a, 1993b, 1995), Wu (1993), Draper and Lin (1995), Deng, Lin, and Wang (1996a,b), Nguyen (1996), Li and Wu (1997), Westfall, Young, and Lin (1998), Balkin and Lin (1998), and Yamada and Lin (1997). These authors discussed such issues as design

construction, evaluation, and analysis methods. Lin (2000) gave a concise overview of recent developments. Holcomb and Carlyle (2002) described the construction of supersaturated designs in terms of computational complexity and proposed some new methods for evaluation and construction of designs. Abraham, Chipman, and Vijayan (1999) discussed the analysis of supersaturated designs and pointed out that the practical interpretation of effect estimates may be very difficult. Beattie, Fong, and Lin (2002) proposed a two-stage Bayesian model selection strategy, and Holcomb, Montgomery, and Carlyle (2003) developed a contrast-based analysis technique for supersaturated designs that provides more reliable results that the usual stepwise-regression based techniques.

Some work explores the possibilities for following up on a 2^{k-p} experiment by adding runs to de-alias effects of potential interest. For example, Mee and Peralta (2000) described semifolding, a technique using half of a standard fold-over design (see also Barnett, Czitrom, John, and Leon (1997)). Meyer, Steinberg, and Box (1996) used a Bayesian criterion for design augmentation. Chipman and Hamada (1996) advocated an effect-based approach and illustrated how the follow-up design selected depends on the family of models selected. Montgomery and Runger (1996) gave rules for optimal foldovers of Resolution IV designs and presented some new foldovers for Resolution III designs. Li and Mee (2002) provided some new results on fold-over of resolution III designs. Nelson et al. (2000) compared augmentation strategies for both 2^{k-p} and Plackett-Burman designs. The method of characterizing projected designs after an initial analysis has identified a subset of the original factors was investigated by Draper and Lin (1995), Box and Tyssedal (1996, 2001), Cheng (1995), and Wang and Wu (1995).

Draper (1985), Draper and Lin (1990), and Lin and Draper (1992) made a successful case for the use of Plackett-Burman designs for the factorial portion of Hartley's small composite designs. In some cases, the design results in a smaller run size than the standard small composite design (SCD). For example, in the case of $k = 7$ design variables, the use of a standard SCD requires a 2^{7-2} fraction plus 14 axial points for a total of 46 points plus centerpoints. On the other hand, the use of a 24 run Plackett-Burman design for the factorial portion requires $24 + 14 = 38$ points, which is two above saturation, plus center points. In addition, the use of a Plackett-Burman design allows a compromise between a saturated SCD and a CCD. For example, for $k = 4$ the standard SCD requires 16 factorial plus center and axial runs, while the use of a Plackett-Burman 12 run design plus axial and center runs allows 6 lack-of-fit degrees of freedom resulting from a total of 21 design points. The complete central composite design requires 24 runs plus the center runs. Furthermore, it was observed by a referee that because Plackett-Burman designs project down to a limited selection of type of designs in fewer dimensions, this makes it possible to run, for example, a 12-run Plackett-Burman design in 11 factors, project down to one of only two possible types in any five factors, and then add axial runs to estimate the parameters of a second-order model in five factors very efficiently (see Draper (1985)). Lin and Draper (1993) discussed generating the alias relationships for the two-level Plackett-Burman designs.

Morris (2000) proposed a new class of three-level designs for fitting the second-order response surface model. He referred to these as augmented pairs designs, because they are constructed by forming the third level of a factor as a linear combination of the levels of every pair of points in a two-level first-order design. The maximum and minimax distance criteria introduced by Johnson, Moore, and Ylvisaker (1990) and used in spatial modeling motivate this construction scheme. The number of runs in these designs for $3 \leq k \leq 10$ factors is larger than in the small composite design (except for $k = 3$, where they are identical), and less than in the Box-Behnken design (except for $k = 4$) and the central composite design (except for $k = 4$ and 5). Morris (2000) showed that the precision of model parameter estimates and the precision of expected response estimation for augmented pairs designs compares favorably to central composite and Box-Behnken designs.

Cheng and Wu (2001) suggested combining the factor screening optimization stages of an RSM study by using 3^{k-p} designs for both screening and optimization. The projection properties of the 3^{k-p} are utilized to provide second-order designs in the important factors. The authors defined projection efficiency criteria for these designs and studied their performance. They noted that the D-efficiencies of the projected designs is generally good, but that the G-efficiencies are worse. They noted that some nonregular designs may enjoy better projection properties. Several discussants of this paper commented on the desirability and potential risks of combining screening and optimization in a single stage, including the choice of an appropriate experi-

mental region, adequacy of the quadratic model if the region is too large, and the possible necessity of adding variables to the experiment between the screening and optimization stages. In their discussion of Cheng and Wu (2001), Montgomery and Borror (2001) noted that with the addition of a small number of runs the G-efficiency of the projected designs can be significantly improved. Lawson (2003) provided two illustrations of one-step screening and optimization experiments.

There have been a number of useful extensions involving well-known response surface designs. Mee (2001) discussed noncentral composite designs formed by adding a second two-level design to an initial 2^k or 2^{k-p} experiment. He also considered asymmetric composite designs consisting of a factorial portion and k axial runs that originate from a new base point not at the design center. He identified several scenarios in which these designs may prove useful. Mee (2002) discussed the use of three-level simplex designs in sequential experimentation, and noted that a number of new second-order designs can be formed in this manner. Ankenman, Liu, Karr, and Picka (2002) proposed a class of split factorial designs that are useful for estimating both a response surface and variance components arising from nested random effects.

The rotatability criterion continues to attract the attention of researchers. Draper, Gaffke, and Pukelsheim (1991) and Draper and Pukelsheim (1994) studied the topic in a very general way, investigating rotatability of designs, information variance surfaces, and moment matrices. Related work includes that by Draper, Gaffke, and Pukelsheim (1993) and Draper, Heiligers, and Pukelsheim (1996). Measures of rotatability were proposed by Khuri (1988, 1992) and Draper and Pukelsheim (1990). These measures are single-number indices that describe the departure of a given design from rotatability.

Uniform designs (Fang et al. (2000)) are another new class of experimental designs that are potentially useful in RSM. These are space-filling designs in which runs are scattered in a deterministic manner over the design space. Because these designs typically have a large number of levels they can be used to fit models considerably more complex than the usual second-order models typically employed in the optimization phase of an RSM study. One of the responses in the example in Fang et al. (2000) involved an 8 order polynomial. Freeny and Landwehr (1995) discussed graphical approaches to the analysis of large designs, and Lewis and Dean (2001) proposed methods for the detection of interactions in large designs.

Draper et al. (1994) showed how the degrees of freedom in a Box-Behnken design beyond those required to fit the model can be used to check model adequacy by forming contrasts in the higher-order terms. When they are orthogonal, these contrasts can be plotted on a normal probability plot to detect model bias.

There have been advances in methods for constructing optimal response surface designs. Haines (1987) and Zhou (2001) applied simulated annealing and Welch (1982) developed a branch and bound algorithm for the construction of alphabetically optimal designs. Both of these techniques are general methods for the solution of complex optimization problems. Recently, genetic algorithms (GAs) have been applied to design construction. An important feature of GAs is that they do not require a candidate set of points from which to select the design, so they can often find designs that are superior to those produced by other methods that require a candidate set. They can also be very efficient for large problems when compared to exchange methods employing a very fine grid of candidate points. Montepiedra, Myers, and Yeh (1998) used an abbreviated version of the GA, omitting the mutation step, and employing the traditional binary gene encoding. These decisions tend to limit the performance of the algorithm, and the designs that result are not generally the best that could have been obtained. Hamada et al. (2001) used the GA to find designs that are optimal in a Bayesian sense. These authors used a full GA procedure and real encoding. They were able to provide designs that would be extremely difficult to generate with traditional exchange algorithms. Heredia-Langner et al. (2003) applied the GA to a number of different response surface design problems, including constrained experimental regions. Their designs are often superior in terms of alphabetic optimality criteria than designs found by exchange algorithms. Borkowski (2003a) used a GA to generate near-optimal D, A, G, and IV exact N-point second-order response surface designs on hypercubes, and provided a catalog of designs for 1, 2, and 3 factors.

In many applications of RSM, the final model fit to the experimental data is different than the model for which the experiment was designed. It is well-known that design optimality properties may suffer quite

substantially when the experimenter fits a model different than the one originally thought to be appropriate. The work by Borkowski and Valeroso (2001) showed that many RSM designs are not robust across reduced models to the D, G, A, and IV criteria. Myers and Montgomery (2002) gave an example involving a reduced second-order model in four factors in which the design for the initial model is singular if the experimenter decides subsequently to fit a slightly different reduced second-order model. Situations such as these identify the need for designs that are model robust (that is, those which maintain good efficiency with respect to design optimality criteria across a range of potential models of interest to the experimenter). Some work has been done in this area; indeed, the Bayesian design work by DuMouchel and Jones (1994) and Andere-Rendon, Montgomery, and Rollier (1997) produced designs that are model-robust, in the sense that they allow one to fit a primary model while protecting against a potential model of higher order. Li and Nachtsheim (2000) proposed a class of model-robust factorial designs that are optimal over a group of models consisting of all main effects and a certain number of two-factor interactions. They modified an exchange algorithm to generate their designs. Heredia-Langner et al. (2004) showed how to use a genetic algorithm to create response surface designs that are model-robust for a wide variety of problem types, including designs for constrained regions, mixture experiments, and mixture-process experiments. They showed that their approach is generally superior to the relatively common practice of designing the experiment to fit the most elaborate model in the family of models of potential interest.

Methods for Evaluating Response Surface Designs

The use of optimal design theory for RSM design selection forces the use of a single criterion and, hence, a single philosophy. Certainly, no simple prescription can be satisfactory in all applications. This fact represents a continuing downside to the use of computer-generated designs, most of which depend on a single number criterion, such as D-optimality. Successful RSM designs need to be based on many considerations (for example, see the 14 criteria listed by Box and Draper (1975, 1987)). The importance of design robustness underscores this concept.

The development of graphical methods for comparing and evaluating designs reflects the need to take a multidimensional aspect of a RSM design into account. Many of these graphical methods evolved from the use of prediction variance; that is, the variance of the predicted response

$$\widehat{y}(\mathbf{x}) = f'(\mathbf{x})\widehat{\boldsymbol{\beta}}, \tag{7}$$

where $f'(\mathbf{x})$ reflects the polynomial model terms in the design variables $x_1, x_2, \ldots, x_k$, and

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \tag{8}$$

is the standard least squares estimator of regression coefficients when the response surface model is written in the general linear model form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{9}$$

where $\boldsymbol{\varepsilon}$ is a vector of random errors assumed to be i.i.d. with mean 0 and variance σ^2. It is a standard practice to use the scaled prediction variance in comparing designs. The prediction variance is

$$\mathrm{Var}[\widehat{y}(\mathbf{x})] = f'(\mathbf{x})(\mathbf{X}'\mathbf{X})^{-1}f(\mathbf{x})\sigma^2, \tag{10}$$

and the scaled prediction variance is

$$\frac{N\mathrm{Var}[\widehat{y}(\mathbf{x})]}{\sigma^2} = Nf'(\mathbf{x})(\mathbf{X}'\mathbf{X})^{-1}f(\mathbf{x}), \tag{11}$$

where N is the design size. The scaled prediction variance depends on the design, the model, and the location at which the prediction is to be made. A "good" design can be defined as one in which the scaled prediction variance is reasonably stable in the design region. This approach certainly underscores the preference of a plot of $\mathrm{Var}[\widehat{y}(\mathbf{x})]/\sigma^2$ or $N\mathrm{Var}[\widehat{y}(\mathbf{x})]/\sigma^2$ to a single number criterion such as the D-efficiency. A design that is produced by the computer as a D-optimal design may well have a very unstable distribution of $N\mathrm{Var}[\widehat{y}(\mathbf{x})]/\sigma^2$ in important portions of the region of interest.

