J. Ind. Soc. Agril. Statist.
62(2), 2008 : 156-168

Analysis of Supersaturated Designs: A Review

Sudhir Gupta and Priya Kohli
Northern Illinois University, DeKalb, IL 60115, U.S.A.

SUMMARY

Identification of correct model is essential for optimization of processes and systems in engineering and the sciences. Supersaturated designs provide an important tool for evaluating a large number of factors in the initial phase of an experiment. Several procedures have been proposed in literature over the last decade or so for identification and estimation of correct model using supersaturated designs. A review of these methods is provided in this paper. A generalization of the contrast variance method to situations where coefficients of active effects are neither all equal nor linearly related is also provided. Since no method seems to be consistently superior over other methods, it is concluded that results obtained using competing methods should be carefully evaluated before a final model is selected.

Key words: Active factors, Adjusted p-values, All-subsets regression, Bayesian approach, Bootstrap method, Contrast variance method, Intrinsic Bayes factor, Monte Carlo, Partial least squares, Penalized least squares, Plackett-Burman design, Search designs, Stepwise regression, Stochastic search variable selection.

1. INTRODUCTION

In preliminary phase of many experiments, especially in industry, simulation experiments or exploratory research, experimenters are forced to study a large number of factors. The main purpose of such experiments is to identify a few important factors with large effects from among a very large number of potentially relevant factors. Due to constraints of cost and time, not too many experimental runs can be performed. For example, Nguyen (1996) described an experiment dealing with passenger impact crash test of cars in which as many as 54 factors were initially thought to be relevant. The number of car prototypes that could be used was at most 28. Often the number of factors that turn out to be really important is far fewer than the number of factors that are initially thought to be relevant. For example, Lin (1995) described a situation related to modeling of AIDS data which involved 150 factors to begin with, but only 5 of these eventually turned out to be important.

If resources are adequate, a suitable main effect plan can be used to design a screening experiment. Often, however, the size of even a main effect plan may be too large. For instance, in Nguyen's (1996) example, an orthogonal main effect plan would require 55 runs, which is clearly too large a number given that at most 28 runs can be afforded. In such cases, a design requiring far fewer runs than the number of factors is needed. Such designs are called supersaturated designs. More specifically, if f is the number of factors involved and n is the number of runs, then for a supersaturated design, $n < f + 1$. Supersaturated designs for two-level factors have received considerable attention over the last few years.

The idea of supersaturated designs was initiated by Satterthwaite (1959). However, the random balance designs of Satterthwaite are not very useful. Booth and Cox (1962) proposed a criterion, $E(s^2)$ criterion, for comparing supersaturated designs. Using this criterion, they constructed some designs with the help of a computer. The criterion is a measure of nonorthogonality strictly for the case of two active factors. Wu (1993) discussed the connection of $E(s^2)$ criterion with A- and D-optimality for the case of more than two active factors and gave a method of constructing supersaturated designs using Hadamard matrices. Since then, several authors have given methods of construction, for which we refer the reader to Xu and Wu (2005) and the references cited therein. In this paper we restrict attention to supersaturated designs for two-level factors.

Supersaturated designs do not afford a transparent analysis as is the case with full-rank models. Since the columns of the design matrix can not all be independent, the effects of active factors get aliased with one another in supersaturated designs making it very diffcult to identify the active factors correctly. Several authors have investigated this problem and have developed novel methods for analyzing them. The purpose of this paper is to provide a review of various methods for analyzing supersaturated designs given in literature over the last decade or so. A generalization of the contrast variance method of Holcomb *et al*. (2003) for situations where coefficients of active effects are neither all equal nor linearly related is also provided. Attention is restricted to supersaturated designs, and methods for analyzing saturated designs are not considered, see Hamada and Balakrishnan (1998). We refer to Wang and Voss (2001) for selection of active factors within the context of saturated designs.

2. ANALYSIS METHODS

Hamada and Wu (1992)

The regular fractional factorial or geometric Plackett-Burman (PB) designs with number of runs a power of two are popular for screening experiments due to their nice projection properties and orthogonal estimation of main effects and unconfounded interactions. In contrast, non-geometric PB designs have complex aliasing patterns, and thus they have traditionally been used to screen only main effects under the absence of all interactions. Hamada and Wu (1992) argued that complex aliasing patterns of non-geometric PB designs may, in fact, be considered as an advantage in the following sense. The regular fractions do not allow estimation of a confounded interaction because the interaction effect is completely aliased with one of the main effects. On the other hand, since interactions are partially aliased with main effects in non-geometric PB designs, it may be possible to estimate them under certain situations. Under effect sparsity, Hamada and Wu (1992) considered models that satisfy the condition of effect heredity, i.e. a two-factor interaction is included only if main effect of a factor involved in the interaction is significant, and proposed a method based on iterative forward selection. They first identify significant main effects using usual methods including the half-normal plot. Then all the significant main effects, all two-factor interactions that involve a significant main effect factor, and other interactions that are thought to be relevant are considered. Forward stepwise procedure is then used to select significant effect from this list. Having identified the significant effect in this list, a second forward stepwise procedure is used to identify additional main effects that may have been missed earlier. This whole procedure may be repeated more than once to identify the final model.

Box and Meyer (1993)

Box and Meyer proposed a Bayesian approach for analysis of designs with complex aliasing that is based on marginal posterior probability of a factor being active by summing over posterior probabilities of subsets of models that contain that factor. The method is considerably more computationally intensive than the all-subsets regression approach under effect heredity because posterior model probabilities have to be computed. The procedure considers models containing all possible combinations of factors as in all-subsets regression. Let the i^{th} model be denoted by M_i. Each model contains main effects and all interactions up to a specified order among the factors included in the model. Let the i^{th} model M_i be written as $Y = X_i\beta_i + \in$, $\beta_i = (\beta_{0i}, \beta_{1i}, ..., \beta_{m_i i})'$ be the parameter vector with corresponding design matrix X_i, and $\in$ is the vector of independent random errors distributed normally with zero mean and constant variance σ^2. The prior probability of model M_i is given by $p(M_i) = \pi^{p_i}(1-\pi)^{k-p_i}$, where p_i is the number of active factors included in M_i, π is the constant probability of a factor being active, and k is the total number of factors. The likelihood of the data is then

$$f(y \mid M_i) \propto \sigma^{-n} \exp\{-(y - X_i\beta_i)'(y - X_i\beta_i)/(2\sigma^2)\}$$

where n is the number of runs. The β_{oi} and $\log(\sigma)$ are assumed to have noninformative priors and $(\beta_{1i},...,\beta_{m_i i})'$ is assumed to follow the prior $N(0, \gamma^2\sigma^2 I)$. The value of γ is one for which $p(\gamma \mid y)$ is maximum, being proportional to $\{p(M_0 \mid y)\}^{-1}$. Here, the model M_0 has only the intercept term and none of the factor effects. The value of γ is then obtained by minimizing the probability of model M_0. The posterior probability $p(M_i \mid y)$ needs to be computed for each value of γ. Box and Meyer (1993) showed that then the posterior probability of the model is given by

