Quality Engineering, 23:15–25, 2011

ISSN: 0898-2112 print/1532-4222 online
DOI: 10.1080/08982112.2010.529483

Taylor & Francis
Taylor & Francis Group

Screening for Fuel Economy: A Case Study of Supersaturated Designs in Practice

Philip R. Scinto[1],
Robert G. Wilkinson[1],
Dennis K. J. Lin[2]

[1]The Lubrizol Corporation, Cleveland, Ohio
[2]Penn State University, University Park, Pennsylvania, and Renmin University of China, Beijing, China

ABSTRACT A successful use of supersaturated design and analysis is demonstrated through a case study completed at the Lubrizol Corporation. In the study, a 28-run supersaturated design is used to screen the effects of more than 70 possible model terms (linear effects, quadratic effects, interactions, and measured covariates) on engine motor oil coefficient of friction (COF). Of the over 70 model terms of interest, 50 are two-way linear interactions. A Lubrizol-developed model-averaging technique known as *Bayesian variable assessment* (BVA) is used to identify the important high-level factors and model terms from the experiment.

This study is unique in the literature due to complications in multiple factor levels, physical correlations and constraints on the factors, curvature, and the desire to screen for a large amount of interactions. The test results are subject to common cause variation and unknown special causes such as operator error and test instrument error.

Due to time and cost constraints, supersaturated designs are necessary to screen for phenomena such as gasoline-powered engine fuel economy. Based on the results from a 10-run follow-up experiment, the use of the supersaturated design analyzed using BVA is concluded to be a success in this case study.

KEYWORDS Bayesian variable assessment, design and analysis of supersaturated design, Markov chain Monte Carlo

INTRODUCTION

The Lubrizol Corporation is an innovative specialty chemical company that produces and supplies technologies that improve the quality and performance of our customers' products in the global transportation, industrial, and consumer markets. One particular area of interest to Lubrizol researchers is the improvement of gasoline-powered engine fuel economy through enhancement and improvement of engine oil additives. Given the large number of possible variables, the potential for interactions, the cost of experimentation, and the limited timeframe in which to deliver on the next engine oil industry standard, a design was necessary to effectively screen the numerous variables of interest in a timely, cost-effective manner.

Address correspondence to Dennis K. J. Lin, The Smeal College of Business Administration, Pennsylvania State University, 483 Business Building, University Park, PA 16802-3603. E-mail: DKL5@psu.edu

Two-level screening designs such as Plackett-Burman (which is a type of Hadamard matrix) and fractional-factorial designs have been widely used to efficiently study potential factors and variables that may have a significant impact on a process under study. In these designs, the number of factors approaches, but is always less than, the number of experimental design points. In supersaturated designs, the number of factors and model degrees of freedom can far exceed the number of design points. In order for this approach to be successful, the assumption that most factors and variables are not important must hold. In industrial settings, this is not unreasonable (see Box and Meyer 1986).

Supersaturated designs are not a new concept; for example, see Satterthwaite (1959). More recently, there has been increased activity into this area. The focus of most of the recent research is on algorithms for generating the designs. In Lin (1993), half fractions of Hadamard matrices were used to create supersaturated designs. A Hadamard matrix is a square matrix with entries of either +1 or −1. The rows of a Hadamard matrix are mutually orthogonal and every two different rows have matching entries in exactly half of their columns and mismatched entries in the remaining columns. These properties hold for the columns as well as the rows. Therefore, a 12×12 Hadamard matrix may be used to study up to 11 linear main effects (columns) in 12 runs (the rows). The analysis in this case would be fairly straightforward and relatively easy because the two-way correlations between linear main effects (the columns) are 0. However, if one were to take one of the 12 columns and eliminate all rows in which there is a +1 in the column (selection may also be based on −1), the resulting matrix is 11 columns and 6 rows. This resulting matrix is a supersaturated design in which 10 linear main effects may be studied in 6 runs (the rows). The analysis is more difficult because of the supersaturation of the main effects, and the two-way correlations between linear main effects (the columns) is now greater than 0 (in this particular example, $|r| = 0.33$ for every two-way correlation).

In Wu (1993), supersaturated designs were created by augmenting Hadamard matrices with interaction columns. In Lin (1995), numeric algorithms were given to construct designs where the pair-wise correlations are all less than a specified threshold. Nguyen (1996) proposed a general method for creating supersaturated designs using balanced incomplete block designs (BIBD). W. Li and Wu (1997) use a column exchange algorithm to search for the "best" supersaturated design. In Yamada and Lin (1997), supersaturated designs containing an orthogonal base were proposed. Other approaches include using cyclic generators (Liu and Dean 2004), utilizing a minimax criterion (Butler 2005), and so on.

Research into the analysis aspect of supersaturated designs is somewhat limited, with mixed results. Westfall et al. (1998) explored forward selection methods for finding active factors. Abraham et al. (1999) used simulation studies to show that model selection can be tricky. Kelly and Voelkel (2000) investigated the asymptotic power issue. In Beattie et al. (2002), a two-stage Bayesian model selection approach was used to select the best model. R. Li and Lin (2003) proposed a penalized least squares approach. Allen and Bernshteyn (2003) combined the analysis and design aspects, proposing designs to maximize the chances of selecting the active factors, assuming a stepwise regression approach. Holcomb et al. (2003), Koukouvinos and Stylianou (2005), and Georgiou (2008) explored approaches based on contrasts. Lu and Wu (2004) suggested a three-stage approach that utilizes stepwise regression. Zhang et al. (2006) utilized partial least squares for selecting active factors, whereas Phoa et al. (2009) used an approach based on minimizing complexity as measured by the l_1-norm (Dantzig selector).

