SANDIA REPORT
SAND99–8236
Unlimited Release
Printed April 1999

Small Sample Experimental Design Optimization and Repair

A. J. Johnson, C. DeCarli, S. V. Crowder

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

Sandia National Laboratories

When printing a copy of any digitized SAND Report, you are required to update the markings to current standards.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Prices available from (615) 576-8401, FTS 626-8401

Available to the public from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd
Springfield, VA 22161

NTIS price codes
Printed copy: A03
Microfiche copy: A01

SAND99-8236
Unlimited Release
Printed April 1999

Small Sample Experimental Design Optimization and Repair

A. J. Johnson, C. DeCarli
Sandia National Laboratories/California

S. V. Crowder
Sandia National Laboratories/New Mexico

Abstract

The primary goal of using Design of Experiment (DOEx) methods is to extract the maximum amount of information concerning experimental factors and their interactions from as few observations as possible. DOEx methodology allows an experimenter to selectively and systematically adjust process settings to learn which factors have the greatest impact on process and product performance. Using information about these factors, process settings can be adjusted until optimum performance is obtained. Classical designs can be used for most types of studies, yet additional methods for constructing designed experiments for non-standard situations are needed. In particular, Sandia National Laboratories' current interest is in further developing and applying special DOEx procedures for low-volume manufacturing process characterization and optimization. In this paper, we introduce and discuss the various "Design Optimality" criteria for constructing and "repairing" small-sample experiments. A case study involving the design, analysis, and repair of an ion source experiment is then presented, using the design of experiments software tool *Statistica*. Finally, we present a cost vs. benefit analysis of using the design optimality approach.

TABLE OF CONTENTS

Introduction 5
Basic Design of Experiment 5
Optimal-Design Concepts 6

DOEx Optimal Design Tool 7
D-optimality / D-Efficiency Design Criterion 7
A-optimality / A-Efficiency Design Criterion 8
G-optimality & Prediction Variance Design Criterion 9
Constructing an Optimal-Design Matrix 9
Application to Experiment Design 10
Application to Experiment Repair Example 10
Example of Optimal Experimental Design Selection 11

Case Study: Ion Source Experiment 15
Ion Source Experiment, as Planned 15
Repairing the Ion Source Experiment 16
Summary of Results 18

Resource Comparison 19

Conclusions 23

References 24

APPENDIX A–Design & Data Matrix 25
APPENDIX B–Data Comparison Tables 27

Introduction

Classical Design of Experiment[1] (DOEx) methods are available for most types of studies, yet special procedures for constructing and performing the analysis of experimental designs for non-routine situations are needed. In particular, Sandia National Laboratories' current interest is in further developing and applying special DOEx procedures for low volume manufacturing process characterization and optimization. Traditional DOEx often leads to experimental design plans with too many runs to be practical for small lot production. The goal of small sample DOEx is to provide experimental design plans for process characterization and startup with the fewest number of experimental runs that meets the specific goals and constraints of the experiment. This paper discusses and illustrates the usefulness of the "Design Optimality" criteria for constructing and "repairing" such small sample experiments. A case study involving the design, analysis, and repair of an ion source experiment is also presented as an example of this technique. A generic cost vs. benefit analysis is also presented.

Basic Design of Experiments

The primary goal of using DOEx methods is to extract the maximum amount of information concerning experimental factors and their interactions from as few observations as possible. Granting that various process settings allow for some range of adjustment, the DOEx methodology allows an engineer to selectively and systematically adjust these settings to learn which factors have the greatest impact on performance. Using this information, the settings can be constantly improved through sequential experimentation until the optimum process settings are obtained. Naturally, the more that is known about the experimental region, the easier it is to select appropriate process factors and settings for experimentation, and to identify the right model for estimating effects and experimental error[2]. In the initial stages of experimentation, DOEx provides relatively simple models that contain critical information, primarily identification of the

---

[1] DOEx concepts have been evolving since the 1920's when agricultural experimentation was used to develop the underlying modeling and analysis methods of designed experimentation.

most important factors. As the investigation progresses, more sophisticated DOEx methods can be applied (such as response surface methods) that are specifically designed for optimizing process settings. There are several references[3] that provide explicit procedures for constructing these experiments, as well as providing guidance with respect to the logistic and economical considerations (e.g. time, money, etc.) involved in experimental work. Included below is a brief summary of the critical ideas of DOEx that are also pertinent to the later discussion of optimal-design experiments.

The most desirable design characteristic is that all-important main effects and interactions can be estimated. When the design allows the separate contribution of each factor to be estimated independently, the design is said to be **orthogonal**. Another very important aspect of planning the experiment is to include as much **replication** as possible. Besides improving the estimate of the average response, replication provides an estimate of the experimental error by calculating the variation between the repeated runs. If the experimental error variation is significantly large, an accurate estimated model[4] may be difficult to identify.

Optimal-Design Concepts

Often uncontrollable circumstances present unique challenges to the experimenter. For example, an equipment malfunction or breakdown could result in missing experimental runs. Consequently, some effects of interest might no longer be estimable without redoing the missing runs of the experiment. The problem is then how to choose a minimum number of additional runs necessary to estimate all the effects of interest. Other difficult

---

[2] An estimate of the variation between samples tested at the same combination of factors.

[3] A widely recognized and excellent reference is Box, Hunter, and Hunter (1978) *Statistics for Experimenters*.