Several software packages, such as Design-Expert, now provide contour plots of the scaled prediction variance as well as the estimated prediction variance computed from Equation (10), with σ^2 replaced by an estimate $\widehat{\sigma}^2$. Obviously, plots of Equation (11) prior to taking data are preferable so that various competing designs can be evaluated. Many software packages also provide an average prediction variance measure over the design region in the output. Borkowski (2003b) pointed out that one must be careful in interpreting this measure relative to the integrated variance or IV design optimality criterion. When a fixed set of points are used to calculate the

average prediction variance, the quantity will be less than the true integrated variance and could lead to the selection of an inferior design relative to the IV criterion.

The Variance Dispersion Graph

Even for as few as three factors, contour plots become difficult to use and to interpret. In such cases, Giovannitti-Jensen and Myers (1989) and Myers et al. (1992) proposed the variance dispersion graph (VDG), which plots the prediction variance information for the entire design region on a two-dimensional graph. These plots display the maximum, minimum, and "spherical," or average, prediction variances over spheres of radius r extending from the design origin to the design perimeter. The average prediction variance over a sphere of radius r is given by

$$V_r = \frac{N \int\limits_{\nu_r} \mathrm{Var}[\widehat{y}(\mathbf{x})]d\mathbf{x}}{\int\limits_{\nu_r} d\mathbf{x}}, \qquad (12)$$

where ν_r is the set of points in the region of interest that satisfy $\sum_{i=1}^{k} x_i^2 = r^2$, and $\int_{\nu_r} d\mathbf{x}$ is the volume of the region. The plot of V_r is then augmented by plots of

$$\max_{\nu_r}\left\{\frac{N\mathrm{Var}[\widehat{y}(\mathbf{x})]}{\sigma^2}\right\} \quad \text{and} \quad \min_{\nu_r}\left\{\frac{N\mathrm{Var}[\widehat{y}(\mathbf{x})]}{\sigma^2}\right\}.$$

The use of the scaled prediction variance $N\mathrm{Var}[\widehat{y}(\mathbf{x})]/\sigma^2$ allows the use of G-efficiency as a frame of reference for the design. In addition, the plot of the maximum scaled prediction variance allows a direct assessment of the stability of prediction variance. Software for generating VDGs for spherical regions can be found in papers by Vining (1993) and Borror (1998). Myers et al. (1992), Borkowski (1995), and Myers and Montgomery (2002) illustrated how one can use VDGs to compare and evaluate response surface designs.

The VDG has been extended to other design scenarios. Software for constructing VDGs for cuboidal designs was presented by Borror (1998). Trinca and Gilmour (1998) modified the original VDG to incorporate the effects of running a response surface design in blocks. Borror, Montgomery, and Myers (2002) developed a VDG methodology for designs for RPD and process robustness studies that incorporate both control and noise variables. They produced VDGs for both the mean and slope of the response surface model. Vining, Cornell, and Myers (1993) proposed and illustrated a VDG methodology that allows plots of prediction variance that are of particular interest in an RSM problem involving mixture designs and mixture models. Their VDGs are constructed along the Cox direction (see Cornell (2002)). Piepel and Anderson (1992) and Piepel, Anderson, and Redgate (1993a,b) considered a different approach involving plotting prediction variance contours on successively smaller polytopes that lie inside the original design region. This alternative can be a useful companion plot to that of Vining, Cornell, and Myers (1993).

Khuri, Kim, and Um (1996) built on the foundation of variance dispersion graphics by constructing quantile plots that are designed to provide more information than presented in the original VDG. They plotted quantiles of the distribution of $\mathrm{Var}[\widehat{y}(\mathbf{x})]/\sigma^2$ instead of only the average, maximum, and minimum values. They presented an example where the standard VDGs for two designs do not differ remarkably, but the quantile plots reveal very different prediction variance properties. Khuri, Harrison, and Cornell (1999) applied these quantile plots to the design of a mixture experiment. Another important recent development is the introduction of fraction of design space (FDS) graphs by Zahram, Anderson-Cook, and Myers (2003). The FDS graph plots the range of scaled prediction variance against the volume of the design region, expressed as a fraction. This provides information about the distribution of scaled prediction variance in the design space and gives additional insight into the predictive capability of a design that is not easily determined from the VDG.

Multiple Responses

Design and Modeling

Many designed experiments involve more than one response. In some industrial settings, such as semiconductor manufacturing, 12 to 15 response variables are not unusual. The robust parameter design problem discussed above is a special case of this problem where there are two responses, the mean and variance of a fundamental response observed during the experiment. Khuri and Cornell (1996) proposed the seemingly unrelated regressions method to estimate the model parameters, because the responses might be correlated. When the same design is used for all responses and the same model degree and form is used for all responses, this is the same as

separate modeling of each individual response. The review paper by Khuri (1990) contained a good summary of techniques for the analysis of multiresponse experiments. Langsrud (2001) presented some methods for identifying significant effects on multiresponse fractional factorial experiments.

Kim and Draper (1994) discussed selecting designs for fitting a first-order model in one variable when there are two correlated responses. They observed that when the responses are positively correlated one should not change the design much from the uncorrelated case, but one should make the design smaller in the case of negative correlation. They also reported similar results for two responses and two predictors and for three response and either one or two predictors. More general D-optimal designs for multiple-response regression models were considered by Kraft and Schaefer (1992). They showed that under some conditions the optimal designs do not depend on the response covariance matrix.

Multiple Response Optimization

A common approach first models each response individually and then graphically superimposes the contour plots. The analyst examines the overlay to find the appropriate set of operating conditions for the process. Rarely does a single set of operating conditions produce the optimum for all responses simultaneously. As a result, one usually searches for a set of conditions (sometimes called the "sweet spot") that comes as close as possible to satisfying all response requirements. Lind, Goldin, and Hickman (1960) illustrated this approach.

In practice, the graphical approach is limited to two or perhaps three dimensions. Therefore, there is considerable interest in more general approaches. One of these is to formulate the multiple response problem as a constrained optimization problem. In general, we suppose that there are m responses, and that for each we have a response surface model $\widehat{y}_i = f_i(\mathbf{x})$. We select one response as the primary response (or the objective function) and solve

$$\begin{aligned} &\max(\min)\widehat{y}_1 \\ &\text{subject to} \\ &l_i \le \widehat{y}_i \le u_i, \quad i = 2, 3, ..., m \\ &\mathbf{x} \in R^k, \end{aligned} \tag{13}$$

where the first response is the objective function, and l_i and u_i are lower and upper bounds on the remaining responses, $i = 2, 3, \ldots, m$. The last constraint in Equation (13) is a "region" constraint ensuring that the solution vector $\mathbf{x}_s$ is inside the experimental region R. For cuboidal designs, this constraint usually takes the form $-1 \le x_i \le 1$, $i = 1, 2, \ldots, k$ (the number of design variables). For spherical designs, the constraint is $\mathbf{x}'\mathbf{x} \le r^2$, where r is the design radius. There are commercially-available experimental design software packages that allow the user to formulate the multiple response optimization problem in this manner. It should be noted that in the special case where the f_i are convex functions, and there are no lower bounds on the constraint functions, the result is a convex optimization, which is a well-solved problem (see, for example, Nesterov and Nemirovskii (1994) and their many references). The majority of instances do not satisfy these conditions for convexity. Even so, in some cases the models can be altered slightly to bring about a convex optimization problem, with little impact on the quality of the optimal solution. The double linear regression method for identifying zero eigenvalues in the canonical form of the second-order model proposed by Ankenman and Bisgaard (1996) is useful in this regard.

There are two broad classes of methods that can be used to solve this problem, direct search methods and mathematical optimization algorithms. Most experimental design software packages that handle multiple response optimization use either the pattern search method of Hooke and Jeeves (1961) or the sequential simplex (see Copeland and Nelson (1996) for a modern presentation in the context of multiple response optimization).

Direct search methods are hill-climbing methods. They start from an initial solution and move in the local gradient direction, or an appropriate approximation of it. Frequently, there are several disjoint feasible operating regions for the process, which causes multiple local optima. Thus, there is no guarantee that one will find the absolute or global optimum solution. In fact, we are usually interested in finding all of the optimum solutions because some solutions may be more desirable than others based on practical considerations. For example, some of the feasible operating regions will typically be larger, which implies more robust operating conditions for the process. Most software packages use multiple starting locations, which somewhat alleviates this problem.

The generalized reduced gradient (GRG) method (see Lasdon, Fox, and Ratner (1974) and Del Castillo

and Montgomery (1993)) is a very popular mathematical optimization algorithm. This method is broadly applicable to many types of mathematical programming problems, and it is widely available in commercial software. The spreadsheet package EXCEL employs a version of the GRG, which is very convenient. Since the GRG makes explicit use of derivatives, it is considerably more efficient than search methods, and may be less likely to be confused by numerous local optima. For more discussion of this method, see Carlyle, Montgomery, and Runger (2000).

There are many possible variations of the optimization problem formulated in Equation (13). Del Castillo (1996) presented a formulation of the multiple response optimization problem that allows one to obtain solutions that simultaneously satisfy confidence region constraints for all responses. His methodology is applicable to both linear and quadratic models; however, he assumes the use of a rotatable design. Del Castillo, Fan, and Semple (1999) and Fan (2000) devised algorithms for global or near-global optimization of the dual response system.

The third general approach to multiple response optimization is to simultaneously optimize all m responses. An obvious possibility is to form some function combining the responses, such as $\widehat{y} = \sum_{i=1}^{m} a_i \widehat{y}_i$, where the a_i are weights, and then optimize the *composite* response $\widehat{y}$. In practice, choosing the weights appropriately is usually difficult, and so this approach is not widely used unless there is some unequivocal way to select the weights.

The most popular approach using computer software is the desirability function approach proposed by Derringer and Suich (1980). In this approach, each response is converted into an individual desirability function, say $d_i = h(\widehat{y}_i)$, where the desirability is $d_i = 0$ if the response is in an unacceptable range, $d_i = 1$ if the response is at the optimum value, and $0 < d_i < 1$ otherwise. It is possible to determine the shape of the desirability function, thereby controlling how important it is that the response achieve the target in the final optimal solution. Then, the solution $\mathbf{x}_s$ is found that maximizes the geometric mean of the individual response desirabilities,

$$D = \left(\prod_{i=1}^{m} d_i \right)^{1/m}.$$

Direct search methods are used to find the optimum solution, because the individual desirability functions are not differentiable. This approach also suffers from the multiple-optima phenomena noted above.

Del Castillo, Montgomery, and McCarville (1996) demonstrated how to replace the non-differentiable desirability functions of Derringer and Suich (1980) with differentiable equivalents that are fourth-order polynomials. One can then use standard mathematical programming optimization algorithms such as the GRG to optimize the overall desirability function. Kim and Lin (2000) used an experimental form of the desirability function and illustrated its application to the simultaneous optimization of the mechanical properties of steel.