For this article, the focus will not be on developing new design or analysis methods but on the application of supersaturated designs in practice. We will use a formulating example consisting of over 70 different possible model terms (17 linear main effects, six quadratic effects, 50 two-way interactions, and three covariates). By using a supersaturated design, far less data was collected (28 experimental runs) than would have been required by either a fractional-factorial or Plackett-Burman design. This example will highlight issues that may arise when designing supersaturated experiments, including nonindependence of factor settings, mixed levels of factors, and the need to study interactions. With regard to analysis, a Bayesian model-averaging approach called *Bayesian variable assessment*

(Meyer and Wilkinson 1998) will be utilized. This approach has been successfully applied in data mining applications, and data from supersaturated design has much in common with data mining in that both situations need to deal with a large number of variables.

In the next section, we will describe the formulating problem in some detail. BVA will then be described. Included in the same section is a simulation study demonstrating the value of BVA in analyzing data from supersaturated designs. Next, the formulating data from the actual experiment are analyzed. The final section contains a summary and discussion.

PROBLEM FORMULATION

Motor vehicle fuel economy is a topic of great interest to consumers, equipment manufacturers, and researchers. It is not difficult to understand that improvements in fuel economy would have a large, beneficial impact on both the economy and the environment. Items ranging from vehicle and engine design, fuel type and quality, tires, road surface, and driving habits all contribute to the rate at which fuel is consumed while driving an automobile. It has been hypothesized that the engine motor oil may also play a significant role in fuel economy. Therefore, research scientists at Lubrizol have studied, and continue to study, chemical recipes and formulations for engine motor oils to improve fuel economy.

One test among several used to evaluate the performance of engine motor oil in fuel economy is a high-frequency reciprocating test that measures the coefficient of friction (COF) of oil. This test is called the *high-frequency reciprocating rig* (HFRR) ramp and the response used in the analysis is the average COF of the oil from 70 to 74 minutes in this 75-minute test. A lower COF is an indication of oils that reduce friction, which would theoretically translate into better fuel economy in a mechanical engine. Engine motor oil is comprised of base oil, a by-product of the crude oil distillation process, and chemical additives such as dispersants, detergents, oxidation inhibitors, rust inhibitors, viscosity modifiers, and friction modifiers. Lubrizol formulators are interested in identifying chemical additives and the combinations of chemical additives that reduce friction. Unfortunately, historically, signals have been small and infrequent when experimenting with chemical additives in the areas of friction and fuel economy.

Given the historic lack of signal in testing for fuel economy and friction, a high-level screening design in which many factors could be tested was selected as the design approach. It was very important to test as many potential variables as possible because we did not expect a high success rate for signals. Collaboration between Lubrizol statisticians and formulators led to the development of the chemical additive factors, linear terms, curvature terms, and interaction terms to study the COF response.

The chemical additive factors in each formulation to study the average coefficient of friction are broken down as follows: There are two dispersant types, D1 and D2, which are exclusive, meaning that they cannot be used in the same formulation. D1 has four levels and D2 has three levels. In addition, D1 is made in both a low-nitrogen version (0.81 N) and a high-nitrogen (1.0 N) version, whereas D2 can only physically be made as the high nitrogen version. D1 and D2 are also made in their original form and in an altered form (altered by adding a chemical element), independent of level and nitrogen version. Three design columns (A, B, C) indicate D1, D2, and N class. D1 is set equal to 7 when columns A and B are both +1, 5.5 when A is +1 and B is −1, and 4 when A is −1 and B is +1. In these cases, column C indicates the nitrogen version of D1. When columns A and B are both −1, D1 is set to 0 and D2 is set to either 4 or 7 depending on column C.

There are four friction modifier (FM) types: FMA, FMB, FMC, and FMD, each with four levels. The total amount of friction modifier is held constant at a theorized commercial level. Five design columns (D, E, F, G, H) are used to construct FMA, FMB, FMC, and FMD. Columns D and E are used to select one of the four friction modifiers and columns F and G are used to select the other. If the same friction modifier is selected by D and E and by F and G, then only that friction modifier is used in the formulation. Column H is used to select the level of total friction modifier divided equally among the friction modifiers in the formulation.

There are two levels of detergent substrate and two levels of detergent ash. Unfortunately, ash and substrate are physically correlated (detergents that

are higher in ash are lower in substrate) and it is very difficult to simultaneously achieve their respective high levels in the experiment without increasing the correlation among the detergent types. There are six detergent types that are contributors to detergent substrate and detergent ash. These detergent types are studied because it is possible that the response may depend on the detergent types instead of, or in addition to, substrate and ash. Given that the six detergent types, T1, T2, T3, T4, T5, and T6, are the contributors to substrate and ash, they are correlated with substrate and ash. There is also correlation among the detergent types because high-substrate detergents must be used to increase substrate, high-ash detergents must be used to increase ash, and some combination of both must be used to increase both substrate and ash. Five design columns (J, K, L, M, N) are used to construct substrate, ash, and the detergent types. Columns J and K are used to set the low and high levels of substrate and ash. Column L is for the use of T4 or T5, column M is for the use of T3 or T6, and Column N is for the level of T3 if column M indicates that T3 is to be used. T1, T2, T3, T4, T5, and T6 are adjusted to meet substrate and ash levels. Note that it was predetermined by collaboration with formulators that T4 and T5 would not occur in the same formulation and that T3 and T6 would not occur in the same formulation. Also note that the columns only give a rough idea of factor settings. Trial and error must be used to achieve substrate and ash targets and limit the correlation among the detergent types.