[4] Specifies which factors are included in the regression analysis; in general, the model may be stated as: $y = F(x_1, x_2, ..., x_n)$ where y, the dependent variable is a function of the independent variables or x's. A first order model includes linear terms, whereas higher order models are polynomial.

DOEx problems, such as constrained region[5] designs, have prompted much interest in developing special procedures to address these situations that classical designs cannot adequately handle.

DOEx Optimal Design Tool

With recent advances in design of experiment software, algorithms that construct "optimal" experimental designs[6] are available to address these problems.

These optimal-design algorithms can be used to:

1. Construct the "best" standard design through evaluating various efficiency measures and prediction variance measures.
2. Construct the "best" minimum run designs.
3. Augment or "repair" a current experimental design.
4. Replace points in a design because of an irregular or constrained experimental region.

Because the optimal-design algorithms provide the ability to generate non-standard designs, an investigator can use the "optimal-design" approach to generate a design matrix that limits the number of experimental runs. This ability provides a method to create a design requiring a small or minimum number of runs. Because these programs can calculate efficiency measures (A, D and G-efficiencies, often referred to as the "alphabetic" optimality criterion) and maximum prediction variance (MPV) measures, a quantitative comparison of one design with another can easily be made. The goal of the optimal-design approach is to choose a design with the highest efficiency or smallest MPV possible while still meeting the specific goals and constraints of the experiment.

D-Optimality / D-Efficiency Design Criterion

This criterion searches for a design matrix that maximizes the determinant[7] 'D' of the design/model matrix. The larger the determinant of a matrix, the more independent the

---

[5] The factors within the design region (or sub-region) are not feasible or constrained (such as in mixture problems where all ingredients must sum to 100%).

[6] Design optimality became practical (~1970-80) when the designs were formulated as efficient in terms of criteria, the results of work by Kiefer et al.

columns, thus, finding a design that maximizes the determinant of the design/model matrix means finding a design where the factor effects are maximally independent of each other. The D-efficiency measure can be interpreted as the relative number of runs (as a percent) that would be required by an orthogonal design to achieve the same value of the determinant $|X'X|$, where X is the design/model matrix that specifies what levels of the factors are to be run in the design and what terms are to be included in the model. Another advantage of using the D-optimality criterion to select a design is that it minimizes the overall variance of the estimated effects. Since an orthogonal design may not be possible or practical, this measure is a relative indicator of efficiency.

D-efficiency = $100 * (|X'X|^{1/p} / N)$,

Where: p = the number of factor effects to be estimated by the design (columns in the design/model matrix X), and
N = the number of requested runs.

A-Optimality / A-Efficiency Design Criterion

Another optimality criterion maximizes the diagonal elements of the X'X matrix while minimizing the off-diagonal elements. The effect of this criterion is to minimize the average variance of the estimated effects. The A-efficiency measure is the relative number of runs (as a percent) that would be required by an orthogonal design to achieve the same value of the trace[8] of $(X'X)^{-1}$. Again this measure is a relative indicator of efficiency that can be used to compare other designs of the same size and constructed from the same design points candidate list.

A-efficiency = $100 * p/\text{trace}\ (N*(X'X)^{-1})$

Where: p = the number of factor effects in the design,
N = the number of requested runs, and
trace = the sum of the diagonal elements (of $(N*(X'X)^{-1})$)

---

[7] The determinant D of a square matrix is a specific numerical value that reflects the amount of independence between the columns and rows of the matrix; the larger the determinant, the closer to orthogonality.

[8] The trace of a square matrix is defined as the sum of the elements on the principal diagonal.

G-Optimality & Prediction Variances Design Criterion

A third criterion, G-efficiency (global efficiency) or G-optimality emphasizes the use of designs for which the maximum prediction variance υ (**x**), in the region of the design is not too large. This measure is different from the D and/or A –optimal criterion because it does not calculate a single number, but is a function of the **x** (x_1, x_2, x_1x_2, x_1x_3, ... x_{q-1} x_q....) in the prediction. If the specified model is correct, and the value of Max υ (**x**) $\leq$ 1, then the model predictions are at least as precise as the measured responses. The prediction variance at a point **x**, is given by (standardized by dividing by σ^2):

$$\upsilon(\mathbf{x}) = \mathbf{x}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}'.$$

The G-efficiency or G-optimality measure is computed as:

G-efficiency = 100 * [square root (p/N)] ÷ Max υ (**x**),

Where: p = the number of factor effects in the design,
N = number of requested runs, and
Max υ (**x**) = the maximum prediction variance (MPV) across the list of candidate points.

The value in using the G-efficiency or max υ (**x**) as a design measure is that both provide useful information about the prediction of the response throughout the design space. A useful rule of thumb[9] is that a design matrix with max υ (**x**) $\leq$ 1.0 and/or G - efficiency $\geq$ 50% provides a reasonable design in many applications.

Constructing an "Optimal-Design" Matrix

To construct an "optimal-design" matrix using a computer software program requires the investigator to provide the following items:

1. a list of "candidate points" that specifies the valid or feasible design points (points in this list can be marked for forced inclusion in the final design),
2. the desired number of runs for the final experiment,
3. the specified exact form of the expected model, and
4. any other constraints.

---

[9] Reference: Snee, Ronald D., "Experimental Designs for Quadratic Models in Constrained Mixture Spaces". *Technometrics*, Vol. 17. No 2. May 1975.