$$p(M_i \mid y) \propto T_{1i}T_{2i}, \quad \hat{\beta}_i = (\Gamma_i + X_i'X_i)^{-1}X_i'Y$$

$$T_{1i} = \left(\frac{\pi}{1-\pi}\right)^{p_i} \gamma^{-m_i} \frac{|X_0' + X_0|^{1/2}}{|\Gamma_i + X_i' X_i|^{1/2}}$$

$$T_{2i} = \left(\frac{s(\hat{\beta}_i) + \hat{\beta}_i' \Gamma_i \hat{\beta}_i)}{s(\hat{\beta}_0)}\right)^{-(n-1)/2}$$

$$\Gamma_i = \gamma^{-2} \begin{pmatrix} 0 & 0 \\ 0 & I_{m_i} \end{pmatrix}$$

$$s(\hat{\beta}_i) = (y - X_i\hat{\beta}_i)'(y - X_i\hat{\beta}_i)$$

T_{1i} is the penalty for having a larger number of parameters in the model and T_{2i} is a measure of how well the model fits the data. The marginal posterior probability of a factor is obtained by summing posterior probabilities $p(M_i|y)$ of models that contain that factor. Active factors are identified using these marginal posterior probabilities.

Samset and Tyssedal (1998) observed that in some situations, Box and Meyer (1993) method appears to be dependent on the size of active effects and size of σ^2. For example, sometimes their method fails to identify an active factor with a small effects size if the range of the effect sizes is increased. They argued that because nonsignificant coefficients are retained in model M_i, the penalty term T_{1i} is larger than what it should really be, resulting in the method missing some small size active effects. They proposed a modification of Box and Meyer (1993) method by removing all non-significant terms from M_i, thereby increasing the value of γ, and reducing the magnitude of the penalty function T_{1i}. They showed with the help of examples that the modified method performs better than Box and Meyer's (1993) method.

Chipman, Hamada and Wu (1997)

Chipman *et al*. (1997) developed a Bayesian approach based on the stochastic search variable selection (SSVS) method of George and McCulloch (1993, 1997). All possible models are explored in SSVS even in the supersaturated set-up. However, rather than selecting a final model from among all possible models, the method yields several potential final models with high model posterior probabilities. The SSVS approach can be briefly explained using the linear regression model $Y_r = X\beta + \in, \in \sim N_n(0, \sigma^2 I_n)$, where Y_r is the vector of residuals obtained from response Y after fitting the intercept and any other coefficients that are to be included, a priori, in the final model. The matrix X is normalized such that the sum of squares of the elements of each column equals unity. Normal mixture priors are specified for the k regression parameters in β. For small size of an effect (inactive effect), the variance is taken to be τ_i^2 and for a large effect size (active effect), the variance is taken to be $(c_i\tau_i)^2$, where c_i and τ_i are appropriately chosen tuning constants, i = 1, 2, ... , k. Whether an effect is considered small or large is determined through Bernoulli prior distribution. This in effect specifies a prior on the model. Chipman *et al*. (1997) restricted the search to models that satisfy the condition of effect heredity. This means that the main effects and interactions can not have independent priors as presence of interaction in a model is contingent upon the presence of the main effect of at least one of the factors involved in the interaction. This hierarchy of effects is incorporated through hierarchical priors developed by Chipman (1996). Normal prior for an inactive effect is taken to be $\beta_i \sim N(0, \tau_i^2)$ and correspondingly for an active effect, $\beta_i \sim N(0, c_i^2\tau_i^2)$. Finally, a prior on σ is specified through $v\lambda/\sigma^2 \sim X_v^2$, where v and λ are tuning parameters for the prior on σ.

Beattie *et al*. (2002) note that the conditional probability that β_i is active, to be used in Gibbs sampler, is given by

$$\left[1 + \frac{(1-\pi_i)c_i}{\pi_i} \exp\left(\frac{-\beta_i^2(c_i^2 - 1)}{2c_i^2\tau_i^2}\right)\right]^{-1}$$

where π_i is the prior probability of the i^{th} factor being active. Thus, large τ_i's produce models with few active factors and conversely, small τ_i's provide models with too many active factors. Chipman *et al*. (1997) use $c_i = 10$ indicating that active effect is substantially larger in magnitude than an inactive effect. Also, they used

$$\tau_i = \frac{\Delta Y}{3\{\max(X_i) - \min(X_i)\}}$$

ΔY is taken to be equal to $s_Y/5$, where s_Y denotes the sample standard deviation of ΔY. The σ is also taken to be approximately equal to ΔY with $v \approx 2$. The posterior probability of a model is used to assess the goodness of a model obtained using SSVS. The iterative stepwise methodology of Hamada and Wu (1992) does not search for models as well as the all-subsets regression approach.

The approach based on SSVS, under condition of effect heredity, stochastically searches all possible models with nonzero probability. As an advantage over all-possible regression, the SSVS approach can identify several possible models capable of explaining the data sufficiently well.

Westfall, Young and Lin (1998)

Wastfall *et al.* (1998) observed that Type 1 error rates in forward selection method can be high under effect sparsity, as high as 70% when 5% significance level for selecting active factors is used. They carried out a simulation study for the half fraction of the 28-run Plackett-Burman design given by Lin (1993) using up to $p = 6$ active effects and significance levels α of 0.05, 0.10 and 0.15. The study showed that it is highly likely that one or more inactive is declared as an active factor. Several inactive factors may be declared active at higher significance levels. The problem is that the mean square error appearing in the denominator of the F test statistic tends to be overly biased downwards resulting in large F values. The authors proposed a resampling method for controlling Type 1 errors at each stage of the forward selection through adjusted p-values.

The method is based on the distribution of the maximal F statistic, which is maximum of all the F test statistics computed at any given step of the forward selection procedure. At any given step, the method assumes that the factors selected in previous steps, and thus appearing in the model at the given step, were forced into the model and none of the remaining factors under consideration are truly active. The distribution of the maximal F statistic is then completely determined by the known design matrix X but it is not analytically tractable due to complex aliasing among the columns of the design matrix. Then, adjusted p-value is obtained using two components - the Bonferroni estimate and an estimate of the remainder to be substracted from the Bonferroni estimate. The remainder is obtained using Monte Carlo simulation. Simulation results for Lin (1993), 14×23 design showed that Bonferroni method provides a better control over Type 1 error than the unadjusted forward selection method. Thus, the proposed adjusted p-value procedure appears to be an improvement over Bonferroni's method.