Problems with highly nonlinear, or multimodal, objective functions can be extremely difficult to solve, and are further complicated by the presence of multiple objectives. A standard approach in this situation is to use a heuristic search procedure. There are several families of methods that have been proposed for finding global optima for such problems, two of which are simulated annealing and genetic algorithms. Neither is guaranteed to find the global optimum, but they have been designed to use different means of avoiding getting stuck at local optima. Carlyle, Montgomery, and Runger (2000) discussed these approaches in more detail.

All of the multiple response optimization techniques we have discussed assume that the responses are independent or uncorrelated. This assumption is often inappropriate, and a method that takes the dependence among responses into account is desirable. Khuri and Conlon (1981) presented an approach in which a generalized distance measure is employed to indicate the weighted distance of each response from its individual optimum value. The variances and covariances of the responses are used in determining the weights. Then, the solution is found that minimizes the generalized distance. Vining (1998) established that the Khuri and Conlon approach is a special case of a weighted squared error loss function, and showed several other plausible weighting schemes. Pignatiello (1993) and Ames et al. (1997) proposed approaches based on squared error loss functions. Chiao and Hamada (2001) proposed a multiple response optimization procedure that takes correlation among the responses into account. They considered the covariance matrix of the responses as dependent on the experimental

factors and modeled the parameters of the response distribution in terms of these factors. They then found the setting of the factors that maximizes the probability that all responses simultaneously meet their respective specifications.

RSM and Generalized Linear Models

The generalized linear model (GLM), developed by Nelder and Wedderburn (1972) and discussed in detail by McCullagh and Nelder (1989) and Myers, Montgomery, and Vining (2002), is a unified modeling strategy that allows for maximum likelihood estimates of model parameters in situations where the responses are members of the exponential family. This includes the normal, gamma, exponential, Poisson, and binomial distributions. Myers (1999) provided a thorough discussion of the growing importance that GLMs play in response surface design, modeling, and analysis (see that paper for many useful details and some basic references). We concentrate here on work published since Myers (1999) and on some additional references that were not included in that paper.

Papers illustrating the application of GLMs to industrial problems include those by Brinkley, Meyer, and Lu (1996), Lewis, Myers, and Montgomery (1999-2000, 2001a), Heredia-Langner et al. (2000), O'Neill et al. (2000), Borror, Heredia-Langner, and Montgomery (2002), and Lee and Nelder (2003). Myers (1999) noted the connection between GLM and the RPD problem. In fact, the nonhomogeneous variance routinely found in RPD leads naturally to consideration of the GLM for modeling and analysis. Grego (1993), Nelder and Lee (1991), and Lee and Nelder (1998, 2003) discussed and illustrated the connections between robust parameter design and generalized linear models and provided numerous examples.

Lewis, Montgomery, and Myers (2001b) compared the GLM to traditional least squares analysis and showed that the length of the confidence interval on the mean response is a useful measure to evaluate the competing models. Often, the GLM results in shorter confidence intervals, implying more precise response estimation, which can be an important issue in RSM.

The literature that presents examples of the use of GLM in RSM typically makes use of standard designs (that is, factorials or central composite designs). Myers, Montgomery, and Vining (2002) used the notion of design efficiencies to argue that many of these standard designs will work very well with the GLM. However, most of the work has been confined to the first-order model and two-level designs. Consequently, we believe that more work needs to be done in the development of practical design strategy in the use of generalized linear models.

As we have noted, optimal (or even good) designs cannot be found for nonlinear models (including GLMs) without "guesses" of model parameters. As a result, most of the basic research into designs for GLM has been in local design optimality, or the construction of optimal designs under the assumption that the parameters are known. Clearly, this work must be a springboard to the development of creative Bayesian designs, robust designs, or multiple stage designs. Ford, Torsney, and Wu (1992), Sitter and Wu (1993), Kalish (1990), and Rosenberger and Kalish (1978) dealt with locally optimal designs for logistic regression in a single design variable. For the most part, these works involve $D-$optimality, G-optimality, and F-optimality. In this context, D-optimality deals specifically with designs that allow optimal estimation of the effective dose (ED), that is, the value of x that elicits a particular probability. The F(or Fieller)-optimality criterion selects the design that minimizes the squared width of the Fieller interval for estimation of a certain ED. Myers, Myers, and Carter (1994) gave a review of design optimality for logistic regression and introduced the notion of Q-optimality with the logistic regression model. While all of these represent single variable designs with local optimality, the stage was set for further research involving more pragmatic approaches. Atkinson, Chanoler, Herzberg, and Juritz (1993) developed Bayesian designs that are D- and D_s- (optimal for a subset) optimal for a type of compartmental model used in biological applications. Heise and Myers (1996) studied both optimal and robust designs for bivariate logistic regression. The natural setting for this case involves two responses, efficacy and toxicity. Clearly, applications exist for industrial as well as biological processes.

Other Issues and Applications

Response Surface Experiments with Randomization Restrictions

Many industrial experiments involve two types of factors: those with levels that are easy-to-change and

those with levels that are difficult or costly to change. Consider a sintering process for battery anodes involving different densities of nickel powder and different furnace temperatures. The temperature profiles in a furnace require a great deal of time to reach equilibrium. Experimenters prefer to obtain as much information as they can for a single setting of the temperature conditions; consequently, furnace temperature is a hard-to-change factor. On the other hand, operators can change nickel powder densities relatively easily. In this situation, a reasonable experiment fixes the level of furnace temperature and then runs all levels of density for the nickel powder. The experimenter then changes the setting of temperature and repeats the process. This experiment generates more information about nickel powder density, because it is replicated more often, than about furnace temperature. As a result, analysis involves two different error terms. Proper analysis of data must take this consequence of the restricted randomization scheme into account.

Nelson (1985) used a split-plot example involving furnaces to show how important it is to correctly specify the model as a split-plot model instead of a completely randomized design (CRD) model. Lucas and Hazel (1997) ran the helicopter experiment as a CRD and as a split-plot to illustrate the differences in analysis. Ganju and Lucas (1997, 1999) discussed inadvertent split-plotting, which occurs when factors are not independently reset for consecutive runs requiring the same level of the factor. Ganju and Lucas (1997) demonstrated the underestimation of the variance for the main effects of factors that are not reset. Ganju and Lucas (1999) showed the difficulty in detecting the effect of the randomization restrictions.

Letsinger, Myers, and Lentner (1996) introduced bi-randomization designs (BRDs). These BRDs are designs with two randomizations, similar to split-plot designs. Letsinger, Myers, and Lentner break BRDs into two classes, crossed and non-crossed. In crossed BRDs, every level of the whole plot factors is "crossed" with every level of the subplot factors, which produces the usual split-plot designs. For the non-crossed BRD, the whole plot factors have different levels of the sub-plots and need not have the same number of levels. The distinction between these two can be thought of in terms of the sub-plot factors. The crossed BRD might be represented by a 2^k factorial in the sub-plot factors, while the non-crossed BRD might use a 2^{k-p} fractional-factorial in the sub-plot factors. The authors compared three methods for analyzing the second order case: OLS, iterative reweighted least squares (IRLS), and restricted maximum likelihood (REML). Though IRLS and REML appear to be better methods, the "best" method depends on the design, model, and any prior information.

Minimum-aberration (MA) designs for split-plot experiments where both the whole plot and subplot factors have two-levels are discussed by Huang, Chen, and Voelkel (1998) and Bingham and Sitter (1999a). They provided methods for determining the MA designs and then gave tables for various combinations of whole plot and subplot factors. Some design issues with two-level fractional-factorial split-plot experiments, including where to split and where to fractionate, are presented by Bingham and Sitter (2001). Theoretical justification of these types of split-plot designs is given by Bingham and Sitter (1999b).

Lucas and Ju (1992) used a simulation study to investigate the use of split-plot designs in industrial experiments where one factor was difficult to change and served as the whole plot. Their results confirm that split-plot designs will produce increased precision on the subplot factors while sacrificing precision on the whole plot factors. Kowalski (1999) also studied split-plot designs and confounding issues using simulation. Box (1996) explained that completely randomized experiments are often impractical in industry, and indicated that split-plot experiments are often very efficient and easier to run.

Miller (1997) considered various fractional-factorial structures in strip-plot experiments. Strip-plot configurations can be applied when the process being investigated can be separated into two distinct stages and it is possible to apply the second stage simultaneously to groups of the first-stage product. Miller (1997) proposed a method for constructing strip-plot configurations for fractional-factorial designs. The method is applied for two-level designs and then extended to m-level and mixed-level designs. Mee and Bates (1998) considered split-lot experiments involving silicon wafers. These experiments are performed in steps where a different factor is applied at each step. Mee and Bates (1998) proposed a model and provided a general approach for constructing split-lot designs.

Draper and John (1998) considered modifications to the central composite and Box-Behnken designs based on rotatability when some of the factors are harder to change than others. Trinca and Gilmour

(2001) presented a method for designing response surface experiments that have multiple strata. Beginning with the highest stratum, designs are built stratum by stratum. Then, the trials are arranged so that they are nearly orthogonal to the units in the higher strata. This strategy produces optimal designs that minimize the maximum of a function of the variances of the parameter estimates.

Schoen and Wolff (1997) discussed the design and analysis of a mixed fraction split-plot experiment. They use confounding and a half-normal plot to determine active between-runs effects and within-runs effects. Schoen (1999) proposed a method for manipulating the division of contrasts over the different error strata of two-level experiments with nested errors. Goos and Vandebroek (2001) proposed an exchange algorithm for obtaining D-optimal split-plot designs. They showed that the design matrices for the D-optimal split-plot designs and D-optimal CRDs are typically different, and that split-plot experiments are often more efficient than CRDs. Others doing interesting work with split-plot experiments include Federer and Meredith (1992), Mathew and Sinha (1992), and Remmenga and Johnson (1995).

The analysis of data from mixture experiments containing process variables is presented by Cornell (1988). The factor-level combinations of the process variables are considered whole plot treatments and the mixture component blends the subplot treatments. Kowalski, Cornell, and Vining (2002) proposed new designs for split-plot experiments that have mixture components and process variables, which are much smaller than those discussed by Cornell (1988). They also used simulation to compare several estimation methods in terms of the size of the confidence ellipsoid around the parameters.

Industrial experimentation often involves the use of both hard-to-change and easy-to-change factors. Engineers routinely run split-plot experiments without fully realizing the implications for analysis. As a result, they often reach inappropriate conclusions from their experiments. More work needs to be done in extending basic split-plot analysis techniques to different types of industrial experiments.

Split-Plots in Robust Parameter Designs

Box and Jones (1992) considered three experimental arrangements for robust parameter design with the use of split-plot designs:

(a) the whole plots contain the environmental factors and the subplots contain the design factors;

(b) the whole plots contain the design factors and the subplots contain the environmental factors;

(c) and the subplot factors are assigned in 'strips' across the whole plot factors (this is commonly called a strip-block experiment).

They point out that, if at all possible, experimenters should pursue arrangement (a) because it gives more information about the design factors. When there is no replication, they suggest using two separate normal probability plots, one for the whole plot effects and another for the subplot and subplot by whole plot interactions, to detect significant effects.

Box and Jones (2001) pointed out that split-plot experiments are ideal for robustness studies because the interactions between design factors and environmental factors are estimated with the smaller subplot error. Bisgaard and Kulachi (2001) showed that more information is gained in RPD experiments from switching fractions (split-plot confounding) than from using the same fraction (inner-outer array designs). Based on this, they recommended using split-plot confounding in RPD to reduce the size of the experiment.