Using this information as the basis, the optimal design algorithm searches for the "best design" that is orthogonal. Because an orthogonal design may not be possible, nearly orthogonal designs are generated. Other uncertainties or difficulties that the designer should be aware of include the following:

- Expect to have low efficiency measures if the design is "saturated[10]" or near saturated.
- There are usually no exact solutions. There can be several "local minima," therefore it is important to try a number of different initial designs and algorithms.
- Because the design intent is often to minimize the number of runs, an estimate of the experimental error[11] may not be available. Replication of some design points will be necessary to obtain an estimate of experimental error.
- If the specified model is not correct, then the generated design may not maximize the optimality criterion and may not be the best design.

Application to Experiment Design

The ability to quantitatively assess design efficiency is particularly valuable in creating small sample designs. Specifically in the low-volume manufacturing of very expensive parts, the number of experimental runs required in classical designs may not be possible. Other considerations such as production schedules provide additional incentives to use optimal-design methods with small sample experiments. Other design problems can also be accommodated (i.e., irregular or constrained design region). Such problems are commonly encountered in mixture problems, but other types of experiments can have factors where the upper or lower bounds of the design space must be constrained. (An example of such a situation is when the operating range of test equipment restricts the experimental region.)

Application to Experiment Repair

Since missing experimental runs can pose a real dilemma, the capability to "repair" a design becomes very important. The optimal-design algorithms allow the inclusion of

---

[10] It is possible to construct designs for investigating up to k factors = N-1 in only N runs; these designs are said to be saturated (where N is the number of experimental runs). Example: a 2^{3-1} design requires four experimental runs to examine 3 factors.

[11] There is no information to test for model inadequacy.

existing design points and use the remaining candidate points to complete the design matrix (within the specified number of runs). The investigator can select the design that will provide the appropriate data for analysis with a limited number of additional runs. The design efficiency measures can be used to quantify the value of additional runs or to identify which part of the design region of greatest interest is best examined by the different designs. This information helps the experimenter decide which additional run (or runs) should be added to "repair" the existing experiment.

Example of Optimal Experimental Design Selection

Included here is a simple example of how the optimal design criteria can be applied to generate an experimental design. This example will compare the optimal design measures (A, D, G, and Max Prediction Variance) for different run sizes. The final design matrix for this example will be selected using the recommended max prediction variance (MVP) criteria that satisfies max υ (x) $\leq$ 1.0.

For the example, a 4-factor (A, B, C, & D), 3-level (3^4), experiment is considered. The traditional full factorial design[12] requires an experimental matrix of 81 runs; this design matrix is entered into the optimal design software program as the "candidate list". Next, the factors to include in the regression model[13] must be determined. Selecting the model is the most challenging aspect of using the "optimal" design procedure since the experimenter is rarely certain of the model prior to designing any experiment. For this example, a linear and quadratic main effects with 2-way interactions model was specified[14]. To perform an

---

[12] A full factorial design includes all possible combinations of the factors at their respective level settings; a fractional factorial would require at least 27 experimental runs.

[13] Specifies which factors are included in the regression analysis; in general, the model may be stated as: $y = F(x_1, x_2, ..., x_n)$ where y, the dependent variable is a function of the independent variables or x's. A first order model includes linear terms, whereas higher order models are polynomial.

[14] Regression model includes linear main effects for the 4 factors (A, B, C, & D), squared factor effects (A^2, B^2, C^2, & D^2) and two way interactions (A*B, A*C, A*D, B*C, B*D, & C*D)

analysis, the lower bound on the number of experimental runs is 15 (selected from the candidate list of 81 experimental runs). However, as shown in the following table and graph, to satisfy the selection criteria for the prediction variance (max υ (x) $\leq$ 1.0)[15], the final design must include at least 24 runs. The value for the maximum prediction variance continues to improve (decrease) incrementally as more runs are included in the design. However, as the number of runs in the design exceeds the number required to achieve an orthogonal design, the improvement in the D, A and G- efficiencies measures do not improve (in this example at approximately 60). In comparing the improvement in the amount of the prediction variance within the first 30 experimental runs with the improvement over the next 20 runs, the point of "diminishing returns" can be evaluated. Assuming that there are no other limitations (i.e., economic considerations) and the experimental objectives can be met, then the design matrix associated with the 24 runs would provide the "optimal design matrix" for the 4-factor, 3-level experiment. Using such a graph, the size of an experiment can be appropriated (i.e., the minimum number of runs for which max υ (x) $\leq$ 1.0 can be obtained vs. the point of "diminishing returns"). While this example uses the value of the prediction variance, other considerations such as cost per experiment, time to conduct each experiment, or some component performance measure could be used as the determining factor.