Finally, the altered form of dispersant is indicated by column O, the two levels of a Z agent are indicated by column P, and there are three covariates based upon measured physical characteristics of the formulated oil.

In studying the main effects, including curvature, covariates, and the interactions, factorial designs, fractional-factorial designs, and even highly fractionated factorials are of little practical use given the size of such designs for this particular study. Given the sheer number of variables and possible model terms, the constraints on the variables, and constraints on the size of the study, a supersaturated design was constructed. Due to the nonindependence of variables setting and mixed levels of factors, it is not difficult to notice that such a supersaturated design cannot be constructed by simply taking a half fraction of a Hadamard matrix. Therefore, a Plackett-Burman was augmented in order to create the supersaturated design.

The design was initially constructed using a 24-run Plackett-Burman. Three columns were used to construct the dispersant effects, five columns were used to construct the friction modifier effects, and five columns were used to construct the detergent effects. The Z effect and the altered form of the dispersant were determined using one column apiece. Four additional experimental runs were added, through trial and error, to improve the correlation structure; better decouple the factors, variables, and possible model terms of interest; and allow for experiments where the formulations contain all of the friction modifiers and none of the friction modifiers.

The design settings for the linear main effects and the COF response results from the experiment are listed in Table 1. The first 24 runs in the table are from a Plackett-Burman, and the last 4 runs in the table (shaded in the table) are the added, augmented runs. After constructing columns for curvature and interactions, there are over 70 different possible model terms (17 linear main effects, six quadratic effects, 50 two-way interactions, and three covariates) to study in 28 experimental runs. It is quite possible, and even likely, that a design with reduced correlation among the variables and greater power to determine effects could have been developed. However, this design was constructed to be "good enough" in terms of the correlation structure in the interest of time.

To some, this form of design may not seem satisfactory, and it may not even seem to qualify as a supersaturated design. Note that this design is a supersaturated design. It is supersaturated because there are more degrees of freedom for the variables and model terms studied (none of which are completely confounded) than design points. However, it is quite possible that the design is not optimal in terms of correlation structure among the variables. Given that time is a limited resource, the optimal correlation structure was not, and could not be, extensively researched.

One reason we were comfortable with the design was that approximately 80% of the two-way correlations among the main effects were less than 0.25, and only a very small percentage of the two-way correlations were greater than 0.5 (as displayed in Table 2).