Chen and Lin (1998)

Srivastava (1975) showed that when $\sigma^2 = 0$, a necessary and sufficient condition for identifying and estimating p active effects is that every submatrix of the design matrix X with 2p columns has full column rank. However, in practice we have $\sigma^2 > 0$, in which case the condition is still necessary. Shirakura *et al.* (1996) gave some results on probability of correctly searching one active effect. Chen and Lin (1998) derived a lower bound on the probability that the factor with the largest estimated effect is indeed the largest true effect. The bound depends on the relative size of the largest effect and the maximum correlation between columns of the design matrix. Their results have important implications for construction of suitable designs with suciently large value of this probability. Chen and Lin (1998) proved that the absolute value of the correlation between any two columns of the design matrix X should be less than $1/(p - 1)$ in order to estimate any set of p effects. For p odd, this correlation could be equal to $1/(p - 1)$. Using simulation, they obtained probability of correct identification of the largest active effect for the designs given by Lin (1991, 1993, 1995) and Wu (1993).

Abraham, Chipman and Vijayan (1999)

Abraham *et al*. (1999) pointed out that all-subsets regression is to be preferred over stepwise regression since it can identify several candidate models that may explain the data adequately. These candidate models can be further investigated to determine the final set of active factors. However, as they illustrated with the help of Williams (1968) rubber data, all-subsets regression can still be misleading in some cases. They compared the results obtained for the rubber data using eight alternative designs constructed through Lin's (1993) half fractions of Hadamard matrices method. They considered 5 model sizes, with 1 to 5 active factors. Using all-subsets regression, different designs gave different choices of active factors for various model sizes. They showed that selection of active factors is sensitive to design choice. They also considered best fitting models for each of the 8 designs, for each of the five models. However, the findings were not satisfactory, since important active factor(s) were not selected in many cases. Although several candidate models is a better alternative over single best model, their results emphasize one of the main di culties in analysis of supersaturated designs, i.e. different designs may identify different factors as active. Abraham *et al.* (1999) also compared stepwise and all-subsets

regression through simulation under model misspecification. They concluded that all-subsets regression performs significantly better than stepwise regression in identifying the correct model. However, one has to be very careful as frequently it is possible to select inactive factors as active factors.

Beattie, Fong and Lin (2002)

Beattie, Fong and Lin proposed a two-stage Bayesian model selection strategy, combining two recent methodologies: the stochastic search variable selection (SSVS) method of George and McCulloch (1993, 1997) and the intrinsic Bayes factor (IBF) method of Berger and Pericchi (1996). Although the main interest is in the analysis of supersaturated designs, the approach can be used in analyzing any type of dataset, especially when (a) the number of independent variables is large, (b) a relatively small number of these factors are likely to be active, (c) only a relatively small number of observations can be taken, and (d) all active factors have first order effects which are at least as large as interactions and higher order effects. The SVSS used in the first-stage was also utilized by Chipman *et al*. (1997). As discussed above, the SVSS is able to keep all possible models under consideration, however, informative priors are used instead to allow proper posterior distributions. Beattie *et al*. (2002) demonstrated that when SSVS and IBF are combined into a single two-stage procedure, a powerful tool for the analysis of supersaturated designs is obtained. The limitation of IBF that it cannot be used for supersaturated designs data is not shared by SSVS. However, the limitation of SSVS is that it is very sensitive to tuning constant choices and it is unlikely to select one best "objective" model. The strength of IBF is that non-informative priors result in objectivity in model selection. Since SSVS identifies several possible models that provide good fit to the data, the use of informative priors is not so disadvantageous. The main use of SSVS in the two-stage procedure is to find a short list of potential final models, which are then feeded into the IBF method, thereby resulting in objective selection of the final model from this list through use of a noninformative prior in IBF approach. Thus, the two-stage strategy takes advantage of the strengths of each approach.

As already discussed above, one needs to specify tuning constants (c_i, τ_i) in the prior for factor coefficients β. Beattie *et al.* (2002) used the following four values as suggested by George and McCulloch (1993): $(\hat{\sigma}_{\beta_i}, 5)$, $(\hat{\sigma}_{\beta_i}, 10)$, $(0.10\hat{\sigma}_{\beta_i}, 100)$ and $(0.10\hat{\sigma}_{\beta_i}, 500)$

Here $\hat{\sigma}_{\beta_i}$ is the standard error of the least squares estimate of β_i, which were obtained by fitting a model for each β_i separately. One also needs to specify the tuning constants v and λ in the prior for σ. These were taken as v = 3 and $\lambda = \hat{\sigma}^2/3$, where σ^2 is an estimate of σ^2, which was taken from the final model selected by stepwise regression. In the second stage of the approach, noninformative priors of the form $f(\beta, \sigma) \propto \sigma^{-(1+q)}$ with q = 0 were used in the IBF approach.

Beattie *et al.* (2002) described the following step-by-step procedure for their method.

1. Identify all the candidate factors.
2. Center all the predictors and scale them so that each has sum (across observations) of squares equal to 1.
3. Run stepwise regression using a p-value criterion of p = 0.05. Using the final model that stepwise selects, obtain an estimate of the residual variance.
4. Run the SSVS procedure of George and McCulloch (1993) on the full dataset using several different choices of tuning constants (e.g., $(c_i, \tau_i) = (\hat{\sigma}_{\beta_i}, 5), (\hat{\sigma}_{\beta_i}, 10), (0.10\hat{\sigma}_{\beta_i}, 100)$, and $(0.10\hat{\sigma}_{\beta_i}, 500)$, where $\hat{\sigma}_{\beta_i} = \hat{\sigma}$).
5. From each of these four SSVS runs, identify the model with the highest estimated posterior probability. Select only models that are distinguishable from the other sampled models. Also identify any other apparent important factors that may not have been included in the top model, such as factors that exist in 5 of the top 10 sampled models.
6. Combine all the models and factors selected in the previous step into one "encompassing" model.
7. Run the IBF procedure of Berger and Pericchi (1996a) on the encompassing model using the reference prior (or with several different prior choices).
8. Identify the best model selected by IBF using either A IBF or G IBF averaging.

9. Proceed with regression diagnostics on the final model choice to assure oneself of the accuracy of the final selection.