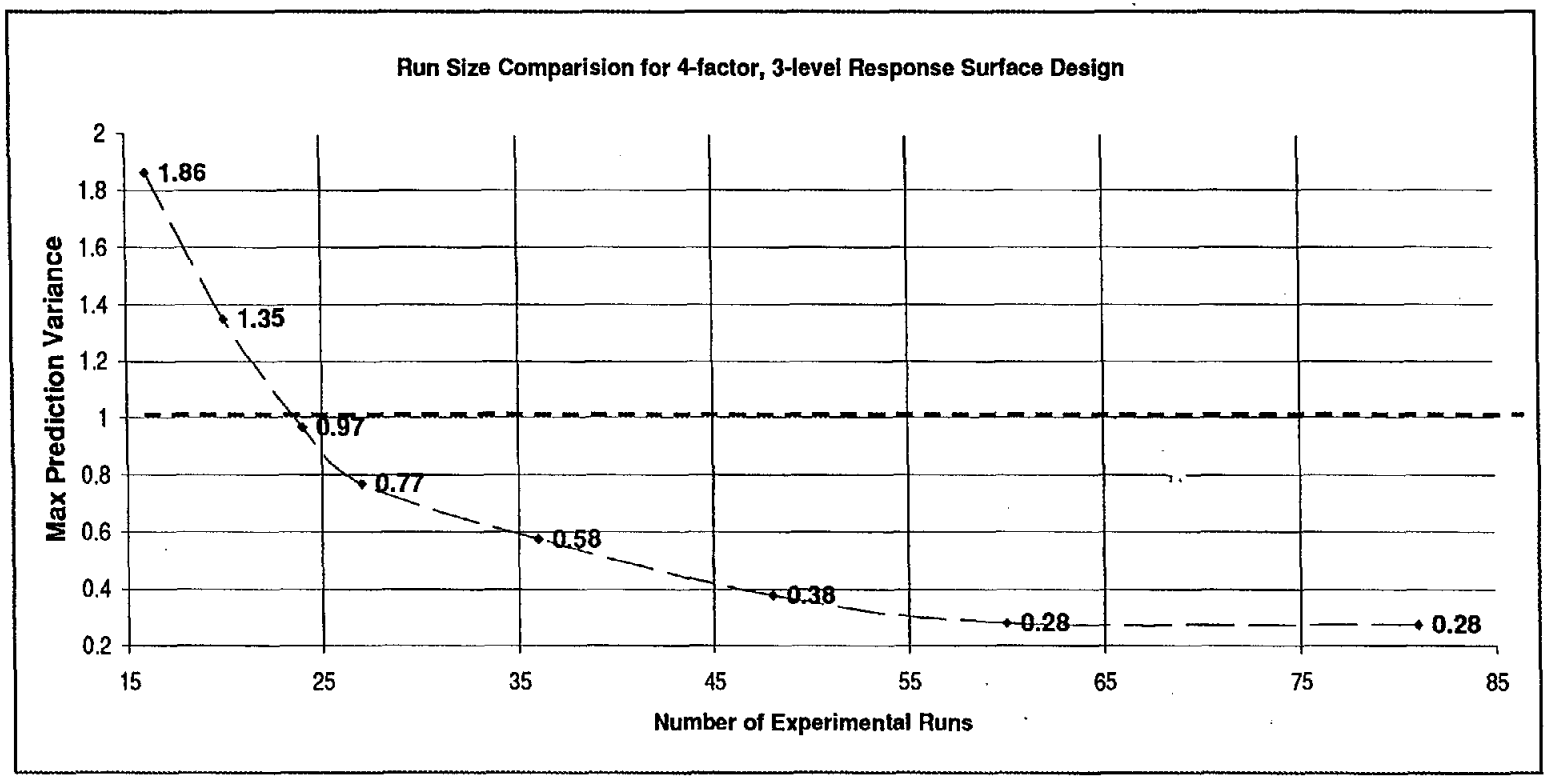
---

[15] As previously stated a design matrix with max υ (x) $\leq$ 1.0 and/or G - efficiency $\geq$ 50% provides a reasonable design. Reference: Snee, Ronald D., "Experimental Designs for Quadratic Models in Constrained Mixture Spaces". *Technometrics*, Vol. 17. No 2. May 1975.

Table Summary of Efficiency Measures 4-Factor, 3-Level Linear & Quadratic Main Effects with 2-way Interactions Model[16]

Measure	16 Runs	20 Runs	24 Runs	27 Runs	36 Runs	48 Runs	60 Runs	81 Runs
D-Efficiency	43.02	46.28	46.32	47.67	48.15	48.44	48.60	43.45
A-Efficiency	17.85	24.14	23.99	25.50	26.29	25.23	26.47	32.26
G- Efficiency	70.94	74.53	80.32	84.86	85.16	90.92	94.23	81.65
Max Prediction Variance	1.86	1.35	0.97	0.77	0.58	0.38	0.28	0.28

[16] Regression model includes linear main effects for the 4 factors (A, B, C, & D), squared factor effects (A^2, B^2, C^2, & D^2) and two way interactions (A*B, A*C, A*D, B*C, B*D, & C*D)



NOTE: The "optimal" number of runs to include in the design matrix is based on the selection criteria for the prediction variance (max υ (x)) $\leq$ 1.0.

Case Study: Ion Source Experiment

The objective of the Ion Source Experiment was to identify whether certain processing conditions had an impact on the total number of shots to failure (shot life) of the neutron generator tubes. Four factors were included in this investigation:

A. Chromium Thickness (1000 or 3500 A), 2 levels
B. Chromium Deposition Rate (10 or 200 A/s), 2 levels
C. Cleaning Agent (Brulin or TCE), 2 levels and
D. Getter Type (Getter, Throttled Getter, or No Getter), 3 levels.

If any of these factors could be shown to impact shot life performance, those results would be used to optimize the process and maximize shot life. Other design and processing factors (Substrate surface finish (15 micron), Scandium temperature, Scandium deposition rate (200 A/s), Oxygen Passivation (None), and Exhaust profile (Standard)) were held constant throughout the experiment. To be able to readily quantify the impact on the product in terms of producing a longer shot life, a standard full factorial mixed-level design[17] was desired. This design (the full factorial matrix is given in Appendix A, Table 1) allows an estimate of all important factors and interactions between factors to be identified. Also, anticipating some sample-to-sample variation, it was recommended to obtain at least one replication for each combination of factor levels included in the experiment. This improves the estimate of the average response at each combination and provides an estimate of the experimental error.