TABLE 1 **Design Settings Recorded COF Test Results**

D1	D2	Altered	N Class	FMA	FMB	FMC	FMD	Substrate	Ash	T1	T2	T3	T4	T5	T6	Z Agent	KV100C	HTHS	CCS	COF
7.0	0	YES	0.81	0	2	2	0	0.372	0.7010	0.975	0.00	0.250	0.20	0.000	0.0	0.045	11.638	3.21	6072	0.321
5.5	0	NO	1.00	1	0	0	1	0.378	0.6573	1.150	0.00	0.000	0.10	0.000	0.2	0.076	10.840	3.11	5632	0.259
7.0	0	YES	1.00	1	0	1	0	0.806	0.2626	0.000	0.00	0.000	0.60	0.000	1.0	0.076	11.625	3.15	6539	0.298
0.0	7	NO	1.00	0	0	4	0	0.751	0.7010	0.900	0.00	0.1 25	0.85	0.000	0.0	0.045	11.292	3.24	6428	0.203
5.5	0	YES	1.00	0	2	0	2	0.719	0.3916	0.000	0.30	0.000	0.00	0.550	1.0	0.045	10.873	3.11	5983	0.367
7.0	0	YES	1.00	4	0	0	0	0.373	0.7133	1.250	0.00	0.000	0.00	0.100	0.2	0.045	11.500	3.30	61 82	0.300
5.5	0	NO	0.81	0	1	0	1	0.779	0.6820	0.000	1.00	0.250	0.70	0.000	0.0	0.045	11.071	3.18	5778	0.31 9
5.5	0	NO	0.81	2	0	2	0	0.372	0.3840	0.450	0.00	0.1 25	0.40	0.000	0.0	0.076	10.965	3.13	5696	0.309
0.0	4	YES	1.00	0	2	0	0	0.378	0.6573	1.150	0.00	0.000	0.10	0.000	0.2	0.076	10.075	2.93	5022	0.325
4.0	0	YES	1.00	1	0	0	1	0.756	0.3450	0.000	0.00	0.250	1.10	0.000	0.0	0.076	10.395	3.00	5235	0.344
0.0	4	NO	1.00	2	0	0	0	0.379	0.4003	0.550	0.00	0.000	0.00	0.150	0.5	0.045	10.818	3.06	5058	0.304
4.0	0	NO	1.00	0	4	0	0	0.838	0.6893	0.900	0.00	0.000	0.50	0.000	0.8	0.045	10.493	2.92	5401	0.286
7.0	0	NO	0.81	0	0	2	2	0.373	0.7133	1.250	0.00	0.000	0.00	0.100	0.2	0.076	11.858	3.33	5916	0.260
7.0	0	NO	1.00	1	0	0	1	0.380	0.6860	1.000	0.00	0.1 25	0.00	0.400	0.0	0.076	11.670	3.31	5966	0.288
0.0	4	YES	1.00	0	0	2	0	0.732	0.7093	0.800	0.00	0.000	0.00	1.250	0.2	0.045	10.355	2.98	51 58	0.284
0.0	7	YES	1.00	0	2	2	0	0.374	0.7030	0.900	0.00	0.250	0.00	0.375	0.0	0.076	11.292	3.26	61 03	0.300
4.0	0	YES	0.81	0	1	1	0	0.728	0.7110	0.200	0.70	0.1 25	0.00	1.200	0.0	0.076	10.575	3.06	51 74	0.332
4.0	0	YES	0.81	0	2	2	0	0.372	0.3840	0.450	0.00	0.1 25	0.40	0.000	0.0	0.045	10.571	3.04	5114	0.31 9
0.0	7	NO	1.00	2	2	0	0	0.623	0.4260	0.000	0.00	0.200	0.00	1.500	0.0	0.076	11.403	3.23	6334	0.297
4.0	0	NO	1.00	0	0	1	1	0.365	0.3933	0.600	0.00	0.000	0.25	0.000	0.2	0.045	10.457	2.87	51 29	0.297
7.0	0	NO	0.81	1	1	0	0	0.623	0.4260	0.000	0.00	0.200	0.00	1.500	0.0	0.045	11.959	3.36	6277	0.311
4.0	0	NO	0.81	0	0	0	4	0.379	0.3973	0.500	0.00	0.000	0.00	0.500	0.2	0.076	10.541	2.90	51 63	0.232
5.5	0	YES	0.81	2	0	0	2	0.728	0.2083	0.000	0.00	0.000	1.00	0.000	0.2	0.076	11.1 64	3.08	5529	0.329
5.5	0	YES	1.00	0	0	0	2	0.377	0.3870	0.300	0.00	0.1 25	0.00	0.750	0.0	0.045	11.118	3.14	5586	0.348
7.0	0	NO	0.81	1	1	1	1	0.376	0.7030	0.650	0.88	0.000	0.00	0.000	0.0	0.045	11.967	3.36	6262	0.290
4.0	0	NO	1.00	1	1	1	1	0.770	0.4389	0.000	0.50	0.000	0.00	0.000	1.5	0.045	10.732	3.06	5551	0.300
7.0	0	NO	1.00	0	0	0	0	0.756	0.3450	0.000	0.00	0.250	1.10	0.000	0.0	0.076	11.821	3.33	6514	0.361
4.0	0	NO	0.81	0	0	0	0	0.376	0.7030	0.650	0.88	0.000	0.00	0.000	0.0	0.076	10.753	3.03	51 60	0.31 6

TABLE 2 Linear Correlations among Design Variables and Measured Covariates

Correlation coefficient range	Main effects frequency (%)	Main by 2-way frequency (%)	2-Way interactions frequency (%)
0 to 0.25	76	76	75
0.25 to 0.50	18	19	17
0.5 to 0.75	2	2	4
0.75 to 1	4	3	4

Note that Main Effects Include Non-Linear Main Effects and Measured Covariates.

Even though we are comfortable with the correlation structure, it is still a daunting task to identify the model terms of importance among the 70 plus studied in 28 runs. This highlights the importance of using a model-averaging approach, such as BVA, in analyzing the data.

MODEL BUILDING: BAYESIAN VARIABLE ASSESSMENT

Modeling historical unstructured data and analyzing supersaturated designs have much in common. In particular, the predictor variables are not orthogonal. Further, the difference in fit among the best models may often be very small. The goals are also similar; that is, assess the importance that each predictor variable has on a response of interest. In such settings, selecting a single model ignores information from nearly equivalent competing models. This can result in misleading conclusions and inferior predictive models. We propose using an approach that has been quite effective in modeling historical unstructured data by averaging information across competing models; that is, BVA (Meyer and Wilkinson 1998).

In BVA, generic prior distributions are assigned to unknown parameters. In particular, each factor is assigned (the same) "activity probability" that reflects the prior belief that the variable (X_i) is in the true model.

$$\Pr(X_i \text{ is in the true model}) = \pi$$

Under an independence assumption, the prior probability that model M_k is the true model is given by

$$\Pr(M_k) = \pi^{a(k)}(1-\pi)^{p-a(k)}$$

where p is the total number of factors, and a(k) is the number of active factors in model M_k. Once we observe data (**y**), these prior model probabilities are updated using the full prior distribution (the details on the full prior distribution can be found in Meyer and Wilkinson [1995, 1998]) and the observed data. This results in posterior model probabilities:

$$\Pr(M_k|\mathbf{y}) = \text{constant} \times \Pr(\mathbf{y}|M_k)\Pr(M_k)$$

Typically, when there are many factors, there is no single model that stands out as having a posterior probability far higher than any other model. So selecting the one best model is not obvious. Instead, we compute the marginal probability that a factor is active by averaging across all models,

$$\Pr(X_i \text{ is in the true model}|\mathbf{y}) = \Sigma\, I(X_i \text{ is in model } M_k) \times \Pr(M_k|\mathbf{y})$$

where I is the indicator function, and the sum is taken over possible models. These posterior activity probabilities are used for variable selection. That is, variables with activity probabilities greater than a threshold are included in the model.