Li and Lin (2002)

The approach introduced by Li and Lin is based on the nonconcave penalized likelihood variable selection (NPL) method of Fan and Li (2001). With a suitable choice of penalty function and regularization parameter in the NPL method, coecients that are close to zero are automatically estimated as zero and other coefficients, i.e. non-zero coefficients, are estimated like the true submodel was known in advance (oracle property). The method, however, requires the design matrix to be of full-rank, and hence it can not be used for analysis of supersaturated designs. Li and Lin (2002) extended the NPL method to nonconvex penalized least squares and non-full rank design matrix. The iterative procedure requires initial value for the regression parameters, which are obtained using stepwise regression. Penalized least squares estimates are obtained by minimizing

$$Q(\beta) = \frac{1}{2n}\sum_{i=1}^{n}(Y_i - x_i'\beta)^2 + \sum_{j=1}^{k} p_\lambda(|\beta_j|)$$

where $p_\lambda(.)$ is a penalty function and is a tuning parameter. The penalty function $p_\lambda(|\beta_j|)$ can be locally approximated by the quadratic function

$$[p_\lambda(|\beta_j|)]' = p_\lambda(|\beta_j|)\,\mathrm{sgn}(\beta_j) \approx [p'_\lambda(|\beta_j^{(0)}|)/|\beta_j^{(0)}|]\beta_j$$

where $\beta_j^{(0)}$ is an initial value of β_j close to its true value, the $\hat{\beta}_j$ is taken to be zero if $\beta_j^{(0)}$ is very close to zero. As mentioned above, initial values $\beta_j^{(0)}$ are obtained using stepwise regression. The penalty function is taken to be the smoothly clipped absolute deviation (SCAD) penalty of Fan and Li (2001), defined by the first-order derivative

$$p'_\lambda(\beta) = \lambda\left[I(\beta \le \lambda) + \frac{(\alpha\lambda - \beta)_+}{(\alpha - 1)\lambda} I(\beta > \lambda)\right]$$

with $p_\lambda(0) = 0$, and I(.) denotes the indicator function. The tuning parameter λ is selected using generalized cross-validation approach of Craven and Wahba (1979). Let $V = \lim_{n\to\infty}(1/n)X'X$, and let V_1 and X_1 be respectively the submatrix of V and the submatrix of X corresponding to the active effects in β. The asymptotic normality and the oracle property for the penalized least squares estimator holds under the conditions $\max_{1\le i\le n} x'_{1i}(X_1'X_1)^{-1}x_{1i} \to 0$, as $n \to \infty$, and V is finite with $V_1 > 0$. To implement the procedure, the SCAD penalty function is locally approximated by the quadratic approximation described above. The penalized least squares estimates are then obtained iteratively from the following ridge regression type expression

$$\beta^{(1)} = \{X'X + n\Sigma_\lambda(\beta^{(0)})\}^{-1}X'Y\ (|\beta_1^{(0)}|)/\beta_1^{(0)}$$

where

$$\Sigma_\lambda(\beta^{(0)}) = \mathrm{diag}\{p'_\lambda(|\beta_1^{(0)}|)/\beta_1^{(0)},\ \ldots,\ p'_\lambda(|\beta_d^{(0)}|)/|\beta_d^{(0)}|\}$$

and d denotes the number of active effects. Li and Lin (2002) showed effectiveness of their procedure using several examples.

Holcomb, Montgomery and Carlyle (2003)

For orthogonal fractions of two-level factorials, least squares estimates of estimable main effects and interactions, except for normalizing constants, are given by $C = X'Y$ where the design matrix X contains effect contrasts in its columns. This simple way of estimating contrasts has sometimes been used in non-orthogonal fractions as well. Holcomb *et al.* (2003) showed that contrast estimates C follow permuted multivariate hypergeometric distribution, which may be approximated by normal distribution. From $Y = X\beta + \in$, it follows that $C = X'X\beta + X' \in$. Partition $\beta = (\beta_p'\ \beta_r')'$ with $\beta_r = 0$ and $\beta_p = (\beta_{p1},...,\beta_{pp}), X = (X_p.X_r)$, $C = (C_p'C_r')'$ with $C_p' = (C_{p1}, C_{p2},..., C_{pp})$ where p and r correspond to active and inactive effects respectively. Then, Holcomb *et al.* (2003) showed that $E(C_r) = (\beta_p' 1)\mu_s, E(C_{pj}) = \beta_{pj}\{n - \mu_s\} + (\beta_p' 1)\mu_s, \sigma_r^2 = (\beta_p'\beta_p)\sigma_s^2\sigma_{pj}^2 = (\beta_p'\beta_p - \beta_{pj}^2)\sigma_s^2$

where $\mu_s = E(s), \sigma_s^2 = E(s^2) - E(s)^2$ are respectively the expected value and variance of an off-diagonal element of the X′X matrix, and 1 is a column vector of 1's of appropriate order. They proposed two methods for

selecting active contrasts, namely the bootstrap method and the contrast variance method (CVM).

The bootstrap method is based on bootstrap samples from the n responses of a supersaturated experiment. Each bootstrap sample is also of size n. For each bootstrap sample, a contrast estimate is obtained by summing the values in the first half of the bootstrap sample and subtracting from it the values in the second half of the bootstrap sample. This process generates the empirical distribution of the contrasts. Then, contrasts falling in the equal tails of this distribution are selected as active effects.

The main objective of their second method, the CVM, is to control Type 2 error rate. The following steps are as given by them.

1. Estimate the number of active factors in the model. (...the process is more robust to moderate over-estimation of p, and ..... that the number of active factors is half the number of experimental runs.)
2. Form all factor contrasts and remove the p contrasts having the largest absolute values.
3. Estimate the variance of the inactive factor contrasts from the remaining contrasts.
4. ... variance of the active factors is approximately equal to the variance of the inactive factors, so assume $\hat{\sigma}_p^2 = \hat{\sigma}_r^2$.
5. Use this estimate of variance, the design properties E(s) and E(s^2), and the bounds on the summation terms to estimate the maximum absolute coefficient

$$\hat{\beta}_{max} = \sqrt{\hat{\sigma}_r^2 / (E[s^2] - E[s]^2) \sum_{i=1}^{P} d_i^2}$$

6. The mean of the extreme contrasts may be estimated as $-n\beta_{max}$ and $+ n\beta_{max}$.
7. Form the upper and lower limits on the critical region with the following : ucl = $n\hat{\beta}_{max} - z_\gamma \hat{\sigma}_p$ where z_γ is the standard normal percentile with an upper tail area of γ.

Contrasts that are not within the interval (lcl, ucl) are selected as active. In the above procedure, $d_i = \beta_{pi}/\beta_{max}$ where β_{max} is the maximum value of the active effects β_p, in absolute value.

Holcomb *et al.* (2003) demonstrated the superiority of the bootstrap and CVM methods through Monte Carlo simulations for two cases

(i) All the regression coecients of active effects are equal, i.e. $\beta_p = \beta_{max} 1$.

(ii) The regression coecients of active effects vary linearly from $-\beta_{max}$ to $+\beta_{max}$.

For these two cases, they presented Type 1 and Type 2 probabilities for their methods, and stepwise regression and all-models regression with some models selectively excluded. However, we remark that although Type 2 error seems to be generally small for their methods, frequently Type 1 error can be as high as 80%. It appears that CVM and bootstrap methods are particularly useful when a significantly high value of Type 1 error is not a disadvantage for the experimenter.