For various reasons the experiment as planned could not be completed, and the data available for analysis was very limited. In summary, only 29 of the originally planned 48 tests were performed. Of these 29 tests, only 7 provided useful shot life data (the data matrix is provided in the Appendix, Table 2). The other 22 shot life tests were truncated at approximately 20 shots, resulting in data that could not be used in the analysis. The situation raised the following questions:

---

[17] A mixed-level design includes factors with unequal levels. (i.e., for a 3-factor experiment, 2 of the factors have only 2-levels, high and low, while the other factor has 3-levels, high, medium, and low.)

- By using the DOEx tools for optimal designs, could we the "repair" the experiment? (i.e., is there an optimum number and choice of additional runs that could be performed to allow a meaningful analysis?)
- If only a single additional run could be conducted (due to material, time, or other resource constraints), which run would provide the greatest amount of information?

Ion Source Experiment Design Repair

The data set used for this investigation included only the 7 runs from the Ion Source Experiment that had useful shot life test results. The additional runs necessary to complete the full factorial design matrix for a mixed-level 3-factor[18] design were provided as the other candidate design points. (This design/data matrix is provided in Appendix A, Table 2.) The objective of our investigation was to answer the questions, which were raised as a result of the missing experimental runs. Using the "optimal-design[19]" software, several potential "repair" strategies were investigated by varying the following factors:

1. Number of runs in repaired design (e.g., 7 vs 8, vs 9, etc...),
2. Using different criterion to evaluate the designs (A-optimality vs. D-optimality vs. Maximum Prediction Variance (MPV) criterion),
3. Forcing the inclusion of design points (6 points vs. 8 points), and
4. Specifying different models to fit in the analysis (e.g., main effects vs. inclusion of 2-way interactions).

A summary of these comparisons is provided here and a comprehensive set of comparison tables is included in the appendix for further review.

As expected, increasing the number of runs had a positive impact on the A and D-efficiency measures as well as the Maximum Prediction Variance criterion. This result was consistently seen, regardless of the complexity of the model specified. Throughout the investigation, there was no significant difference in the final recommended design when

---

[18] A decision was made to eliminate the type of cleaner from the initial design as one of the factors of interest in the repair effort, since it was believed to have no effect on the response variable, shot life.

[19] The computer software package used was *Statistica's Industrial System Design and Analysis of Experiments*[19] module, from Statsoft, Inc.

comparing the D-optimality and A-optimality criteria. For most of the trials, the points that were marked for forced inclusion in the final design were the six unique points from the existing shot life data set. However, for the sake of this investigation, we performed a few comparisons with 8 points forced into the final design. The efficiency measures for the designs with 8 forced points were consistently lower (D-efficiency measures for these comparison runs are provided in Appendix B.).

As seen in the results of the Ion Source Experiment, the model specified in the investigation can make a considerable difference in the efficiency and maximum prediction variance measures. And as one would expect, as the complexity of the model increases, the efficiency measures decrease. An example that demonstrates the significant difference in both the efficiency and prediction variance measures is in the comparison of the models that include linear, quadratic and interaction effects with the model that includes only main effects. The first difference is that the more complex model cannot be estimated completely with only 7 runs, since more than 7 terms must be estimated. The next comparison, between the two models of 9 runs, shows that the efficiency is considerably better (as expected) in the main effects model. It is of interest, however, that the 9 run main effects design is still better than the 12 or 13 run higher order model. This shows the impact that model selection has on the efficiency and maximum prediction variance measures.

D-Optimal Model includes linear, quadratic and interaction effects:

Number Runs	7 RUNS[20]	9 RUNS	12 RUNS	13 RUNS
Selected Runs (added runs in bold)	--	1, 2, 3, 4, 6, 7, **8, 11, 13**	1,2, 3, 4, 6, 7, **8, 9, 10, 11, 12, 13**	all
D-efficiency (%)	--	64.63	71.17	69.11
A-efficiency (%)	--	31.47	53.33	53.38
Max pred. var.	--	3.8	.75	.745

[20] No measures could be calculated for 7 runs because requested interactions in the model.

D-Optimal Model includes main effects only:

Measure	7 RUNS	9 RUNS	12 RUNS	13 RUNS
Selected Runs (added runs in bold)	1, 2, 3, 4, 6, 7, **13**	1, 2, 3, 4, 6, 7, **8, 11, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13, 6, 11**	all
D-efficiency (%)	85.45	91.86	94.74	88.19
A-efficiency (%)	79.50	89.19	93.55	85.87
Max pred. var.	.91	.64	.36	.37

Another means of comparing the different trials is shown in the table below. The addition of just a single run (from 9 to 10) provides a considerable improvement to the maximum prediction variance (2.43 to 0.875).

D-Optimal model includes Linear Effects/Interacts A*B, A*C, B*C:

Number Runs	7 RUNS	9 RUNS	10 RUNS	12 RUNS	13 RUNS
Selected Runs (added runs in bold)	--	1, 2, 3, 4, 6, 7, **8, 11, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13, 13, 11**	all
D-efficiency (%)	--	83.70	89.83	89.53	80.84
A-efficiency (%)	--	74.24	88.42	86.34	78.32
Max pred. var.	--	2.43	.875	.86	.71

Results Summary

As previously stated, the objective of this effort was to address the following questions:

- By using the DOEx tools for small designs, could we the repair the experiment? (i.e., Is there an optimum number and choice of additional runs that could be performed to allow a meaningful analysis?)
- If only a single run could be conducted, (added to the existing runs) which run would provide the greatest amount of information?