BVA computations are performed via Markov chain Monte Carlo (MCMC) approximation. For further mathematical details of this approach, see Meyer and Wilkinson (1995, 1998).

To illustrate the benefits of BVA, we consider the simulation study found in Beattie et al. (2002). This particular example is based on a design matrix from Lin (1993) and has been used by many authors for simulation purposes. The design matrix, provided in Appendix A, consists of 23 factors and 14 experimental runs (note that there is no factor X16). The response data are generated from the model

$$Y = E(Y) + \varepsilon$$

where the errors ε are distributed as i.i.d. standard normal. Three cases are examined:

- Case 1: $E(Y) = 10(X1)$
- Case 2: $E(Y) = -15(X1) + 12(X5) - 8(X9) + 6(X13) - 2(X17)$
- Case 3: $E(Y) = -15(X1) + 8(X5) - 6(X9) + 3(X5)(X9)$

For the simulation, the candidate set of models consists of all possible first-order models using the 23 different factors. Note that in Case 3, the true model contains an interaction term. This interaction is added to the true model to complicate the search for active factors; that is, because only first-order models are considered, the interaction term adds structured noise to the response. We ran BVA on 1,000 replicates for the simulation. The prior probability that a factor is important was set at 0.25, a standard value used at Lubrizol for many years in our data mining effort (e.g., see Meyer and Wilkinson 1998). A posterior activity probability above a certain threshold is an indication that the factor should be in the model. Two thresholds were considered; that is, 0.35 and 0.50. Though 0.50 is an obvious choice, in practice, we often use a value less than 0.50 as an indication that a factor is active. This protects against the possibility of missing active factors. In addition, a posterior probability of 0.35 is higher than our prior probability of 0.25, providing some evidence of an effect.

The BVA method was compared to a partial least squares variable selection method (PLSVS) proposed in Zhang et al. (2006), a penalized least squares method (SCAD) proposed by R. Li and Lin (2003), a three-stage stepwise regression approach (3STAGE) proposed in Lu and Wu (2004), and the Dantzig selector approach (DANTZIG) proposed in Phoa et al. (2009). Whenever the method selects the correct factors, and only those correct factors, then it is concluded that the true factor core has been identified. For example, for Case 3, this would be whenever X1, X5, and X9 are the only variables selected as active. Appendix B contains the results from the simulations. The table shows that the BVA technique performs well across a wide range of situations. As stated earlier, we believe that a key feature of the BVA approach is the concept of averaging over many models to determine which factors are important.

ANALYSIS OF THE FORMULATION EXPERIMENT

Though the simulation study in the previous section demonstrates the viability of the BVA algorithm, BVA should not be used as black-box software that is run once on the data set to yield the final model. In this actual experiment and analysis, as in any analysis, residuals must be checked to verify assumptions, and outliers and points of high influence must be investigated. Many iterations of the BVA analysis were performed on the COF data and results to determine the active variables from the experiment. Assumptions were checked, outliers and points of influence were investigated, and nonlinearity was assessed in all iterations of BVA. Though the analysis was not easy, it was helped by the effect sparsity, in which only a few of the variables appeared to have an influence on the test results.

As mentioned above, many iterations of the BVA analysis were performed on the COF data and results to determine the active variables from the experiment. In initial iterations, in order to respect hierarchy, first-order terms were forced into the model (activity probability set equal to 1) and BVA was used to determine the activity probabilities of second-order terms. In subsequent iterations, for convenience, first- and second-order terms were given the same prior distribution in the BVA analysis.

Initially, detergents T1 and T3 were estimated to have posterior activity probabilities greater that 50%, and detergent ash was estimated to have an activity probability of over 40%. The FMC by detergent substrate posterior activity probability was estimated at over 40% and the FMD by altered dispersant posterior activity probability was estimated to be over 80%. Of course, this initial iteration of BVA could not be considered as the final model without checking residuals, especially because the variables are correlated. The residuals revealed nonlinearity that was incorporated into the model. Using notation from Table 1, the final model consisted of the following terms: Altered, T3, Altered × FMD, $I(FMC \geq 2) \times FMC^2$ and $I(FMD \geq 2) \times FMD^2$. Here I(.) denotes the indicator variable. The effects of FMC and FMD are shown graphically in Figure 1.