Bisgaard and Steinberg (1997) considered the design and analysis of prototype experiments. They presented examples that use split-plot designs and showed clearly how to carry out a two-stage analysis of the data. Bisgaard (2000) used inner and outer arrays, with factors at two-levels, and illustrated several ways to confound in order to reduce the number of experimental runs needed. By looking at these different alias structures, he showed that the goals of the experiment can dictate the appropriate defining contrasts. Bisgaard (2000) also provided the standard errors for various contrasts among the whole plot and subplot factors.

Kowalski (2002) also considered split-plot experiments in robust parameter design. He constructed 24-run designs in two ways: using the properties of a balanced incomplete block design, and semifolding a 16 run design. When there are only a few noise factors and a few design factors, the designs can estimate all main effects and almost all two-factor interactions. The designs are intended to be a compromise between 16 and 32 run screening designs.

There are times in RPD experiments where a measure of a product's quality is its worst perfor-

mance over the environmental conditions. Pan and Santner (1998) proposed two procedures for these situations. First, they determined the number of replicates of the split-plot design that are needed to assure, with a specified confidence level, the selection of the product design whose worst performance is within δ of the performance of the optimal product design. Second, if a particular split-plot experiment has already been conducted, they provided a procedure which will screen the product designs and provide a subset that contains the optimal product design with a specified confidence level.

Phadke (1989) presented an example, analyzed using Taguchi's SNRs, involving a polysilicon deposition process. The actual structure of this experiment is a split-split-plot design, because there are three sizes of experimental units with different sources of variation. Therefore, using Taguchi's SNRs to analyze this experiment results in a complete loss of information in the design by noise factor interactions. Cantell and Ramirez (1994) reanalyzed the data using a split-split-plot design.

Computer Experiments

Computer experiments are experiments performed on a computer model. Sometimes these computer models are stochastic simulations, and sometimes they are deterministic models of system behavior. Experimentation with a stochastic simulation model is much like experimentation with a physical system, although specific variance reduction techniques can often be employed to simplify analysis of the output. Some examples of RSM applications for such systems were given by Myers, Khuri, and Carter (1989). In a deterministic computer model, every detail of the system is fixed, explicitly or implicitly; consequently, there is no random error component associated with the output. Examples of such models include circuit design tools used in the semiconductor industry and finite element analysis models used in the design of discrete parts.

Many analysts feel that the model for a deterministic response should match the observed response at every design point as closely as possible. Furthermore, analysts often want to do a more global representation of the system rather than the local fitting standard in the RSM approach. As a result, the ranges of the design variables are generally larger than in standard RSM, and RSM designs may not be appropriate. Standard RSM designs are often characterized by spherical, cuboidal, or tetrahedral geometry in the case of mixture designs. In addition, the number of levels employed in the design is a function of the order of the model. However, in experiments with deterministic computer models, space filling designs are often used. This assumes that data points close to the point of interest enhance the prediction at that point. The Latin hypercube designs developed by McKay, Beckman, and Conover (1979) are widely used in such situations. See Welch, Yu, Kang, and Sacks (1990), Barton (1992), Donohue (1994), and Fang et al. (2000) for some approaches to design and response modeling. Many of the modeling techniques for deterministic computer experiments have their basis in nonparametric and semiparametric methods, which we discuss in the next section.

Nonparametric and Semiparametric Response Surface Methods

Nonparametric methods are applicable when a simple polynomial model in the region of interest cannot describe the response. The term "nonparametric RSM" implies that there is no specific model and that the primary focus of the analysis is prediction, which leads to the use of such techniques as kernel regression. Semiparametric RSM signifies that a model is used, but that it is not a standard polynomial model. As in deterministic computer experiments, these methods are potentially of interest when one wishes to fit a model that covers a larger portion of the factor space than is typical in standard RSM applications.

Applications of semiparametric or nonparametric response surface methods (NPRSM) have been couched in a way similar to traditional methods. The methods often employ 2-level factorial and fractional factorial experiments as well as central composite designs. In addition, analysis often uses steepest ascent, despite the fact that NPRSM seeks to achieve a global model of the process. Thus, the experiment covers the entire operability region rather than a local region of interest. For this reason, the space-filling designs mentioned in the previous section are used extensively.

The methods used for response modeling include (i) Gaussian stochastic process models, (ii) thin plate splines, and (iii) neural networks. The Gaussian stochastic process procedure was popularized by Sacks, Schiller, and Welch (1989) and Sacks, Welch, Mitchell, and Wynn (1989) in computer experiments. Welch, Buck, Sacks, Wynn, Mitchell, and Morris (1992) illustrated its use in the RSM scenarios

involving screening experiments and prediction. Once again, computer experiments provide the basis for their work. The methodology employs the model

$$y_i = f'(\mathbf{x}_i)\boldsymbol{\beta} + Z(\mathbf{x}_i) + \varepsilon_i \qquad (14)$$

for the i experimental run y_i, where $f'(\mathbf{x}_i)\boldsymbol{\beta}$ is a standard RSM model and $Z(\cdot)$ is a univariate Gaussian stochastic process on the design space. A correlation function of the type

$$\begin{aligned} R(\mathbf{x}_i, \mathbf{x}_j) &= Corr(Z(\mathbf{x}_i), Z(\mathbf{x}_j)) \\ &= \prod_k \exp\{-\theta_k |x_{ik} - x_{jk}|^\rho\} \end{aligned}$$

is assumed for $Z(\mathbf{x})$, where $\theta_k \geq 0$ and $0 \leq \rho \leq 2$.

The thin plate spline approach is similar to the Gaussian stochastic process model. It is a generalization of the classic polynomial model in which a smoothing parameter varies from a polynomial model at one limit to an interpolating spline at the other limit. There is a need for work that determines what designs are appropriate for different methods and conditions. Interesting distance-based designs are discussed by Johnson, Moore, and Ylvisaker (1990). See also Haaland, McMillan, Nycha, and Welch (1994).

Applications of NPRSM are abundant. As we indicated earlier in this section, the application roots are in computer experiments. Other applications include the optimization of reaction buffers to use in DNA for detection of infectious bacteria (see Spargo, Haaland, Jurgensen, Shank, and Walker (1993)). Space filling designs and neural networks are used to detect an injection molding process window by O'Connell, Haaland, Hardy, and Nychka (1995). These papers are a small sample of applications in a field that will doubtless grow to be a major area of RSM.

Concluding Remarks

The last fifteen years have seen extensive growth in both the underlying theory and practical applications of RSM. Advances have occurred on a broad front, including new designs and improved techniques for design construction, evaluation, and analysis. The solution of the robust parameter design problem in an RSM framework and the growing use of the generalized linear model as an analysis tool have been key developments. We expect that these two areas will continue to fuel much of the research in RSM. As the research of the last decade has shown, there is still much to do in the area of response surface designs, including the integration of randomization restrictions into RSM designs, efficient designs for larger problems, designs that are robust to the form of the final model chosen by the experimenter, designs for the GLM, and designs for computer experiments. Design, analysis, and optimization techniques for multiple responses will also continue to be an important area for research.

We also anticipate continued growth in RSM applications. Biological, biomedical, engineering design and development, computer-aided design, and transactional business processes are likely to be important areas of application. Because the application research problems are challenging and stimulating, RSM will remain one of the most active areas in applied statistics.

Appendix: List of Abbreviations

BRD	Bi-randomization design
CRD	Completely randomized design
GLM	Generalized linear model
GRG	Generalized Reduced Gradient
IRLS	Iteratively reweighted least squares
MA	Minimum aberration
NID	Normally and Independently Distributed
NPRSM	Nonparametric response surface methods
REML	Restricted maximum likelihood
RPD	Robust parameter design
RSM	Response surface methodology
VDG	Variance dispersion graph

Acknowledgments

The authors would like to thank the Editor, Professor Norman R. Draper, and an anonymous reviewer for their insightful comments on a earlier version of this paper.

References

ABRAHAM, B.; CHIPMAN, H.; and VIJAYAN, K. (1999). "Some Risks in the Construction and Analysis of Supersaturated Designs". *Technometrics* 41, pp. 135–141.

AGGARWAL, M. L. and BANSAL, A. (1998). "Robust Response Surface Designs for Quantitative and Qualitative Factors". *Communications in Statistics—Theory and Methods* 27, pp. 89–106.

AGGARWAL, M. L.; GUPTA, B. C.; and BANSAL, A. (2000).

"Small Robust Response-Surface Designs for Quantitative and Qualitative Factors". *American Journal of Mathematical and Management Sciences* 39, pp. 103–130.

AMES, A. E.; MATTUCCI, N.; MACDONALD, S.; SZONYI, G.; and HAWKINS, D. M. (1997). "Quality Loss Functions for Optimizations across Multiple Response Surfaces". *Journal of Quality Technology* 29, pp. 339–346.

ANDERE-RENDON, J.; MONTGOMERY, D. C.; and ROLLIER, D. A. (1997). "Design of Mixture Experiments Using Bayesian *D*-Optimality". *Journal of Quality Technology* 29, pp. 451–463.

ANKENMAN, B. E. and BISGAARD, S. (1996). "Standard Errors for the Eigenvalues in Second-Order Response Surface Models". *Technometrics* 38, pp. 238–246.

ANKENMAN, B. E.; LIU, H.; KARR, A. F.; and PICKA, J. D. (2002). "A Class of Experimental Designs for Estimating a Response Surface and Variance Components". *Technometrics* 44, pp. 45–54

ATKINSON, A. C.; CHANOLER, K.; HERZBERG, A. M.; and JURITZ, J. (1993). "Optimum Experimental-Designs for Properties of a Compartmental Model". *Biometrics* 49, pp. 325–337.

ATKINSON, A. and DONEV, A. N. (1992). *Optimum Experimental Design*. Clarendon Press, Oxford.

ATKINSON, A. C. and HAINES, L. M. (1996). "Designs for Nonlinear and Generalized Linear Models" in the *Handbook of Statistics* North Holland Publishing Co., Amsterdam, pp. 437–475.

BALKIN, S. D. and LIN, D. K. J. (1998). "A Graphical Comparison of Supersaturated Designs". *Communications in Statistics—Theory and Practice* 27, pp. 1289–1303.

BARNETT, J.; CZITROM, V.; JOHN, P. W. M; and LEON, R. V. (1997). "Using Fewer Wafers to Resolve Confounding in Screening Experiments" in *Statistical Case Studies for Industrial Process Improvement* edited by V. Czitrom and P.D. Spagon, SIAM, Philadelphia, PA., pp. 235–250.

BARTON R. R. (1992). "Metamodels for Simulation Input-Output Relations". *Winter Simulation Conference*, pp. 289–299.

BEATTIE, S. D.; FONG, D. K. H.; and LIN, D. K. J. (2002). "A Two-Stage Bayesian Model Selection Strategy for Supersaturated Designs". *Technometrics* 44, pp. 55–63.

BERGMAN, B. and HYNEN, A. (1997). "Dispersion Effects from Unreplicated Designs in the 2^{k-p} Series". *Technometrics* 39, pp. 191–198.

BERUBE, J. and NAIR, V. N. (1998). "Exploiting the Inherent Structure in Robust Parameter Design Experiments". *Statistica Sinica* 8, pp. 43–66.