Given the results of this investigation the answers are:

1. Depending on the specified model, the "optimal-design" tools could readily repair the experiment in as few as 7 runs for the main effects only model. If an estimate of all of the 2-way interaction terms is desired, then a 10 run design is required. In each case, the recommended number of runs to repair the design reduces the maximum prediction variance to below one, the recommended value of an adequate design.
2. Concerning the question of adding a **single** run, the main effects model design with just 7 runs showed a remarkable improvement in the efficiency and maximum prediction variance measures:

Measure	6 RUNS	7 RUNS
D-efficiency (%)	65.6	85. 5
A-efficiency (%)	40.	79.5
Max pred. var.	4.3	.91

For the model including the 2-way interaction terms, adding a single run would not have a meaningful benefit using the max υ (**x**)) $\leq$ 1.0 criterion suggested. For this model the smallest feasible design (i.e., the one with the fewest number of runs) requires at least 3 additional runs. However, in reviewing the efficiency and maximum prediction variance measures, an additional 4 runs (for a total of 10 runs) provides a much better design.

Number Runs	7 RUNS	9 RUNS	10 RUNS
D-efficiency (%)	--	83.7	89.8
A-efficiency (%)	--	74.2	88.4
Max pred. var.	--	2.4	.875

Resources Comparison

One of the benefits realized in using the DOEx approach is that statistical experiment design plans assure that the experimental work is done efficiently and the limitations of the traditional one-at-a-time scientific method are avoided. This benefit is further exploited with the ability to quantitatively evaluate different designs and minimize the number of runs using optimal design measures.

To illustrate this aspect of the optimal design approach, an example has been created. Using a design of experiment software program, three standard Central Composite Designs (CCDs) were generated. By specifying the number of factors to include in the experimental design, the program generated the following design parameters and number of recommended experimental runs:

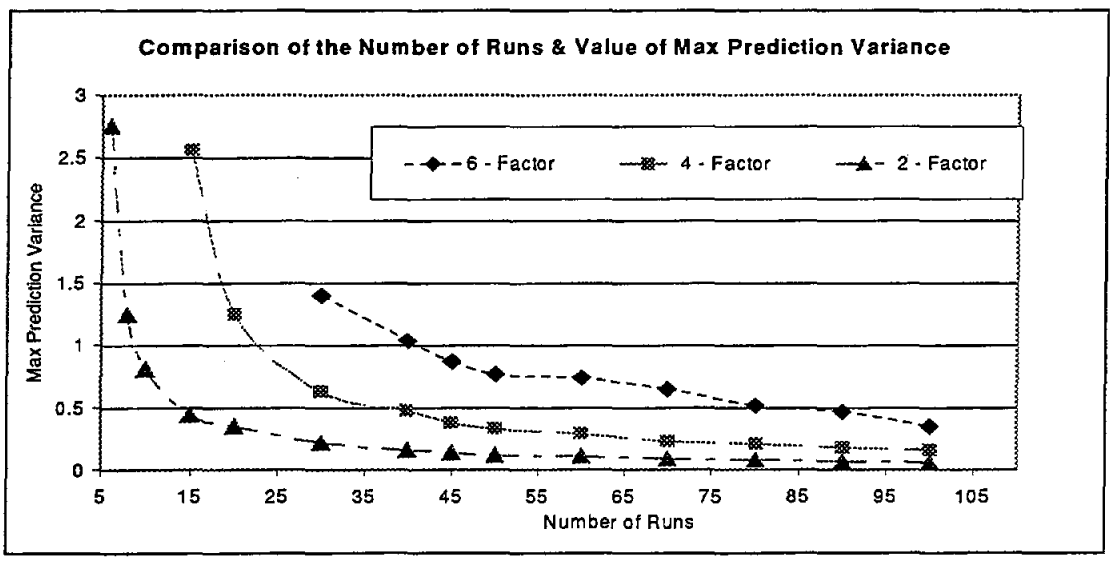
2 Factors	Center points: 5	Runs: 13
4 Factors	Center points: 7	Runs: 31
6 Factors	Center points: 14	Runs: 90

Similar to the example presented earlier, each of the three design matrices was provided as candidate lists in the optimal design software program. The analysis model used for all three cases included linear and quadratic main effects with 2-way interactions. The following table is provided to show the optimal design results of the minimum number of runs that could be performed while satisfying the MPV criterion max υ (x) $\leq$ 1.0.

Comparison of Traditional DOEx & Optimal Design

Factors in Experiment	Recommended Number of Runs Using Traditional DOEx CCD	Recommended Number of Runs Using Optimal Design Criteria
2	13	9
4	31	24
6	90	45

A graphical representation of the maximum prediction variance vs the number of experimental runs is shown in the following chart. This type of graph provides an easy way to appropriately size an experiment. It can be used to determine the number of runs required to satisfy the MVP$\leq$ 1.0. It can also be used to identify the point of "diminishing returns," that point at which additional experimental runs provide very little additional information.