Generalization of Contrast Variance Method

A limitation of CVM is that it requires the effect coefficients to be either all equal or linearly related. In practical situations, it is quite likely that effects corresponding to distinct factors do not bear such a systematic relationship. We now develop a generalization of CVM (GCVM) to cover other situations. Let $\mu_{\beta p} = \beta_p' 1/p, \sigma_{\beta p}^2 = (\beta_p' \beta_p - p\mu_{\beta p}^2)/p$ and $\theta_{\beta p} = \mu_{\beta p}/\sigma_{\beta p}$ be the mean, variance and coefficient of variation of active effects respectively. We assume that the experimenter has some prior knowledge about possible spread in the magnitudes of active effects β_p, i.e. $\sigma^2_{\beta_p}$, and the average size of active effects as measured by the coefficient of variation θ_{β_p}. The range of active effects can then be taken to be approximately $(\theta_{\beta p} - 3)\sigma_{\beta p}$ and $(\theta_{\beta p} + 3)\sigma_{\beta p}$ which ideally corresponds to normal distribution. We then have the following theorem.

Theorem 1. Estimates of the minimum and maximum active effects are given by

(i) $$\hat{\beta}_{min} = \sqrt{\frac{\hat{\sigma}_r^2}{\sigma_s^2} - \{(p-1)\theta_{\beta p}^2\sigma_{\beta p}^2 + (p-9)\sigma_{\beta p}^2 + 6\theta_{\beta p}\sigma_{\beta p}^2\}}$$

(ii) $$\hat{\beta}_{max} = \sqrt{\frac{\hat{\sigma}_r^2}{\sigma_s^2} - \{(p-1)\theta_{\beta p}^2\sigma_{\beta p}^2 + (p-9)\sigma_{\beta p}^2 - 6\theta_{\beta p}\sigma_{\beta p}^2\}}$$

Proof. From the definitions of $\theta_{\beta p}$ and $\sigma_{\beta p}$ it can be seen that $\beta_p'\beta_p = p(1+\theta_{\beta p}^2)\sigma_{\beta p}^2$. Adding $\beta_{min}^2 - \{(\theta_{\beta_p} - 3)\sigma_{\beta_p}\}^2$, which approximately equals zero, to the right hand side of $\beta_p'\beta_p$ and simplifying we get

$$\hat{\beta}_{min} = \hat{\beta}_p'\beta_p - \{(p-1)\theta_{\beta p}^2\sigma_{\beta p}^2 + (p-9)\hat{\sigma}_{\beta p}^2 + 6\theta_{\beta p}\sigma_{\beta}^2\}$$

Also, from $\sigma_r^2 = (\beta_p'\beta_p)\sigma_s^2$ we get $\hat{\beta}_p'\beta_p = \hat{\sigma}_r^2/\sigma_s^2$. Substituting $\hat{\beta}_p'\beta_p$ in the expression proves (i). The proof of (ii) follows along similar lines by adding $\beta_{max}^2 = \{(\theta_{\beta p} + 3)\sigma_{\beta p}\}^2$ which approximately equals zero, to the right hand side of the expression for $\beta_p'\beta_p = p(1+\theta_{\beta p}^2)\sigma_{\beta p}^2$.

The means of the extreme contrasts are then estimated by $\hat{E}(C_{min}) = \hat{\beta}_{min}\{n - \mu_s\} + (\beta_p'1)\mu_s$ and $\hat{E}(C_{max}) = \hat{\beta}_{max}\{n - \mu_s\} + (\beta_p'1)\mu_s$ with $\beta_p'1 = p\sigma_{\beta_p}\theta_{\beta_p}$. Then, following Holcomb *et al.* (2003), form upper and lower limits of the critical region as lcl = $\hat{E}(C_{min}) + z_\gamma\hat{\sigma}_p$ and ucl = $\hat{E}(C_{max}) - z_\gamma\hat{\sigma}_p$. Finally, contrasts that are not within the interval (lcl, ucl) are selected as active.

A simulation study was carried out to study the performance of CVM, bootstrap, GCVM and stepwise regression methods. Only the uniform and linearly varying effects were considered for CVM. For stepwise regression, α_{entry} and α_{out} were both kept at 0.15, and the critical value z_γ for CVM and GCVM was taken to be 2.4, the values used by Holcomb *et al.* (2003). Several different values of $\theta_{\beta p}$ in the range 3 to 10 were used, and the values of $\sigma_{\beta p}^2$ were taken to be in the range 0.1 to 1.0. For uniform and linearly varying effects in CVM, β_{max} = 2, 3, 4, 5 were used. The above set of values were also used for the bootstrap method. Simulations were carried out for the 8-run supersaturated design with 10 factors and the 12-run supersaturated design with 22 factors. The Type 1 and Type 2 errors were based on 5000 simulation runs, while the number of samples for each bootstrap run was kept at 1000. The results of simulation showed less than consistent behavior in Type 1 and Type 2 error rates for both CVM and GCVM. The Type 2 error rates were generally quite low but high values of Type 1 error rates were observed in the simulation runs. Our simulation results for CVM for the two designs we considered were less convincing than those reported by Holcomb *et al*. (2003). Similarly, results we observed for GCVM and the bootstrap methods were also less than convincing. Results obtained using stepwise regression method were significantly better than CVM, GCVM and bootstrap methods.

The simulation study will be extended to all the designs considered by Holcomb *et al*. (2003) and some other designs with larger number of factors and/or runs, results of which will be reported elsewhere.

Allen and Bernshteyn (2003)

The analysis of a design is closely related to its statistical properties. The authors argue that although the popular $E(s^2)$ criterion is statistically meaningful in the supersaturated set-up, by far it is not the best or the only meaningful criterion. Unless a design has good statistical properties, the experimenter can not hope to achieve correct results from analysis of the design. The authors consider four criteria from the perspective of the analysis of supersaturated designs. Evaluation of available supersaturated designs on the basis of these criteria underscores the importance of the criteria in constructions of designs as well. The four criteria are (i) p_{cs}, the probability that the analysis identifies all active and inactive effects correctly, (ii) p_{cov}, the probability the analysis correctly identifies all active effects, (iii) d, the average model size, and (iv) w, the probability of identifying any given active factor. Allen and Bernshteyn (2003) evaluated the 14-run design in 23 factors given by Lin (1993) on the basis of these criteria and concluded

that users of this design "should not expect to achieve correct selection in realistic scenarios." They constructed several designs for number of runs 6, 8, and 10 for number of factors from 7 to 16 using $E(s^2)$, D_2 criterion of Wu (1993), Half-fractions of Plackett-Burman designs, and p_{cov}. Comparisons of these competing designs on the basis of the proposed four criteria using stepwise regression on simulated data resulted in the following conclusions. The designs that are optimal with respect to the p_{cov} are approximately optimal with respect to the other criteria as well. The proposed criteria may sometimes result in unbalanced designs which may be useful when costs of treatment runs are taken into consideration. Further, among several possible $E(s^2)$ optimal designs, one should select a design that is optimal with respect to other criteria.