Now the question is whether we truly identified the active variables from this design and analysis or have we overfit the model. To answer this question,

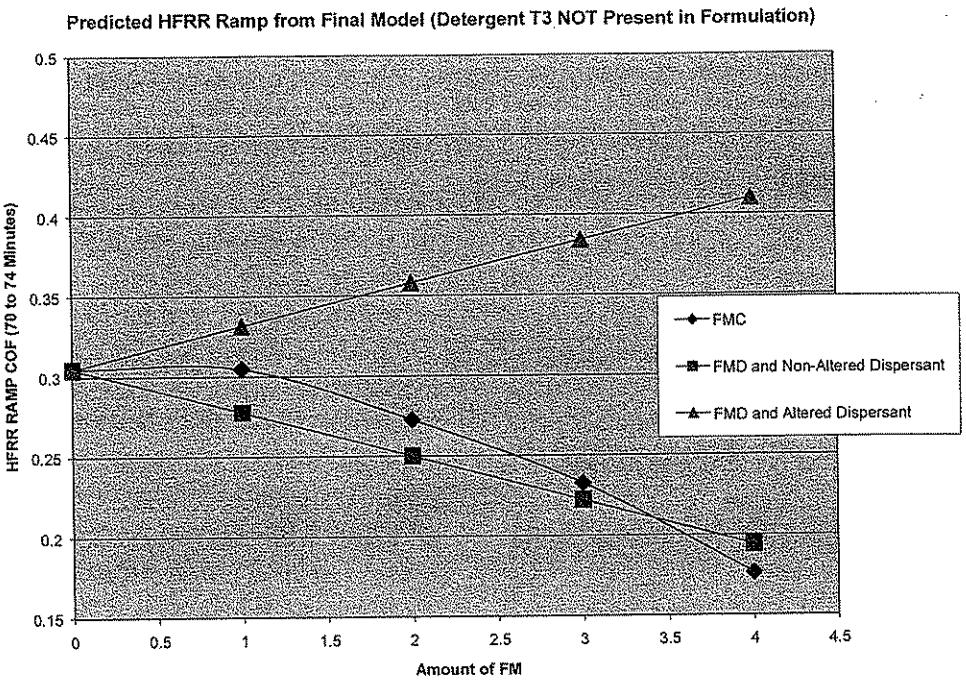

FIGURE 1 The COF response as a function of friction modifier.

we set up a 10-run design involving the variables identified as active, plus T1, T2, T4, and two different Z agents (Z1 and Z2) at a fixed level. All other variables from the first design were fixed. Note, however, that instead of altering the dispersant, the chemical element used to alter the dispersant was added to the formulation in a different form and tested as an indicator variable. This was done to test the hypothesis that the chemical element in another form would have the same directional effect identified in the first study. The follow-up design and the COF results from the follow-up design are included in Table 3.

Except for detergent T3, the analysis of the follow-up design confirmed all previously identified active variables as active. Detergents T1, T2, and T4 were not identified as active as in the original design and analysis, and there was no evidence that a change in the Z agent had an influence on the results. In addition, the model developed from the original design and analysis was used to predict the results from the follow-up design. In Figure 2, it can be seen that there is a correlation between the model predictions using the first design and the actual test results from the follow-up design. It may be noticed that the predictions, although correlated with the actual results, are not particularly close to the actual results. This may be for several reasons, including, but not limited to, a difference in the size of the effect of the

TABLE 3 Follow-up Design to the Supersaturated Design

FMC	FMD	Substrate	Ash	T1	T2	T3	T4	Z Agent	Interactive element	COF
0	4	0.250	0.661	1.25	0.0	0.00	0.0	Z1	YES	0.269
0	4	0.748	1.170	0.00	2.5	0.25	0.0	Z1	NO	0.218
0	4	0.248	0.679	1.00	0.0	0.25	0.0	Z1	YES	0.252
0	4	0.750	1.152	0.25	2.5	0.00	0.0	Z2	NO	0.233
0	4	0.804	0.600	0.80	0.0	0.00	1.0	Z1	YES	0.343
0	4	0.322	0.088	0.00	0.0	0.00	0.5	Z1	NO	0.198
0	0	0.248	0.679	1.00	0.0	0.25	0.0	Z1	NO	0.322
4	0	0.804	0.600	0.80	0.0	0.00	1.0	Z2	NO	0.224
4	4	0.250	0.661	1.25	0.0	0.00	0.0	Z1	NO	0.176
4	0	0.560	0.816	0.00	2.0	0.00	0.0	Z1	YES	0.198