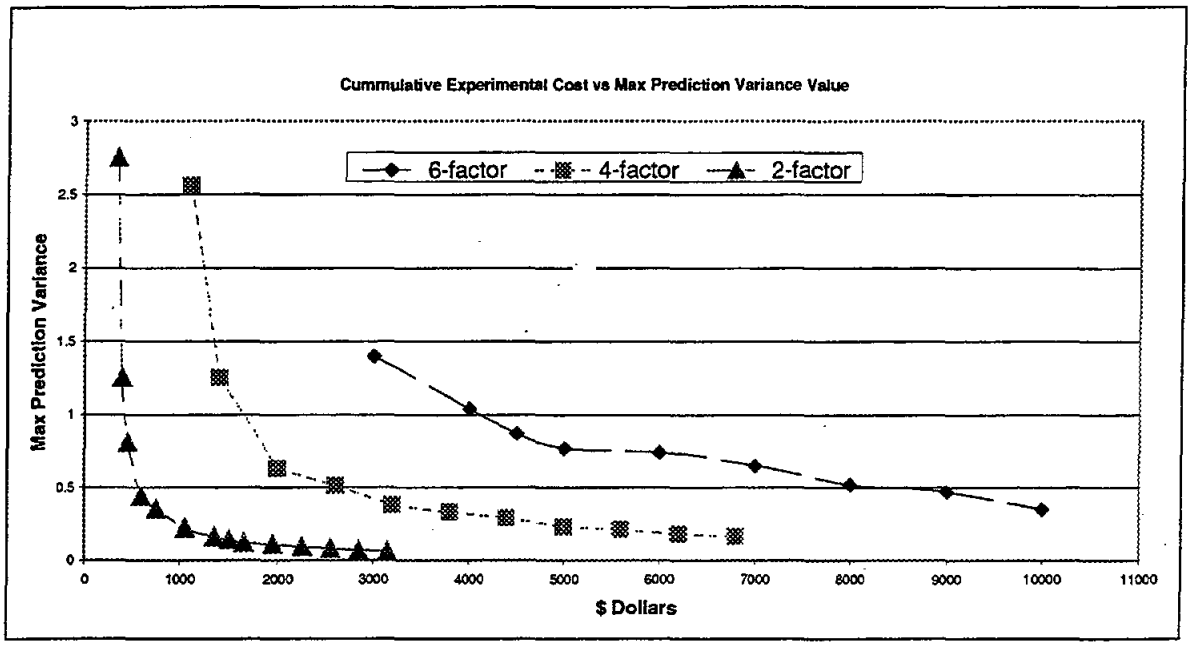
Using the graph and the next table, the value of the maximum prediction variance for the traditional CCD design and the optimal design can be compared. In the 2-factor case the maximum prediction variance is less than 1 after 9 runs (0.81) and the 13 run has a

maximum prediction variance of 0.79. From this analysis we can readily conclude that if four additional runs are conducted, the maximum prediction variance will improve by only 0.02. The investigator can then determine if this amount of improvement is necessary and or of sufficient value to perform the additional 4 experimental runs.

Number of Experimental Runs	6-Factor	4-Factor	2-Factor
6	--	--	2.75
8	--	--	1.25
9	--	--	0.81
13	--	--	0.79
15	--	2.56	0.44
20	--	1.25	0.35
30	1.4	0.63	0.22
40	1.04	0.48	0.16
45	0.87	0.38	0.14
50	0.77	0.33	0.12
60	0.74	0.29	0.11
70	0.65	0.23	0.09
80	0.52	0.21	0.08
90	0.47	0.18	0.07
100	0.35	0.16	0.06

Given a situation where the number of factors or the number of experimental runs are limited due to budget constraints, the optimal design measurements may be of further use in evaluating the trade-offs, or as a method to justify additional experimental budget requests. (For example, to justify cost associated with performing the four additional runs in the Ion Source experiment, a graphical representation of the maximum prediction variance vs cost could provide meaningful support to a budget request.) The next example illustrates using the optimal design measures in a systematic method to evaluate the trade-offs and determine the number of runs when cost is the underlining constraint. For this example, hypothetical costs per run for each of the cases (6-factor, 4-factor, and 2-factor

experiments) were assumed[21] and the plot shows the cumulative cost versus the maximum prediction variance value.



For the 4-factor experiment, approximately $2000 would be necessary to conduct the number of experimental runs that satisfy the criteria MVP≤ 1.0. In the 6-factor case, approximately $4500 would be needed to conduct the desired number of runs to meet the criteria MVP≤ 1.0. This information can be used to establish the overall DOEx budget; in keeping with general guidelines approximately 25% of the available experimental budget should be allotted to the initial screening phase, 50-60% allocated to the model building phase and the remaining 15-25% spent on verification. Therefore, given that this effort is part of the model building phase, the budgeted amount would represent 50-60% of the overall DOEx budget. While it was mentioned earlier, other considerations such as time to conduct each experiment or some component performance measure could be used as the factor in the analysis.

---

[21] Hypothetical Cost Assumptions:
Cumulative Cost for 6-Factor: base cost $300, + $100/run
Cumulative Cost for 4-Factor base cost $200, plus $60/run
Cumulative Cost for 2-Factor base cost $150, plus $30/run

Conclusions

Although the classic experimental design approach provides a reliable means of generating and analyzing experiments, the ability to generate "optimal designs" remains very important in the field of DOEx. Because there will often be experiments where the sample size is limited (as in small sample process characterization), the region of experimentation is unusual, or (as in the Ion Source Experiment investigation) the need to repair the design exists, an optimal design tool must be provided to support the design of experiments. In this paper, we introduced and discussed various "Design Optimality" criteria for constructing and "repairing" small sample experiments. A case study involving the design, analysis, and repair of an Ion Source Experiment was then presented, using the design of experiments software tool *Statistica*. The case study demonstrated how to use the "optimal-design" approach to generate small yet efficient designs. Finally, we presented how a cost vs. benefit analysis could be performed using the design optimality approach.