Yamada (2004)

Yamada conducted an extensive simulation study to investigate Type 1 and Type 2 errors in selecting active factors using stepwise regression. His recommendations reinforce the results of Liu *et al*. (2007) concerning selection of one or two active factors. The simulation results highlight that one should focus on selection of the most important active factors, *e.g.* one or two largest active factors.

The probability of selecting the second largest active factor is maximum in many cases at the third stage of the stepwise regression. Based on this the author recommended that for controlling Type 1 error, it is preferable to select active factors only up to the third stage in stepwise regression. However, if Type 2 error is more important, then as a compromise, active factors may be selected up to the fifth stage in stepwise regression. Furthermore, as observed by Abraham *et al*. (1999), the simulation results indicate that active factors may not be correctly selected if there are several potentially important active factors, e.g. as many as five or more.

Lu and Wu (2004)

It has been established by Abraham *et al.* (1999) and several other authors that all-subsets regression performs better than stepwise regression. However, a major disadvantage of all-subsets regression is that it is too time consuming since the number of potential factors is much larger compared to the number of active factors. Several authors have thus proposed various strategies to reduce the number of model evaluations without substantially sacrificing the efficiency of all-subsets regression. Hamada and Wu (1992) proposed iterative use of forward stepwise regression in the setting where the model may include some significant interactions. Lu and Wu (2004) proposed selecting some large effects in the first stage, followed by selecting active factors from these relatively large effects using stepwise regression. The active factors selected using stepwise regression are combined with the factors that were left out in the first stage and stepwise regression is applied for the second time to come up with the final selection of active factors from this combined list of factors. The selection of factors in the first stage is based on a preestimate of effects. Based on simulation results, the authors provide some guidelines for the minimum size of the effects and the number of effects to include in the first stage.

Based on simulation results, the authors claimed that their procedure was more accurate than the usual stepwise regression procedure in identifying active factors and the correct model size. However, the results were not uniformly superior across various supersaturated designs.

Zhang, Zhang and Liu (2006)

Zhang, Zhang and Liu introduced a variable selection method based on partial least squares (PLS) regression. The PLS regression blends the frameworks of principal components and canonical correlation with multiple regression, see *e.g.* Bastien *et al.* (2005). The i^{th} PLS component is defined by $t_i = X_s w_i$, where the weights $w_i = [w_{i1}, w_{i2}, \ldots, w_{ik}]'$, with $w_i' w_i = 1$ are determined using maximal correlation of Y_s with PLS components. Here Y_s and X_s are standardized versions of the response Y and the design matrix X, i.e. each of the columns of $[Y_s X_s]$ has zero mean and unit variance. Also, note that $corr(Y_s, X_{sj}) = cov(Y_s, X_{sj})$, as both Y_s and X_{sj}, the j^{th} column of X_s, are standardized.

For the first PLS component t_1, weights w_{il} as given below are obtained by maximizing $corr(Y_s, t_1)$, and they correspond to the standardized eigen vector of $X_s' Y_s Y_s' X_s$ with the largest eigen value

$$w_{1l} = \frac{corr(Y_s, X_{sl})}{\sqrt{\sum_{j=1}^{k} \{(corr(Ys, X_{sj})\}^2}}$$

The main feature of PLS regression is that now it treats the first PLS component t_1 as an independent variable not only for the response variable Y_s, but also for each of

the variables $X_{s1}, X_{s2}, \ldots, X_{sk}$. Let Y_{s1} be the residual vector obtained from Y_s after fitting t_1 to it. Similarly, X_{sj1} be the residual vector obtained from X_{sj} after fitting t_1 to it, j = 1, ... , k. The weights for the second PLS component $t_2 = \sum_{l=1}^{k} w_{2l} X_{sl1}$ are then given by

$$w_{2l} = \frac{\text{cov}(Y_{s1}, Y_{sl1})}{\sqrt{\sum_{j=1}^{k} \{\text{cov}(Y_{s1}, X_{sj1})\}^2}}$$

The next PLS component is obtained in a similar way by considering t_2 as an independent variable for the residuals Y_{s1}, X_{sj1}, j = 1, ... , k, and the procedure is repeated until the partial covariances are no longer significant. Finally, in PLS regression, a multiple linear regression model fits to Y_s using the PLS components t_1, t_2, ... , t_m as independent variables, where m denotes the number of PLS components. The relative importance of the l effect variable is judged through the variable importance in projection (VIP) criterion defined by

$$\text{VIP}_l = \sqrt{\frac{k}{\sum_{j=1}^{m} \{\text{corr}(Y_s, t_j)\}^2} \sum_{j=1}^{m} \{\text{corr}(Y_s, t_j)\}^2 w_{jl}^2}$$

A variable with high VIP value tends to be more important and it is likely to be chosen as an active factor. However, the selection of active factors is not based on VIP values alone. The selection is fine-tuned using another criterion called Mpress. The whole process of selecting active factors involves repeated application of the PLS regression. Let e_{ij} be the residual for the i^{th} response obtained from the model with only the j^{th} predictor in it and fit after excluding the ith response. Then, for the j^{th} variable, $\text{Press}_j = \sum_{i=1}^{n} e_{ij}^2$. Likewise, let e_{ijl} be the corresponding residual obtained from a model with l – 1 predictors, and the j^{th} predictor in it; clearly these l – 1 predictors, and the j^{th} predictor are all distinct. Then, for the j^{th} variable, $\text{Press}_{jl} = \sum_{i=1}^{n} e_{ijl}^2$ Zhang *et al.* (2006) integrate a penalty function into Press_{jl} and define

$$\text{Mpress}_{jl} = \frac{\text{Press}_{jl}}{2(n-l)} + \frac{2l}{n}$$

They use Mpress_{jl} to decide which factor to select among two candidate active factors with largest VIP values, and also to decide when to stop the process of active factor selection. The active factor selection continues as long as there is a decrease in Mpress_{jl} criterion value; the procedure terminates when the criterion value increases as a result of adding another active factor in the model. From the two variables with the largest VIP values, the variable having smaller value of Mpress_{jl}, with l = 1, is chosen to be the first active factor. Let Y_1 be the residual vector after fitting the first active factor to the response Y. The m PLS components are now recomputed using the remaining k – 1 factors and the residual vector Y_1. The VIP criterion values are similarly computed for each of the remaining k – 1 factors. From the two candidate factors having largest VIP values, the factor having smaller value of Mpress_{jl} with l = 2, is chosen as the second active factor. This procedure of selecting active factors is repeated as long as there is a reduction in Mpress_{jl} criterion value, otherwise the procedure concludes. Zhang *et al.* (2006) refer to their procedure as PLSVS method. The PLSVS method is applicable to multi-level and mixed level supersaturated designs as well.