BINGHAM, D. R. and LI, W. (2002). "A Class of Optimal Robust Parameter Designs". *Journal of Quality Technology* 34, pp. 244–259.

BINGHAM, D. R. and SITTER, R. S. (1999a). "Minimum-Aberration Two-Level Fractional Factorial Split-Plot Designs". *Technometrics* 41, pp. 62–70.

BINGHAM, D. R. and SITTER, R. S. (1999b). "Some Theoretical Results for Fractional Factorial Split-Plot Designs". *The Annals of Statistics* 27, pp. 1240–1255.

BINGHAM, D. R. and SITTER, R. S. (2001). "Design Issues in Fractional Factorial Split-Plot Experiments". *Journal of Quality Technology* 33, pp. 2–15.

BISGAARD, S. (2000). "The Design and Analysis of $2^{k-p} \times 2^{q-r}$ Split Plot Experiments". *Journal of Quality Technology* 32, pp. 39–56.

BISGAARD, S. and KULAHCI, M. (2001). "Robust Product Design: Saving Trials With Split-Plot Confounding". *Quality Engineering* 13, pp. 525–530.

BISGAARD, S. and STEINBERG, D. M. (1997). "The Design and Analysis of $2^{k-p} \times s$ Prototype Experiments". *Technometrics* 39, pp. 52–62.

BLOMKVIST, O.; HYNEN, A.; and BERGMAN, B. (1997). "A Method to Identify Dispersion Effects from Unreplicated Multilevel Experiments". *Quality and Reliability Engineering International* 13, pp. 127–138.

BOOTH, K. H. V. and COX, D. R. (1962). "Some Systematic Saturated Designs". *Technometrics* 4, pp. 489–495.

BORKOWSKI, J. J. (1995). "Spherical Prediction-Variance Properties of Central Composite Designs and Box-Behnken Designs". *Technometrics* 37, pp. 399–410.

BORKOWSKI, J. J. (2003a). "Using a Genetic Algorithm to Generate Small Exact Response Surface Designs". *Journal of Probability and Statistical Science* 1, pp. 65–88.

BORKOWSKI, J. J. (2003b). "A Comparison of Prediction Variance Criteria for Response Surface Designs". *Journal of Quality Technology* 35, pp. 70–77.

BORKOWSKI, J. J. and LUCAS, J. M. (1997). "Designs of Mixed Resolution for Process Robustness Studies". *Technometrics* 37, pp. 63–70.

BORKOWSKI, J. J. and VALEROSO, E. S. (2001). "Comparison of Design Optimality Criteria of Reduced Models for Response Surface Designs in the Hypercube". *Technometrics* 43, pp. 468–477.

BORROR, C. M. (1998). "Response Surface Methods for Experiments Involving Noise Variables". Ph.D. Dissertation, Arizona State University, Tempe, AZ.

BORROR, C. M.; HEREDIA-LANGNER, A.; and MONTGOMERY, D. C. 2002. "Generalized Linear Models in the Analysis of Industrial Experiments". *Journal of Propagations in Probability and Statistics* (International Edition) 2, pp. 127–144.

BORROR, C. M. and MONTGOMERY, D. C. (2000). "Mixed Resolution Designs as Alternatives to Taguchi Inner/Outer Array Designs for Robust Design Problems". *Quality and Reliability Engineering International* 16, pp. 117–127.

BORROR, C. M.; MONTGOMERY, D. C.; and MYERS, R. H. (2002). "Evaluation of Statistical Designs for Experiments Involving Noise Variables". *Journal of Quality Technology* 34, pp. 54–70.

BOX, G. E. P. (1985). "Discussion of 'Off-Line Quality Control, Parameter Design and the Taguchi Method". *Journal of Quality Technology* 17, pp. 198–206.

BOX, G. E. P. (1988). "Signal-to-Noise Ratios, Performance Criteria, and Transformations" (with discussion). *Technometrics* 30, pp. 1–41.

BOX, G. E. P. (1996). "Split-Plot Experiments". *Quality Engineering* 8, pp. 515–520.

BOX, G. E. P. (1999). "Statistics as a Catalyst to Learning by Scientific Method Part II—A Discussion". *Journal of Quality Technology* 31, pp. 16–29.

BOX, G. E. P.; BISGAARD, S.; and FUNG, C. A. (1988). "An Explanation and Critique of Taguchi's Contributions to

Quality Engineering". *Quality and Reliability Engineering International* 4, pp. 123–131.

BOX, G. E. P. and DRAPER, N. R. (1975). "Robust Designs". *Biometrika* 62, pp. 347–352.

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

BOX, G. E. P. and JONES, S. (1992). "Split-Plot Designs for Robust Product Experimentation". *Journal of Applied Statistics* 19, pp. 3–26.

BOX, G. E. P. and JONES, S. (2001). "Split-Plots for Robust Product and Process Experimentation". *Quality Engineering* 13, pp. 127–134.

BOX, G. E. P. and LIU, P. Y. T. (1999). "Statistics as a Catalyst to Learning by Scientific Method Part I—An Example". *Journal of Quality Technology* 31, pp. 1–15.

BOX, G. E. P. and TYSSEDAL, J. (1996). "Projective Properties of Certain Orthogonal Arrays". *Biometrika* 83, pp. 950–955.

BOX, G. E. P. and TYSSEDAL, J. (2001). "Sixteen Run Designs of High Projectivity for Factor Screening". *Communications in Statistics: Simulation and Computation* 30, pp. 217–228.

BOX, G. E. P. and WILSON, K. B. (1951). "On the Experimental Attainment of Optimum Conditions". *Journal of the Royal Statistical Society* B 13, pp. 1–45.

BRENNEMAN, W. A. and NAIR, V. N. (2001). "Methods for Identifying Dispersion Effects in Unreplicated Factorial Experiments: A Critical Analysis and Proposed Strategies". *Technometrics* 43, pp. 388–405.

BRENNEMAN, W. A. and MYERS, W. R. (2003). "Robust Parameter Design with Categorical Noise Variables". *Journal of Quality Technology* 35, pp. 335–341.

BRINKLEY, P; MEYER, K.; and LU, J. (1996). "Combined Generalized Linear Modelling-Non-linear Programming Approach to Robust Process Design—A Case-study in Circuit Board Quality Improvement". *Journal of the Royal Statistical Society Series* C 45, pp. 99–110.

CANTELL, B. and RAMIREZ, J. G. (1994). "Robust Design of a Polysilicon Deposition Process Using a Split-Plot Analysis". *Quality Reliability Engineering International* 10, pp. 123–132.

CARLYLE, W. M.; MONTGOMERY, D. C.; and RUNGER, G. C. (2000). "Optimization Problems and Methods in Quality Control and Improvement" (with discussion). *Journal of Quality Technology* 32, pp. 1–17.

CHANOLER, K. and VERDINELLI, I. (1995). "Bayesian Experimental Design: A Review". *Statistical Science* 10, pp. 273–304.

CHENG, C. S. (1995). "Some Projection Properties of Orthogonal Arrays". *Annals of Statistics* 23, pp. 1223–1233.

CHENG, S.-W. and WU, C. F. J. (2001). "Factor Screening and Response Surface Exploration" (with discussion). *Statistica Sinica* 11, pp. 553–604.

CHIAO, C.-H. and HAMADA, M. (2001). "Analyzing Experiments with Correlated Multiple Responses". *Journal of Quality Technology* 33, pp. 451–465.

CHIPMAN, H. (1998). "Handling Uncertainty in Analysis of Robust Design Experiments". *Journal of Quality Technology* 30, pp. 11–17.

CHIPMAN, H. and HAMADA, M. S. (1996). "Discussion: Factor-Based or Effect-Based Modeling? Implications for Design". *Technometrics* 38, pp. 317–320.

COPELAND, K. A. F. and NELSON, P. R. (1996). "Dual Response Optimization via Direct Function Minimization". *Journal of Quality Technology* 28, pp. 331–336.

CORNELL, J. A. (1988). "Analyzing Data from Mixture Experiments Containing Process Variables: A Split-Plot Approach". *Journal of Quality Technology* 20, pp. 2–23.

CORNELL, J. A. (2002). *Experiments with Mixtures* 3rd Ed. John Wiley and Sons, New York, NY.

DEL CASTILLO, E. (1996). "Multiresponse Process Optimization via Constrained Confidence Regions". *Journal of Quality Technology* 28, pp. 61–70.

DEL CASTILLO, E.; FAN, S. K.; and SEMPLE, J. (1999). "Optimization of Dual Response Systems: A Comprehensive Procedure for Degenerate and Nondegenerate Problems". *European Journal of Operational Research* 112, pp. 174–186.

DEL CASTILLO, E. and MONTGOMERY, D. C. (1993). "A Nonlinear Programming Solution to the Dual Response Problem". *Journal of Quality Technology* 25, pp. 199–204.

DEL CASTILLO, E.; MONTGOMERY, D. C.; and MCCARVILLE, D. (1996). "Modified Desirability Functions for Multiple Response Optimization". *Journal of Quality Technology* 28, pp. 337–345.

DENG, L. Y.; LIN, D. K. J.; and WANG, J. N. (1996a). "Marginally Oversaturated Designs". *Communications in Statistics* 25, pp. 2557–2573.

DENG, L. Y.; LIN, D. K. J.; and WANG, J. N. (1996b). "A Measure of Multifactor Nonorthogonality". *Statistics and Probability Letters* 28, pp. 203–209.

DERRINGER, G. C. and SUICH, R. (1980). "Simultaneous Optimization of Several Response Variables". *Journal of Quality Technology* 12, pp. 214–219.

DONOHUE, J. M. (1994). "Experimental Designs for Simulation". *Winter Simulation Conference* pp. 200–206.

DRAPER, N. R. (1985). "Small Composite Designs". *Technometrics* 27, pp. 173–180.

DRAPER, N. R.; DAVIS, T. P.; POZUETA, L.; and GROVE, D. (1994). "Isolation of Degrees of Freedom for Box-Behnken Designs". *Technometrics* 36, pp. 283–291.

DRAPER, N.; GAFFKE, N.; and PUKELSHEIM, F. (1991). "First and Second Order Rotatability of Experimental Designs, Moment Matrices, and Information Surfaces". *Metrika* 38, pp. 129–161.

DRAPER, N.; GAFFKE, N.; and PUKELSHEIM, F. (1993). "Rotatability of Variance Surfaces and Moment Matrices". *Journal of Statistical Planning and Inference* 36, pp. 347–356.

DRAPER, N.; HEILIGERS, B; and PUKELSHEIM, F. (1996). "An Optimal Third-Order Rotatable Designs". *Annals of Institute of Statistical Mathematics* 48, pp. 395–402.

DRAPER, N. and JOHN, J. A. (1988). "Response Surface Designs for Quantitative and Qualitative Variables". *Technometrics* 30, pp. 423–428.

DRAPER, N. and JOHN, J. A. (1998). "Response Surface Designs where levels of Some Factors Are Difficult to Change". *Australian and New Zealand Journal of Statistics* 40, pp. 487–495.

DRAPER, N. and LIN, D. K. J. (1990). "Small Response Surface Designs". *Technometrics* 32, pp. 187–194.

DRAPER, N. and LIN, D. K. J. (1995). "Characterizing Projected Designs- Repeat and Mirror Image Runs". *Communications in Statistics—Theory and Methods* 24, pp. 775–795.

DRAPER, N. and LIN, D. K. J. (1996). "Response Surface Designs". Chapter 11 in *Handbook of Statistics, Vol. 13, Design and Analysis of Experiments* edited by S. Ghosh and C.R. Rao. Elsevier Science, Amsterdam, pp. 343–375.