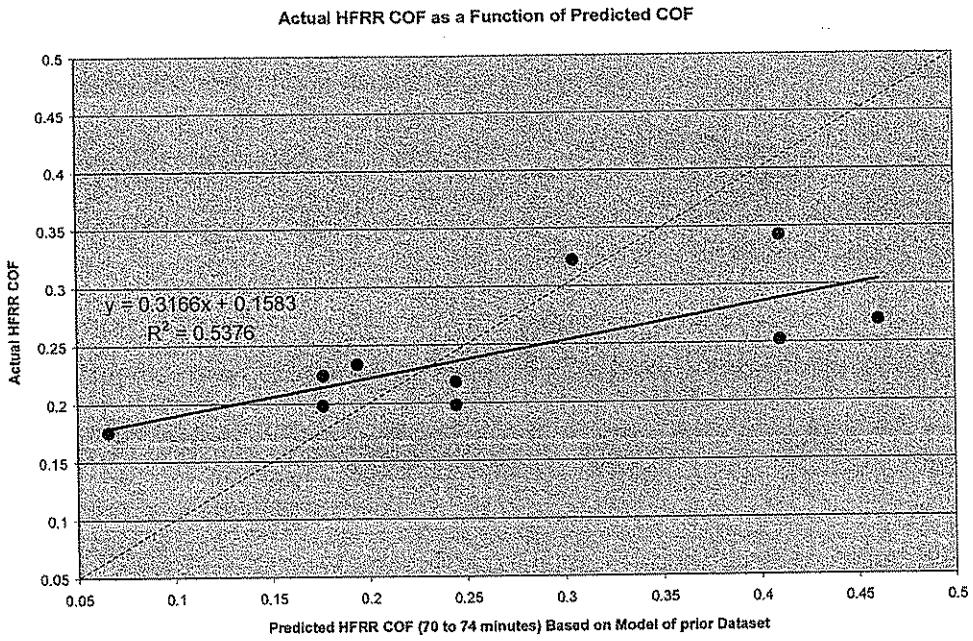

FIGURE 2 **Follow-up design to the supersaturated design.**

chemical element in another form and/or a change in the test parts, operators, etc., between the original and follow-up experiments.

The follow-up design results represent evidence that we were able to identify active variables from our first design and analysis with the supersaturated design and BVA analysis tools. Furthermore, the knowledge generated from the study was in time to incorporate into Lubrizol's oil formulation strategy for the next engine oil industry standard.

DISCUSSION AND CONCLUSIONS

Though our summary of a successful implementation of an actual supersaturated design and analysis does not prove that such designs should be used, it certainly does demonstrate that such designs may be used if coupled with an analysis technique such as model averaging. It is true that supersaturated designs would reduce the detection power of factors and variables common to both the supersaturated design and an unsaturated fractional-factorial of the same size, but it certainly improves the detection power of factors and variables squeezed into the supersaturated design and left out of the unsaturated fractional-factorial.

If our study were forced into an unsaturated 28-run design, we would have likely not tested either friction modifier C or D. We would have likely not tested six detergents, although we would have probably tested T3, and we would have not considered altering the dispersant. Such a design would have been much easier to analyze and the results much easier to accept; however, it would have fallen well short of meeting the need. Fortunately, the supersaturated design was timely, met the screening needs of Lubrizol formulators, and had a positive impact on Lubrizol's formulation strategy. In addition, this study has served as a demonstration and learning tool for supersaturated designs within Lubrizol. Since this successful trial, dozens of supersaturated designs have been launched and used in areas where cost-effective experimentation was thought to be impossible.

It should be noted and emphasized that the BVA analysis method was extremely important to the success of our particular supersaturated design and analysis. In typical designs such as factorials, fractional-factorials, and even Plackett-Burman designs, data analysis is straightforward and even somewhat uneventful. However, in cases where variables are not independent and the number of variables is larger than the number of design points, the analysis method as demonstrated earlier is critical. Supersaturated designs have the potential to be extremely powerful screening tools but only if they are analyzed with diligence and with successful analysis methods, such as BVA.

When trying to decide on the type of design to use for an application, it is extremely important to communicate and collaborate with your scientific partners, clients, and customers. Budget, theory, experience, subject matter, scientific need, product need, marketing need, and long- and short-term goals are all important items to discuss when constructing your design. Keep in mind that if the case were to arise where the number of study variables of interest exceeds the number of design points, instead of beginning design discussions by eliminating variables, it well may be worth the exercise to consider a supersaturated design.

ABOUT THE AUTHORS

Mr. Philip R. Scinto is a Technical Fellow for the Lubrizol Corporation where he has been employed since 1989. He holds a B.S., Cornell University, and M.S., Carnegie Mellon University, in Statistics. Some of his accomplishments in the engine oil industry include the worldwide control charting system for engine calibration known as the "Lubricant Test Monitoring System", the statistical treatment of product approval data, an experimental design approach to product read across, and "virtual"product testing.

Dr. Robert G. Wilkinson is the Manager of the Statistical Sciences Department at the Lubrizol Corporation. Besides managing a global group of statisticians, Dr. Wilkinson provides overall leadership for the trademarked Q.LIFE® Virtual Testing effort, and heads up internal method development efforts in Bayesian modeling, data mining and supersaturated designs. Dr. Wilkinson received his PhD in Statistics from Carnegie Mellon University.

Dr. Dennis Lin is a University Distinguished Professor of Supply Chain and Statistics at Penn State University. His research interests are quality engineering, industrial statistics, data mining and response surface. Dr. Lin is an elected fellow of ASA and ASQ, an elected member of ISI, a lifetime member of ICSA, and a fellow of RSS. He is an honorary chair professor for various universities, including a Chang-Jiang Scholar of China at Renmin University, National Chengchi University (Taiwan), Fudan University, and XiAn Statistical Institute (China).

ACKNOWLEDGMENT

We are grateful to the reviewers for their thorough and constructive comments, which have led to a significant improvement in the final version.

REFERENCES

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

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

Beattie, S. D., Fong, D. K. H., Lin, D. K. J. (2002). A two-stage Bayesian model selection strategy for supersaturated designs. *Technometrics*, 44:55–63.

Box, G. E. P., Meyer, R. D. (1986). An analysis for unreplicated fractional factorials. *Technometrics*, 4:489–495.

Butler, N. A. (2005). Minimax 16-run supersaturated designs. *Statistics & Probability Letters*, 73:139–145.

Georgiou, S. D. (2008). Modelling by supersaturated designs. *Computational Statistics and Data Analysis*, 54:428–435.