Liu, Ruan and Dean (2007)

Liu, Ruan and Dean also emphasize the fact that correct analysis of a design is directly linked to its statistical properties. They investigated selection of correct active factors using forward selection through consideration of the expected values of the quadratic forms involved in the F statistics of different effects. For the case of one active factor, they proved that forward selection method will always select the correct active factor when random errors are small relative to the size of the active effect. In general, there may be more than one active factor present. For the case of two or more active factors, they showed that correct selection of the active factor with the largest effect in absolute terms (the "most active" factor) depends on the extent of the correlation between the columns of the design matrix. They obtained bounds on the maximum correlation between columns of the design matrix to ensure the correct selection of the most active factor. Liu *et al.* (2007) also investigated selection of the two most active factors using best subset selection method. For the case of two active factors, they proved that the best subset method will select the correct two active factors if the maximum correlation between columns of the design matrix is less than 0.33. They also listed several designs that satisfy their bounds on the maximum correlation between the columns of the design matrix.

Srivastava (1975)

As we have already mentioned, Srivastava (1975) derived a condition that must be satisfied by a supersaturated design before the problem of identifying active factors can be solved at all. Let f_0 denote the maximum number of factors such that when the design is projected into any subset of f_0 factors, all the main effects in the projected design are estimable. This means that there is at least one subset of $f_0 + 1$ factors for which not all main effects in the projected design are estimable. Deng *et al*. (1999) defined f_0 as the resolution rank, r-rank of a supersaturated design. Among alternative designs with the same number of runs, it will be desired to select the one with the largest value of r-rank. Srivastava (1975) proved that a supersaturated design will not be able to search the active factors and estimate their effects unless its r-rank is at least 2p. The supersaturated designs satisfying the r-rank condition are called search designs. Shirakura *et al.* (1996) gave some results on searching probabilities for main effects designs for searching one active interaction. Srivastava's (1975) procedure for searching active factors is as follows.

We have the search linear model $Y = \mu 1 + X\beta + \in$, $V(\in) = \sigma^2 I$ where β includes both active and inactive effects. Srivastava (1975) gave the following fundamental formulation of the search problem. Under the above model with $\sigma^2 = 0$, a necessary and sufficient condition that the problem can be completely solved is that for every submatrix X_1 of order $n \times 2p$ of X, $\text{rank}[1\ X_1] = 1 + 2p$ holds. The condition remains necessary for the practical case $\sigma^2 > 0$. For only active effects, the model is $E(Y) = \mu 1 + X_p\beta_p$, where the subscript p refers to the active effects. The sum of squares for this model is given by $ss(\beta_p) = Y'\{I - Q(\beta_p)\}Y$, where $Q(\beta_p) = A(A')^{-1}A'$ with $A = [1 \quad X_p]$. Srivastava (1975) proposed the following procedure for searching active effects.

Step 1: Compute $ss(\beta p)$ by taking all possible $p \times 1$ sub-vectors of β.

Step 2: Find a vector β_p for which $ss(\beta_p)$ is minimized.

However, this does not guarantee selection of correct p active factors. Thus, probability of correct searching for a design d is computed as follows

$$P(d) = \min_{\beta_p \subset \beta} \min_{\beta_{1p} \in A(\beta_{2p})} P(ss(\beta_p) < ss(\beta_{1p}))$$

where $A(\beta_{2p})$ denotes the set of all possible sub-vectors of size p of β of which at least one element is not in β_p. Shirakura *et al.* (1996) suggested the following for comparing two competing designs. Let d and d_1 be two competing supersaturated designs. Calculate the searching probabilities P(d) and $P(d_1)$ for these designs. Then the design d is to be preferred over d_1 if $P(d) > P(d_1)$.

3. CONCLUDING REMARKS

Full-rank models permit estimation of the parameters of the model through usual linear model regression methods. Identification and estimation of active factors in supersaturated designs is considerably more complex. Srivastava (1975) considered identification of the correct model in the non-full-rank setting when, in addition to the parameters that are a priori to be included in the model, e.g. main effects, there may exist few other non-negligible parameters (or active effects) unknown to the experimenter that also need to be estimated. He established that the analysis of a supersaturated design has problems of estimability unless its r-rank is at least 2p. This is a necessary condition for a supersaturated design to be useful - namely, for the analysis to have any hope of success. Since n, the number of runs, is much smaller than the number of factors f, the necessary condition implies that the number of active factors should be small. Furthermore, Chen and Lin (1998) proved that the absolute value of the correlation between any two columns of the design matrix X should be less than $1/(p-1)$ in order to estimate any set of p effects. In other words, larger the number of active effects, smaller the pairwise column correlations must be. It would then seem that either p must be quite small relative to n or the design while supersaturated must be close to being merely saturated. There are several other inherent difficulties in the analysis of supersaturated designs. Although only few factors may be active under effect sparsity, their exact number (or model size) is unknown in general. When the design satisfies the necessary condition of Srivastava (1975), the probability of correct identification of active effects depends on the magnitude of correlations between columns of the design matrix. The results of Abraham *et al.* (1999) emphasize a major difficulty in analysis of supersaturated designs, i.e. different designs may identify different factors as active. They also observed the importance of rather extreme effect sparsity and concluded that active factors may not be correctly selected if there are several potentially important active factors, *e.g.* as many as five or more.

The method of Li and Lin (2002) based on nonconvex penalized least squares appears promising and it should be explored more extensively. Liu *et al*. (2007) emphasize the fact that correct analysis of a design is directly linked to its statistical properties. For the case of (only) two active factors, they proved that the best subset method will select the correct two active factors if the maximum correlation between columns of the design matrix is less than 0.33, which would seem to be a rather severe bound. The contrast variance method and its generalization, and the bootstrap method do not seem to perform better than stepwise regression. Therefore, forward selection, stepwise regression, and all-subsets regression methods should be preferred over ad-hoc methods developed using simulation based procedures.

An additional complexity in identification of active factors is the possibility that some interactions among factors may also be present. Interactions may not only mask the main effects but may also result in different sets of active effects being identified by different methods. Lewis and Dean (2001) introduced two-stage group screening methods to study interactions in addition to main effects. Hamada and Wu (1992) also investigated two-factor interactions but their approach is limited in the sense that it does not really takes into account effect sparsity and supersaturated nature of the experiment, including all two-factor interactions in the second stage. Group screening method was first used by Watson (1961) for screening main effects. In the two-stage group screening method of Lewis and Dean (2001), the factors are partitioned into sets or groups of factors. Using prior information, factor levels are labeled so that the main effects of all the factors within a group are in the same direction, i.e. positive (or negative). They developed two group screening strategies, one that considers only main effect of each grouped factor in stage 1, and the other in which both two-factor interactions and main effects of grouped factors are considered in the first stage. Grouping of factors in the groups selected in the first stage is disbanded in the second stage, and main effects and two-factor interactions among factors are examined in both the strategies. We refer to Lewis and Dean (2001), Dean and Lewis (2002), and Vine *et al.* (2008) for further details of the two-stage group screening methods.