DRAPER, N. and PUKELSHEIM, F. (1990). "Another Look at Rotatability". *Technometrics* 32, pp. 195–202.

DRAPER, N. and PUKELSHEIM, F. (1994). "On Third-Order Rotatability". *Metrika* 41, pp. 137–161.

DRAPER, N. and PUKELSHEIM, F. (1996). "An Overview of Design of Experiments". *Statistical Papers* 37, pp. 1–32.

DRAPER, N. and PUKELSHEIM, F. (1998). "Polynomial Representations for Response Surface Modeling". *New Developments and Applications in Experimental Designs.* IMS Lecture Notes—Monograph Series 34, pp. 199–212.

DRAPER, N. and YING, L. H. (1994). "A Note on Slope Rotatability over all Directions". *Journal of Statistical Planning and Inference* 41, pp. 113–119.

DUMOCHEL, W. and JONES, B. (1994). "A Simple Bayesian Modification of D-optimal Designs to Reduce Dependence on an Assumed Model". *Technometrics* 13, pp. 682–688.

ENGEL, J. (1992). "Modelling Variation in Industrial Experiments". *Applied Statistics* 41, pp. 579–593.

ENGEL, J. and HUELE, A. F. (1996). "Taguchi Parameter Design by Second-Order Response Surfaces". *Quality and Reliability Engineering International* 12, pp. 95–100.

FAN, S.-K. S. (2000). "A Generalized Global Optimization Algorithm for Dual Response Systems". *Journal of Quality Technology* 32, pp. 444–456.

FANG, K.-T.; LIN, D. K. J.; WINKER, P.; and ZHANG, Y. (2000). "Uniform Design: Theory and Applications". *Technometrics* 42, pp. 237–248.

FEDERER, W. T. and MEREDITH, M. P. (1992). "Covariance Analysis for Split-Plot and Split-Block Designs". *The American Statistician* 46, pp. 155–162.

FERRER, A. J. and ROMERO, R. (1993). "Small Samples Estimation of Dispersion Effects from Unreplicated Data". *Communications in Statistics—Simulation and Computation* 22, pp. 975–995.

FERRER, A. J. and ROMERO, R. (1995). "A Simple Method to Study Dispersion Effects from Non-necessary Replicated Data in Industrial Context". *Quality Engineering* 7, pp. 747–755.

FORD, I.; TORSNEY, B.; and WU, C. F. (1992). "The Use of a Canonical Form in the Construction of Locally Optimal Designs for Nonlinear Problems". *Journal of the Royal Statistical Society Series* B 54, pp. 569–583.

FREENY, A. E. and LANDWEHR, J. M. (1995). "Graphical Analysis for a Large Designed Experiment". *Technometrics* 37, pp. 1–14.

GANJU, J. and LUCAS, J. M. (1997). "Bias in Test Statistics When Restrictions in Randomization are Caused by Factors". *Communications in Statistics—Theory and Methods* 26, pp. 47–63.

GANJU, J. and LUCAS, J. M. (1999). "Detecting Randomization Restrictions Caused by Factors". *Journal of Statistical Planning and Inference* 81, pp. 129–140.

GHOSH, S. and DUH, Y.-J. (1992). "Determination of Optimum Experimental Conditions using Dispersion Main Effects and Interactions of Factors in Replicated Factorial Experiments". *Journal of Applied Statistics* 19, pp. 367–378.

GIOVANNITTI-JENSEN, A. and MYERS, R. H. (1989). "Graphical Assessment of the Prediction Capability of Response Surface Designs". *Technometrics* 31, pp. 159–171.

GALDFARB, H. B.; BORROR, C. M.; and MONTGOMERY, D. C. (2003). "Mixture Process Experiments with Noise Variables". *Journal of Quality Technology* 35, pp. 393–405.

GOOS, P. and VANDEBROEK, M. (2001). "Optimal Split-Plot Designs". *Journal of Quality Technology* 33, pp. 436–450.

GREGO, J. M. (1993). "Generalized Linear Models and Process Variation". *Journal of Quality Technology* 25, pp. 288–295.

HAALAND, P.; MCMILLAN, N.; NYCHA, D.; and WELCH, W. (1994). "Analysis of Space-Filling Designs," *Computing Science and Statistics, Computationally Intensive Statistical Methods.* Proceedings of the 26^{th} Symposium on the Interface. Interface Foundation of North America, Fairfax Station, VA, pp. 111–120.

HADER, R. and PARK, S. H. (1978). "Slope-Rotatable Central Composite Designs". *Technometrics* 20, pp. 413–417.

HAINES, L. M. (1987). "The Application of the Annealing Algorithm to the Construction of the Exact Optimal Designs for Linear-Regresison Models". *Technometrics* 29, pp. 4394–437.

HAMADA, M.; MARTZ, H. F.; REESE, C. S.; and WILSON, A. G. (2001). "Finding Near-Optimal Bayesian Experimental Designs via Genetic Algorithms". *The American Statistician* 55, pp. 175–181.

HEISE, M. A. and MYERS, R. H. (1996). "Optimal Designs for Bivariate Logistic Regression". *Biometrics* 52, pp. 613–624.

HEREDIA-LANGNER, A; CARLYLE, W. M.; MONTGOMERY, D. C.; BORROR, C. M.; and RUNGER, G. C. (2003) "Genetic Algorithms for the Construction of D-Optimal Designs". *Journal of Quality Technology* 35, pp. 28–46.

HEREDIA-LANGNER, A.; CARLYLE, W. M.; MONTGOMERY, D. C.; and BORROR, C. M. (2004). "Model-Robust Optimal Designs: A Genetic Algorithm Approach". To appear in the *Journal of Quality Technology*

HEREDIA-LANGNER, A.; LOREDO, E. N.; MONTGOMERY, D. C.; and GRITTON, A. H. (2000). "Optimization of a Bonded Leads Process Using Statistically Designed Experiments". *Robotics and Computer Integrated Manufacturing* 16, pp. 377–382.

HILL, W. J. and HUNTER, W. G. (1966). "A Review of Response Surface Methodology: A Literature Review". *Technometrics* 8, pp. 571–590.

HOLCOMB, D. R. and CARLYLE, W. M. (2002). "Some Notes on the Construction and Evaluation of Supersaturated Designs". *Quality and Reliability Engineering International* 18, pp. 299–304.

HOLCOMB, D. R.; MONTGOMERY, D. C.; and CARLYLE, W. M. (2003). "Analysis of Supersaturated Designs". *Journal of Quality Technology* 35, pp. 13–27.

HOOKE, R. and JEEVES, T. A. (1961). "Direct Search

Solution of Numerical and Statistical Problems". *Journal of the Association of Computing Machinery* 8, pp. 212–229.

HUANG, P.; CHEN, D.; and VOELKEL, J. O. (1998). "Minimum-Aberration Two-Level Split-Plot Designs". *Technometrics* 40, pp. 314–326.

JIA, Y. and MYERS, R. H. (2001). "Experimental Designs for Logistic Models in Two Design Variables". Technical Report number 01-9, Department of Statistics, Virginia Tech, Blacksburg, VA.

JOHNSON, M. E.; MOORE, L. M.; and YLVISAKER, D. (1990). "Minimax and Maximin Distance Designs". *Journal of Statistical Planning and Inference* 36, pp. 131–148.

KACKAR, R. (1985). "Off-Line Quality Control, Parameter Designs, and the Taguchi Methods" (with discussion). *Journal of Quality Technology* 17, pp. 176–188.

KALISH, L. A. (1990). "Efficient Design for Estimation of Median Lethal Dose and Quantal Dose-Response Curves". *Biometrics* 46, pp. 737–748.

KHATTREE, R. (1996). "Robust Parameter Design: A Response Surface Approach". *Journal of Quality Technology* 28, pp. 187–198.

KHURI, A. I. (1988). "A Measure of Rotatability for Response Surface Designs". *Technometrics* 30, pp. 95–104.

KHURI, A. I. (1990). "Analysis of Multiresponse Experiments: A Review" in *Statistical Design and Analysis of Industrial Experiments* edited by S. Ghosh, Marcel Dekker, New York, NY, pp. 231–246.

KHURI, A. I. (1992). "Diagnostic Results Concerning a Measure of Rotatability". *Journal of the Royal Statistical Society Series* B 54, pp. 253–267.

KHURI, A. I. and CONLON, M. (1981). "Simultaneous-Optimization of Multiple Responses Represented by Polynomial Regression-Functions". *Technometrics* 23, pp. 363–375.

KHURI, A. I. and CORNELL, J. A. (1996). *Response Surfaces: Designs and Analyses* 2nd ed. Marcel Dekker, New York.

KHURI, A. I.; HARRISON, J. M.; and CORNELL, J. A. (1999). "Using Quantile Plots of the Prediction Variance for Comparing Designs for a Constrained Mixture Region: An Application Involving a Fertilizer Experiment". *Journal of the Royal Statistical Society Series* C 48, pp. 521–532.

KHURI, A. I.; KIM, H. J.; and UM, Y. (1996). "Quantile Plots of the Prediction Variance for Response Surface Designs". *Computational Statistics and Data Analysis* 22, pp. 395–407.

KIM, W. B. and DRAPER, N. R. (1994). "Choosing a Design for Straight Line Fits to Two Correlated Responses". *Statistica Sinica* 4, pp. 275–280.

KIM, K-J. and LIN, D. K. J. (1998). "Dual Response Surface Optimization: A Fuzzy Modeling Approach". *Journal of Quality Technology* 1, pp. 1–10.

KIM, K-J. and LIN, D. K. J. (2000). "Simultaneous Optimization of Mechanical Properties of Steel by Maximizing Experimental Desirability Functions". *Journal of the Royal Statistical Society Series* C 49, pp. 311–325.

KOWALSKI, S. M. (1999). "The Design and Analysis of Split-Plot Experiments in Industry". Ph.D. Dissertation, University of Florida, Gainesville, FL.

KOWALSKI, S. M. (2002). "24 Run Split-Plot Experiments for Robust Parameter Design". *Journal of Quality Technology* 34, pp. 399–410.

KOWALSKI, S. M., CORNELL, J. A.; and VINING, G. G. (2002). "Split-Plot Designs and Estimation Methods for Mixture Experiments with Process Variables". *Technometrics* 44, pp. 72–79.

KRAFT, O. and SCHAEFER, M. (1992). "D-Optimal Designs for a Multivariate Regression Model". *Journal of Multivariate Analysis* 42, pp. 130–140.

KUHN, A. M.; CARTER, W. H.; and MYERS, R. H. (2000). "Incorporating Noise Factors into Experiments with Censored Data". *Technometrics V42, pp. 376–383.*

LANGSRUD, O. (2001). "Identifying Significant Effects in Fractional Factorial Multiresponse Experiments". *Technometrics* 43, pp. 415–424.

LASDON, L.; FOX, R.; and RATNER, M. (1974). "Nonlinear Optimization Using Generalized Reduced Gradient Method". *Revue Francaise D Automatique Informatique Recherche Operationnelle* 8, pp. 73–103.

LAWSON, J. (2003). "One-Step Screening and Process optimization Experiments". *The American Statistician* 57, pp. 15–20.

LEE, Y. and NELDER, J. A. (1998). "Joint Modeling of Mean and Dispersion". *Technometrics* 40, pp. 168–171.