Holcomb, D. R., Montgomery, D. C., Carlyle, W. M. (2003). Analysis of supersaturated designs. *Journal of Quality Technology*, 35:13–27.

Kelly, H. W., Voelkel, J. O. (2000). Asymptotic-power problems in the analysis of supersaturated designs. *Statistics and Probability Letters*, 47:317–324.

Koukouvinos, C., Stylianou, S. (2005). A method for analyzing supersaturated designs. *Communications in Statistics – Simulation and Computation*, 34:929–937.

Li, R., Lin, D. K. J. (2003). Analysis methods for supersaturated design: Some comparisons. *Journal of Data Science*, 1(3):249–260.

Li, W., Wu, C. F. J. (1997). Columnwise-pairwise algorithms with applications to the constructions of supersaturated designs. *Technometrics*, 39:171–179.

Lin, D. K. J. (1993). A new class of supersatured designs. *Technometrics*, 35:28–31.

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

Liu, Y., Dean, A. (2004). *k*-Circulant supersaturated designs. *Technometrics*, 46:32–43.

Lu, X., Wu, X. (2004). A strategy of searching active factors in supersaturated screening experiments. *Journal of Quality Technology*, 36:392–399.

Meyer, R. D., Wilkinson, R. G. (1995). Variable selection or variable assessment? Center for Quality and Productivity Improvement, University of Wisconsin-Madison. Report #126.

Meyer, R. D., Wilkinson, R. G. (1998). Bayesian variable assessment. *Communications in Statististics - Theory and Methods*, 27:2675–2705.

Nguyen, N. K. (1996). An algorithm approach to constructing supersaturated designs. *Technometrics*, 38:69–73.

Phoa, F. K. H., Pan, Y., Xu, H. (2009). Analysis of supersaturated designs via the Dantzig selector. *Journal of Statistical Planning and Inference*, 139:2362–2372.

Satterthwaite, F. (1959). Random balance experimentation. *Technometrics*, 1:111–137 (with discussion).

Westfall, P. H., Young, S. S., Lin, D. K. J. (1998). Forward selection error control in the analysis of supersaturated designs. *Statistisca Sinica*, 8:101–117.

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

Yamada, S., Lin, D. K. J. (1997). Supersaturated design including an orthogonal base. *Canadian Journal of Statistics*, 25:203–213.

Zhang, Q. Z., Zhang, R. C., Liu, M. Q. (2006). A method for screening active factors in supersaturated designs. *Journal of Statistical Planning and Inference*, 137:2068–2079.

APPENDIX A: SIMULATION STUDY DESIGN MATRIX

Run	1	2	3	4	5	6	7	8	9	10	11	12	13	14
X1	1	1	1	1	−1	−1	−1	−1	−1	1	−1	1	1	−1
X2	1	−1	1	1	−1	−1	−1	1	−1	1	1	−1	1	−1
X3	1	−1	−1	−1	1	1	−1	1	−1	1	−1	−1	1	1
X4	−1	−1	1	1	1	1	−1	−1	−1	1	1	−1	1	−1
X5	−1	−1	1	−1	1	1	1	−1	−1	−1	1	1	1	−1
X6	−1	−1	−1	1	1	1	−1	1	1	1	−1	1	−1	−1
X7	1	1	−1	−1	−1	1	−1	−1	1	1	−1	1	1	−1
X8	1	1	−1	−1	1	−1	1	1	−1	1	1	−1	−1	−1
X9	1	1	−1	−1	1	1	−1	−1	−1	−1	1	1	1	−1
X10	1	−1	1	1	−1	1	1	1	−1	−1	−1	1	−1	−1
X11	1	−1	−1	1	−1	1	−1	−1	1	−1	1	1	−1	1
X12	−1	−1	1	−1	−1	−1	1	−1	1	1	−1	1	1	1
X13	1	1	1	1	1	−1	1	−1	−1	−1	−1	1	−1	−1
X14	−1	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1	1
X15	−1	1	1	1	1	1	−1	−1	1	1	−1	−1	−1	−1
X17	1	−1	1	−1	1	1	1	−1	1	1	−1	−1	−1	−1
X18	−1	1	−1	1	−1	1	1	−1	1	−1	1	−1	1	−1
X19	−1	−1	−1	1	−1	1	1	1	−1	1	1	1	−1	−1
X20	1	−1	−1	1	1	1	1	−1	−1	−1	−1	−1	1	1
X21	−1	1	−1	−1	−1	1	1	1	−1	1	−1	1	1	−1
X22	−1	1	1	−1	1	1	−1	1	−1	−1	−1	1	−1	1
X23	−1	−1	1	−1	1	−1	−1	1	1	−1	1	1	1	−1
X24	1	−1	−1	−1	1	−1	1	−1	1	1	1	1	−1	−1

APPENDIX B: RESULTS FROM THE SIMULATION

	Percentage of time true factor core identified			Percent of time active factors identified[a]		
Method	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
SCAD	76	70	N/A	100	99	N/A
PLSVS	61	74	N/A	100	95	N/A
3STAGE	53	53	42	100	100	100
DANTZIG	99	79	N/A	100	91	N/A
BVA (threshold = 0.35)	87	92	88	100	97	100
BVA (threshold = 0.50)	94	87	98	100	88	100

[a]Sometimes reported in the literature as smallest effect identified.