In a practical situation, identification and estimation of active factors should be based on a careful evaluation of the results obtained across several competing methods. Finally, irrespective of the method of analysis used, it is to be expected that the active factors selected in the screening phase of the experiment will be evaluated further in a follow-up confirmatory experiment.

ACKNOWLEDGEMENTS

The authors thank the referee for helpful comments that led to substantial improvements, especially in the concluding remarks section.

REFERENCES

Abraham, B., Chipman, H., Vijayan, K. (1999). Some risks in the construction and analysis of supersaturated designs. *Technometrics,* **41,** 135-141.

Allen, T.T. and Bernshteyn, M. (2003). Supersaturated designs that maximize the probability of identifying active factors. *Technometrics*, **45,** 90-97.

Bastien, P., Vinzi, V.E. and Tenenhaus, M. (2005). PLS generalized linear regression. *Comp. Statist. Data Anal.*, **48**, 17-46.

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,** 55-63.

Berger, J.O. and Pericchi, L.R. (1996). The intrinsic Bayes factor for model selection and prediction. *J. Amer. Statist. Assoc*., **91,** 109-122.

Booth, K.H.V. and Cox, D.R. (1962). Some systematic supersaturated designs. *Technometrics*, **4,** 489-495.

Box, G. E. P. and Meyer, R. D. (1993). Finding the active factors in fractionated screening experiments. *J. Qual. Tech*., **25,** 94-105.

Chen, J., and Lin, D.K.J. (1998). On the identifiability of a supersaturated design. *J. Statist. Plann. Inf*., **72,** 99-107.

Chipman, H. (1996). Bayesian variable selection with related predictors. *Canadian J. Statist*., **24,** 17-36.

Chipman, H., Hamada, M. and Wu, C. F. J. (1997). A Bayesian variable selection approach for analyzing designed experiments with complex aliasing. *Technometrics*, **39,** 372-381.

Craven, P. and Wabha, G. (1979). Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numer. Math*., **31,** 377-403.

Dean, A.M. and Lewis, S.M. (2002). Comparison of group screening strategies for factorial experiments. *Comp. Statist. Data Anal*., **39,** 287-297.

Deng, L.-Y., Lin, D.K.J. and Wang, J. (1999). A resolution rank criterion for supersaturated designs. *Statistica Sinica*, **9,** 605-610.

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.,* **96,** 1348-1360.

George, E.I. and McCulloch, R.E. (1993). Variable selection via Gibbs sampling. *J. Amer. Statist. Assoc.,* **88,** 881-889.

George, E.I. and McCulloch, R.E. (1997). Approaches for Bayesian variable selection. *Statistica Sinica*, **7,** 339-373.

Hamada, M. and Balakrishnan, N. (1998). Analyzing unreplicated factorial experiments: A review with some new proposals. *Statistica Sinica*, **8,** 1-41.

Hamada, M. and Wu, C.F.J. (1992). Analysis of designed experiments with complex aliasing. *J. Qual. Tech.,* **24,** 130-137.

Holcomb, D.R., Montgomery, D.C. and Carlyle, W.M. (2003). Analysis of supersaturated designs. *J. Qual. Tech.*, **35,** 13, 27.

Lewis, S. M., and Dean, A. M. (2001). Detection of interactions in experiments on large numbers of factors (with discussion). *J. Roy. Statist. Soc.,* **B63,** 633-672.

Li, R. and Lin, D.K.J. (2002). Data analysis in supersaturated designs. *Statist. Prob. Lett.*, **59,** 135-144.

Lin, D.K.J. (1991). Supersaturated designs. Working Papers, 264, Coll. Bus. Admn., Univ. of Tennessee.

Lin, D.K.J. (1993). A new class of supersaturated designs. *Technometrics*, **35,** 2831.

Lin, D.K.J. (1995). Generating systematic supersaturated designs. *Technometrics*, **37,** 213-225.

Liu, Y., Ruan, S. and Dean, A.M. (2007). Construction and analysis of Es^2 ecient supersaturated designs. *J. Statist. Plann. Inf.*, **137,** 1516-1529.

Lu, X. and Wu, X. (2004). A strategy of searching active factors in supersaturated designs. *J. Qual. Tech.,* **36,** 392-399.

Nguyen, N.-K. (1996). An algorithmic approach to constructing supersaturated designs. *Technometrics*, **38,** 69-73.

Samset, O. and Tyssedal, J. (1998). Study of the Box-Meyer method for finding active factors in screening experiments. *Tech. Rep*., 5, Dept. Math. Sci., Norwegian Univ. of Sci. & Tech., Norway.

Satterthwaite, F. (1959). Random balance experiments (with discussion). *Technometrics*, **1,** 111-137.

Shirakura, T., Takahashi, T. and Srivastava, J.N. (1996). Searching probabilities for non-zero effects in search designs for the noisy case. *Ann. Statist*., **24,** 2560-2568.

Srivastava, J.N. (1975). Designs for searching non-negligible effects. In: *A Survey of Statistical Designs and Linear Models*, J.N. Srivastava, ed., North Holland, Amsterdam, 507-519.

Vine, A.E, Lewis, S.M., Dean, A.M., and Brunson, D. (2008). A critical assessment of two-stage group screening through industrial experimentation. *Technometrics*, **50,** 15-25.

Wang, W. and Voss, D.T. (2001). On the analysis of nonorthogonal saturated designs using effect sparsity. *Stat. Appl.*, **3,** 177-192.

Watson, G.S. (1961). A study of the group screening method. *Technometrics*, **3,** 371-388.

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,** 101-117.

Williams, K.R. (1968). Designed experiments. *Rubber Age*, **100,** 65-71.

Wu, C.F.J. (1993). Construction of supersaturated designs through partially aliased interactions. *Biometrika*, **80,** 661-669.

Xu, H. and Wu, C.F.J. (2005). Construction of optimal multi-level supersaturated designs. *Ann. Statist*., **33**, 2811-2836.

Yamada, S. (2004). Selection of active factors by stepwise regression in the data analysis of supersaturated designs. *Qual. Engg*., **16,** 501-513.

Zhang, Q.-H., Zhang, R.-C. and Liu, M.-Q. (2006). A method for screening active effects in supersaturated designs. *J. Statist. Plann. Inf*., **137,** 2068-2079.