LEE, Y. and NELDER, J. A. (2003). "Robust Designs via Generalized Linear Models". *Journal of Quality Technology* 35, pp. 2–12.

LETSINGER, J. D.; MYERS, R. H.; and LENTNER, M. (1996). "Response Surface Methods for Bi-Randomization Structures". *Journal of Quality Technology* 28, pp. 381–397.

LEWIS, S. L.; MONTGOMERY, D. C.; and MYERS, R. H. (1999-2000). "The Analysis of Designed Experiments with Nonnormal Responses". *Quality Engineering* 12, pp. 225–244.

LEWIS, S. L.; MONTGOMERY, D. C.; and MYERS, R. H. (2001a). "Examples of Designed Experiments with Nonnormal Responses". *Journal of Quality Technology* 33, pp. 265–278.

LEWIS, S. L.; MONTGOMERY, D. C.; and MYERS, R. H. (2001b). "Confidence Interval Coverage for Designed Experiments Analyzed with Generalized Linear Models". *Journal of Quality Technology* 33, pp. 279–292.

LEWIS, S. M. and DEAN, A. M. (2001). "Detection of Interactions in Experiments on Large Numbers of Factors". *Journal of the Royal Statistical Society Series* B 63, pp. 633–672.

LI, H. and MEE, R. W. (2002). "Better Foldover Fractions for Resolution III 2^{k-p} Designs". *Technometrics* 44, pp. 278–283.

LI, W. and NACHTSHEIM, C. J. (2000). "Model-Robust Factorial Designs". *Technometrics* 42, pp. 345–352.

LI, W. W. and WU, C. F. J. (1997). "Columnwise-Pairwise Algorithms with Applications to the Construction of Supersaturated Designs". *Technometrics* 39, pp. 171–179.

LIN, D. K. J. (1993a). "A New Class of Supersaturated Designs". *Technometrics* 35, pp. 28–31.

LIN, D. K. J. (1993b), "Another Look at First-Order Saturated Designs: The p-Efficient Designs". *Technometrics* 35, pp. 284–292.

LIN, D. K. J. (1995). "Generating Systematic Supersaturated Designs". *Technometrics* 37, pp. 255–264.

LIN, D. K. J. (2000). "Recent Developments in Super-

saturated Designs" in *Statistical Process Monitoring and Optimization* edited by S. H. Park and G. G. Vining, Marcel Dekker, New York, NY, pp. 305–320.

LIN, D. K. J. and DRAPER, N. (1992). "Projection Properties of Plackett and Burman Designs". *Technometrics* 34, pp. 423–428.

LIN, D. K. J. and DRAPER, N. (1993). "Generating Alias Relationships for Two-Level Plackett-Burman Designs". *Computational Statistics and Data Analysis* 15, pp. 147–157.

LIN, D. K. J. and TU, W. (1995). "Dual Response Surface Optimization". *Journal of Quality Technology* 27, pp. 248–260.

LIN, H. F.; MYERS, R. H.; and YE, K. Y. (2000). "Bayesian Two-stage Optimal Design for Mixture Models". *Journal of Statistical Computation and Simulation* 66, pp. 209–231.

LIND, E. E.; GOLDIN, J.; and HICKMAN, J. B. (1960). "Fitting Yield and Cost Response Surfaces". *Chemical Engineering Progress* 56, p. 62.

LUCAS, J. M. (1989). "Achieving a Robust Process Using Response Surface Methodology". Paper presented at American Statistical Association Conference, Washington, D.C.

LUCAS, J. M. (1994). "How to Achieve a Robust Process Using Response Surface Methodology". *Journal of Quality Technology* 26, pp. 248–260.

LUCAS, J. M. and HAZEL, M. C. (1997). "Running Experiments With Multiple Error Terms: How an Experiment is Run is Important". *ASQC Technical Conference Transactions,* American Society for Quality Control, Milwaukee, WI, pp. 283–296.

LUCAS, J. M. and JU, H. L. (1992). "Split Plotting and Randomization in Industrial Experiments". *ASQC Quality Congress Transactions,* American Society for Quality Control, Nashville, TN, pp. 374–382.

LUNANI, M.; NAIR, V. N.; and WASSERMAN, G. S. (1997). "Graphical Methods for Robust Parameter Design with Dynamic Characteristics". *Journal of Quality Technology* 29, pp. 327–338

MATHEW, T. and SINHA, B. K. (1992). "Exact and Optimum Tests in Unbalanced Split-Plot Designs Under Mixed and Random Models". *Journal of the American Statistical Association* 87, pp. 192–200.

MAYS, D. P. and MYERS, R. H. (1996). "Bayesian Approach for the Design and Analysis of a Two Level Factorial Experiment in the Presence of Dispersion Effects". *Communications in Statistics—Theory and Methods* 25, pp. 1409–1428.

MCCULLOUGH, C. and NELDER, J. A. (1989). *Generalized Linear Models* 2nd ed. Chapman and Hall, New York, NY.

MCGRATH, R. N. and LIN, D. K. J. (2001a). "Confounding of Location and Dispersion Effects in Unreplicated Fractional Factorials". *Journal of Quality Technology* 33, pp. 129–139.

MCGRATH, R. N. and LIN, D. K. J. (2001b). "Testing Multiple Dispersion Effects in Unreplicated Fractional Factorial Designs". *Technometrics* 43, pp. 406–414.

MCKAY M. D.; BECKMAN, R. J.; and CONOVER, W. J. (1979). "A Comparison of Three Methods for selecting Values of Input Variables in the Analysis of Output from a Computer Code". *Technometrics* 21, pp. 239–245.

MEAD, R. and PIKE, D. J. (1975). "A Review of Response Surface Methodology from a Biometrics Viewpoint". *Biometrics* 31, pp. 803–851.

MEE, R. W. (2001). "Noncentral Composite Designs". *Technometrics* 43, pp. 34–43.

MEE, R. W. (2002). "Three-Level Simplex Designs and Their Use in Sequential Experimentation". *Journal of Quality Technology* 34, pp. 152–164.

MEE, R. W. and BATES, R. L. (1998). "Split-Lot Designs: Experiments for Multistage Batch Processes". *Technometrics* 40, pp. 127–140.

MEE, R. W. and PERALTA, M. (2000). "Semifolding 2^{k-p} Designs". *Technometrics* 42, pp. 122–134.

MEYER, R. D.; STEINBERG, D. M.; and BOX, G. E. P. (1996). "Follow-Up Designs to Resolve the Confounding in Multifactor Experiments". *Technometrics* 38, pp. 303–313.

MILLER, A. (1997). "Strip-Plot Configurations of Fractional-Factorials". *Technometrics* 39, pp. 153–161.

MILLER, A. (2002). "Analysis of Parameter Design Experiments for Signal-Response Systems". *Journal of Quality Technology* 34, pp. 139–151.

MILLER, A. and WU, C. F. J. (1996). "Parameter Design for Signal-Response Systems: A Different Look at Taguchi's Dynamic Parameter Design". *Statistical Science* 11, pp. 122–136.

MONTEPIEDRA, G.; MYERS, D.; and YEH, A. B. (1998). "Application of Genetic Algorithms to the Construction of Exact D-Optimal Designs". *Journal of Applied Statistics* 25, pp. 817–826.

MONTGOMERY, D. C. (1990-91). "Using Fractional Factorial Designs for Robust Process Development". *Quality Engineering* 3, pp. 193–205.

MONTGOMERY, D. C. (1992). "The Use of Statistical Process Control and Design of Experiments in Product and Process Improvement". *IIE Transactions* 24, pp. 4–17.

MONTGOMERY, D. C. (1999). "Experimental Design for Product and Process Design and Development" (with Commentary). *Journal of the Royal Statistical Society Series* D 48, pp. 159–177.

MONTGOMERY, D. C. (2001). *Design and Analysis of Experiments* 5th ed. John Wiley & Sons, Inc., New York, NY.

MONTGOMERY, D. C. and BORROR, C. M. (2001). "Discussion of 'Factor Screening and Response Surface Exploration' by Cheng and Wu". *Statistica Sinica* 11, pp. 591–595.

MONTGOMERY, D. C. and RUNGER, G. C. (1996). "Foldovers of 2^{k-p} Resolution IV Experimental Designs". *Journal of Quality Technology* 28, pp. 446–450.

MORRIS, M. (2000). "A Class of Three-Level Experimental Designs for Response Surface Modeling". *Technometrics* 42 pp. 111–121.

MUKERJEE, R. and HUDA, S. (1985). "Minimax Second and Third-Order Designs to Estimate the Slope of a Response Surface". *Biometrika* 72, pp. 173–178.

MURTY, V. N. and STUDDEN, W. J. (1972). "Optimal Designs for Estimating the Slope of a Polynomial Regression". *Journal of the American Statistical Association* 67, pp. 869–873.

MYERS, R. H. (1991). "Response Surface Methodology in Quality Improvement". *Communications in Statistics—Theory and Methods* 20, pp. 457–476.

MYERS, R. H. (1999). "Response Surface Methodology—Current Status and Future Directions" (with discussion). *Journal of Quality Technology* 31, pp. 30–44.

MYERS, R. H. and CARTER, W. H. (1973). "Response Surface Techniques for Dual Response Systems". Technometrics 15, pp. 301–317.

MYERS, R. H.; KHURI, A. I.; and CARTER, W. H. (1989). "Response Surface Methodology: 1966-1988". *Technometrics* 31, pp. 137–157.

MYERS, R. H.; KHURI, A. I.; and VINING, G. G. (1992). "Response Surface Alternatives to the Taguchi Robust Parameter Design Approach". *American Statistician* 46, pp. 131–139.

MYERS, R. H.; KIM, Y.; and GRIFFITHS, K. (1997). "Response Surface Methods and the Use of Noise Variables". *Journal of Quality Technology* 29, pp. 429–441.

MYERS, R. H. and LAHODA, S. J. (1975). "A Generalization of the Response Surface Mean Square Error Criterion with a Specific Application to the Slope". *Technometrics* 17, pp. 481–486.

MYERS, R. H. and MONTGOMERY, D. C. (2002). *Response Surface Methodology, Process and Product Optimization Using Designed Experiments* 2nd ed. John Wiley & Sons, New York, NY.

MYERS, R. H.; MONTGOMERY, D. C.; and VINING, G. G. (2002). *Generalized Linear Models with Applications in Engineering and the Sciences*. John Wiley and Sons, Inc., New York, NY.

MYERS, W. R.; MYERS, R. H.; and CARTER, W. H. (1994). "Some Alphabetic Optimal Designs for the Logistic Regression Model". *Journal of Statistical Planning and Inference* 42, pp. 57–77.

MYERS, R. H.; VINING, G. G.; GIOVANNITTI-JENSEN, A.; and MYERS, S. L. (1992). "Variance Dispersion Properties of Second-Order Response Surface Designs". *Journal of Quality Technology* 24, pp. 1–11.

NAIR, V. N. ET AL. (Editors) (1992). "Taguchi's Parameter Design: A Panel Discussion". *Technometrics* 34, pp. 127–161.

NELDER, J. A. and LEE, Y. (1991). "Generalized Linear Models for the Analysis of Taguchi-Type Experiments". *Applied Stochastic Models and Data Analysis* 7, pp. 107–120.

NELDER, J. A. and WEDDERBURN, R. W. M. (1972). "Generalized Linear Models". *Journal of the Royal Statistical Society* A 135, pp. 370–384.

NELSON, B. J.; MONTGOMERY, D. C.; ELIAS, R. J.; and MAASS, E. (2000). "A Comparison of Several Design Augmentation Strategies". *Quality and Reliability Engineering International* 16, pp. 435–449.

NELSON, L. S. (1985). "What Do Low F Ratios Tell You?" *Journal of Quality Technology* 17, pp. 237–238.

NESTEROV, Y. and NEMIROVSKII, A. (1994). *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics, Philadelphia, PA.

NGUYEN, N.-K. (1996). "An Algorithmic Approach to Generating Supersaturated Designs". *Technometrics* 38, pp. 69–73.

O'CONNELL, M.; HAALAND, P.; HARDY, S.; and NYCHKA, D. (1995). "Nonparametric Regression, Kriging, and Process Optimization" in *Statistical Modeling* Springer Verlag, NY, pp. 207–214.

O'NEILL, J. C.; BORROR, C. M.; EASTMAN, P. Y.; FRADKIN, D. G.; JAMES, M. P.; MARKS, A. P.; and MONTGOMERY, D. C. (2000). "Optimal Assignment of Samples to Treatments for Robust Design". *Quality and Reliability Engineering International* 16, pp. 417–422.

PAN, G. and SANTNER, T. J. (1998). "Selection and Screening Procedures to Determine Optimal Product Designs". *Journal of Statistical Planning and Inference* 67, pp. 311–330.

PARK, S. H. (1987). "A Class of Multifactor Designs for Estimating the Slope of Response Surfaces". *Technometrics* 29, pp. 449–453.

PEIPEL, G. and ANDERSON, C. M. (1992). "Variance Dispersion Graphs for Designs on Polyhedral Regions". *Proceedings of the Section on Physical and Engineering Sciences* American Statistical Association, Alexandria, VA. pp. 111–117.

PEIPEL, G.; ANDERSON, C. M.; and REDGATE, P. E. (1993a). "Variance Dispersion Graphs for Designs on Polyhedral Regions—Revisited". *Proceedings of the Section on Physical and Engineering Sciences* American Statistical Association, Alexandria, VA, pp. 102–107.

PEIPEL, G.; ANDERSON, C. M.; and REDGATE, P. E. (1993b). "Response Surface Designs for Irregularly-Shaped Regions (Parts 1, 2, and 3)". *Proceedings of the Section on Physical and Engineering Sciences* American Statistical Association, Alexandria, VA, pp. 205–227.

PHADKE, M. S. (1989). *Quality Engineering Using Robust Design*. Prentice Hall, Englewood Cliffs, NJ.

PIGNATIELLO, J. J., JR. (1993). "Strategies for Robust Multiresponse Quality Engineering". *IIE Transactions* 25, pp. 5–15.

PIGNATIELLO, J. J., JR. and RAMBERG, J. S. (1991). "Top Ten Triumphs and Tragedies of Genichi Taguchi". *Quality Engineering* 4, pp. 211–225.

PLEDGER, M. (1996). "Observable Uncontrollable Factors in Parameter Design". *Journal of Quality Technology* 28, pp. 153–162.

REMMENGA, M. D. and JOHNSON, D. E. (1995). "A Comparison of Inference Procedures in Unbalanced Split-Plot Designs". *Journal of Statistical Computation and Simulation* 51, pp. 353–367.

RICE, J. (1995). *Mathematical Statistics and Data Analysis* 2nd ed. Wadsworth & Brooks/Cole, Pacific Grove, CA.

ROSENBAUM, P. R. (1994). "Dispersion Effects from Fractional Factorials in Taguchi's Method of Quality Design". *Journal of the Royal Statistical Society* B 56, pp. 641–652.

ROSENBAUM, P. R. (1996). "Some Useful Compound Dispersion Experiments in Quality Design". *Technometrics* 38, pp. 354–364.

ROSENBERGER, J. and KALISH, L. (1978). "Optimal Designs

for Estimation of Parameters of Logistic Function". *Biometrics* 34, pp. 746–756.

SACKS, J.; SCHILLER, S. B.; and WELCH, W. J. (1989). "Designs for Computer Experiments". *Technometrics* 31, pp. 41–47.

SACKS, J.; WELCH, W. J.; MITCHELL, T. J.; and WYNN, H. P. (1989). "Design and Analysis of Computer Experiments" (with discussion). *Statistical Science* 4, pp. 409–435.

SATTERTHWAITE, F. E. (1959). "Random Balance Experimentation" (with Discussion). *Technometrics* 1, pp. 111–137.

SCHOEN, E. D. (1999). "Designing Fractional Two-Level Experiments With Nested Error Structures". *Journal of Applied Statistics* 26, pp. 495–508.

SCHOEN, E. D. and WOLFF, K. (1997). "Design and Analysis of a Fractional $4^1 3^1 2^5$ Split-Plot Experiment". *Journal of Applied Statistics* 24, pp. 409–419.

SHOEMAKER, A. C. and TSUI, K.-L. (1993). "Response Model Analysis for Robust Design Experiments". Communications in Statistics—Series A 22, pp. 1037–1064.

SHOEMAKER, A. C.; TSUI, K.-L.; and WU, C. F. J. (1991). "Economical Experimentation Methods for Robust Design". *Technometrics* 33, pp. 415–427.

SITTER, R. R. (1992). "Robust Designs for Binary Data". *Biometrics* 48, pp. 1145–1155.

SITTER, R. R. and TORSNEY, B. (1995). "Optimal Designs for Binary Response Experiments with 2 Design Variables". *Statistica Sinica* 5, pp. 405–419.

SITTER, R. R. and WU, C. F. W. (1993). "Optimal Designs for Binary Response Experiments—Fieller, D, and A Criteria". *Scandinavian Journal of Statistics* 20, pp. 329–341.

SPARGO, C. A.; HAALAND, P. D.; JURGENSEN, S. R.; SHANK, D. D.; and WALKER, G. T. (1993). "Chemiluminescent Detection of Strand Displacement Amplified DNA From Species Comprising the Microbacterium Tuberculosis Complex". *Molecular and Cellular Probes* 7, pp. 395–404.

STEINBERG, D. M. and BURSZTYN (1994). "Dispersion Effects in Robust-Design Experiments with Noise Factors". *Journal of Quality Technology* 26, pp. 12–20.

STEINBERG, D. M. and BURSZTYN (1998). "Noise Factors, Dispersion Effects, and Robust Design". *Statistica Sinica* 8, pp. 67–85.

STEINER, S. H. and HAMADA, M. (1997). "Making Mixtures Robust to Noise and Mixing Measurement Errors". *Journal of Quality Technology* 29, pp. 441–450.

TAGUCHI, G. (1987). *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost*. Quality Resources, White Plains, NJ.

TAGUCHI, G. (1991). *Introduction to Quality Engineering*. UNIPUB/Kraus International, White Plains, New York (Originally published by the Asian Productivity Organization).

TAGUCHI, G. and WU, Y. (1985). *Introduction to Off-Line Quality Control*. Central Japan Quality Control Association (available from American Supplier Institute, Dearborn, MI).

TRINCA, L. A. and GILMOUR, S. G. (1998). "Variance Dispersion Graphs for Comparing Blocked Response Surface Designs". *Journal of Quality Technology* 30, pp. 314–327.

TRINCA, L. A. and GILMOUR, S. G. (2001). "Multistratum Response Surface Designs". *Technometrics* 43, pp. 25–33.

TSUI, K.-L. (1996). "A Critical Look at Taguchi's Modelling Approach for Robust Design". *Journal of Applied Statistics* 23, pp. 81–95.

VINING, G. G. (1993). "A Computer Program for Generating Variance Dispersion Graphs". *Journal of Quality Technology* 25, pp. 45–58.

VINING, G. G. (1998). "A Compromise Approach to Multiresponse Optimization". *Journal of Quality Technology* 30, pp. 309–313.

VINING, G. G. and BOHN, L. L. (1998). "Response Surfaces for the Mean and Variance Using a Nonparametric Approach". *Journal of Quality Technology* 30, pp. 282–291.

VINING, G. G.; CORNELL, J.; and MYERS, R. H. (1993). "A Graphical Approach for Evaluating Mixture Designs". *Journal of the Royal Statistical Society Series* C 42, pp. 127–138.

VINING, G. G. and MYERS, R. H. (1990). "Combining Taguchi and Response Surface Philosophies: A Dual Response Approach". *Journal of Quality Technology* 22, pp. 38–45.

VINING, G. G. and SCHAUB, D. (1996). "Experimental Designs for Estimating Both Mean and Variance Functions". *Journal of Quality Technology* 29, pp. 135–147.

VUCHKOV, I. N. and BOYADJIEVA, L. N. (1992). "Quality Improvement Through Design of Experiments with Both Product Parameters and External Noise Factors" in *Model Oriented Data Analysis* edited by V. Fedorov, W. G. Muller, I. N. Vuchkov. Physica-Verlag, Heidelberg, pp. 195–212.

WANG, J. C. and WU, C. F. J. (1995). "A Hidden Projection Property of Plackett-Burman and Related Designs". *Statistica Sinica* 5, pp. 235–250.

WELCH, W. J. (1982). "Branch and Bound Search for Experimental Designs Based on D-Optimality and Other Criteria". *Technometrics* 24, pp. 41–48.

WELCH, W. J.; BUCK, R. J.; SACKS, J.; WYNN, H. P.; MITCHELL, T. J.; and MORRIS, M. D. (1992). "Screening, Predicting, and Computer Experiments". *Technometrics* 34, pp. 15–25.

WELCH, W. J.; YU, T. K.; KANG, S. M.; and SACKS, J. (1990). "Computer Experiments for Quality Control by Parameter Design". *Journal of Quality Technology* 22, pp. 15–22.

WESTFALL, P. H.; YOUNG, S. S.; and LIN, D. K. J. (1998). "Forward Selection Error Control in the Analysis of Supersaturated Designs". *Statistica Sinica* 8, pp. 101–117.

WOLFINGER, R. D. and TOBIAS, R. D. (1998). "Joint Estimation of Location, Dispersion, and Random Effects in Robust Design". *Technometrics* 40, pp. 62–71.

WU, C. F. J. (1993). "Construction of Supersaturated Designs

Through Partially Aliased Interactions". *Biometrika* 80, pp. 661–669.

WU, C. F. J. and DING, Y. (1998). "Construction of Response Surface Designs for Qualitative and Quantitative Factors". *Journal of Statistical Planning Inference* 71, pp. 331–348.

YAMADA, S. and LIN, D. K. J. (1997). "Supersaturated Designs Including an Orthogonal Base". *Canadian Journal of Statistics* 25, pp. 203–213.

YING, L. A.; PUKELSHEIM, F.; and DRAPER, N. R. (1995a). "Slope Rotatability over all Directions Designs". *Journal of Applied Statistics* 22, pp. 331–341.

YING, L. A.; PUKELSHEIM, F.; and DRAPER, N. R. (1995b). "Slope Rotatability over all Directions Designs for k $\geq$ 4". *Journal of Applied Statistics* 22, pp. 343–354.

ZAHRAN, A.; ANDERSON-COOK, C. M.; and MYERS, R. H. (2003). "Fraction of Design Space to Assess Prediction Capability of Response Surface Designs". *Journal of Quality Technology* 35, pp. 377–386.

ZHOU, J. (2001). "A Robust Criterion for Experimental Designs for Serially Correlated Observations". *Technometrics* 43, pp. 462–467.