References

Box, G. E. P., Hunter J. S. , and Hunter, W. G. (1978) Statistics for Experimenters. John Wiley & Sons, New York.

Box, G. E. P., Draper, N. R. (1987) Empirical Model-Building & Response Surfaces. John Wiley & Sons, New York.

Minitab Statistical Software, Minitab Inc. (1995)

Mitchell, T. J. (1974). "An Algorithm for the Construction of D-Optimal Experimental Designs." *Technometrics*, Vol. 16, No. 2, pp. 203- 210.

Mitchell, T. J. (1974). "Computer Construction of D-Optimal First Order Designs." *Technometrics*, Vol. 16, No. 2, pp. 211- 220.

Myers, Raymond H., and Montgomery, Douglas C. (1995). Response Surface Methodology, Process and Product Optimization Using Design Experiments. John Wiley & Sons, New York.

Nachtsheim, C. J. (1987). "Tools for Computer-Aided Design of Experiments." *Journal of Quality Technology*. Vol. 19, No 3, pp. 132-160.

Snee, Ronald D. (1975), "Experimental Designs for Quadratic Models in Constrained Mixture Spaces". *Technometrics,* Vol. 17, No 2, pp. 149-159.

Statistica Industrial System Design and Analysis of Experiment. Statsoft, Inc. (1997)

APPENDIX A.

The four factors identified for inclusion in the experiment:

A. Chromium Thickness (1000 or 3500 A), 2 levels

B. Chromium Deposition Rate (10 or 200 A/s), 2 levels

C. Cleaning Agent (Brulin or TCE), 2 levels and

D. Getter Type (Getter, Throttled Getter, or No Getter) 3 levels.

Table 1. Experimental Design Matrix for Group 3 Ion Source Experiment

Run Number	Cr Thickness	Cr Dep Rate	Cleaning Agent	Getter Type
1	1000	10	Brulin	Getter
2	1000	10	Brulin	Getter
3	1000	10	Brulin	No Getter
4	1000	10	Brulin	No Getter
5	1000	10	Brulin	Throttled Getter
6	1000	10	Brulin	Throttled Getter
7	1000	10	TCE	Getter
8	1000	10	TCE	Getter
9	1000	10	TCE	No Getter
10	1000	10	TCE	No Getter
11	1000	10	TCE	Throttled Getter
12	1000	10	TCE	Throttled Getter
13	1000	200	Brulin	Getter
14	1000	200	Brulin	Getter
15	1000	200	Brulin	No Getter
16	1000	200	Brulin	No Getter
17	1000	200	Brulin	Throttled Getter
18	1000	200	Brulin	Throttled Getter
19	1000	200	TCE	Getter
20	1000	200	TCE	Getter
21	1000	200	TCE	No Getter
22	1000	200	TCE	No Getter
23	1000	200	TCE	Throttled Getter
24	1000	200	TCE	Throttled Getter
25	3500	10	Brulin	Getter
26	3500	10	Brulin	Getter
27	3500	10	Brulin	No Getter
28	3500	10	Brulin	No Getter
29	3500	10	Brulin	Throttled Getter
30	3500	10	Brulin	Throttled Getter
31	3500	10	TCE	Getter
32	3500	10	TCE	Getter
33	3500	10	TCE	No Getter
34	3500	10	TCE	No Getter
35	3500	10	TCE	Throttled Getter

36	3500	10	TCE	Throttled Getter
37	3500	200	Brulin	Getter
38	3500	200	Brulin	Getter
39	3500	200	Brulin	No Getter
40	3500	200	Brulin	No Getter
41	3500	200	Brulin	Throttled Getter
42	3500	200	Brulin	Throttled Getter
43	3500	200	TCE	Getter
44	3500	200	TCE	Getter
45	3500	200	TCE	No Getter
46	3500	200	TCE	No Getter
47	3500	200	TCE	Throttled Getter
48	3500	200	TCE	Throttled Getter

In the design/data matrix that follows, the levels for each of the factors are also shown designated as -1, 0, and 1.

Table 2. Group 3 Ion Source Design/Data Matrix[22] Using 7 Life Test Runs

List #	Run #	Cr Thk	Cr Thk	Cr Rate	Cr Rate	Getter	Getter	Shot Life
1	31	1000	-1	200	1	Getter	-1	47
2	38	3500	1	10	-1	Getter	-1	42
3	33	1000	-1	200	1	NG	1	87
4	39	3500	1	10	-1	NG	1	85
5[23]	40	3500	1	10	-1	NG	1	86
6	5	1000	-1	10	-1	TG	0	73
7	11	1000	-1	200	1	TG	0	65
8		1000	-1	10	-1	Getter	-1	
9		1000	-1	10	-1	NG	1	
10		3500	1	200	1	NG	1	
11		3500	1	200	1	Getter	-1	
12		3500	1	10	-1	TG	0	
13		3500	1	200	1	TG	0	

[22] Factor C, Cleaning Agent (Brulin or TCE) was not included in the effort to repair the design.

[23] Run 4 and 5 are replicate runs.

APPENDIX B.

Comparison between the number of runs & using the D vs A-Criterion for main effects model.

Requested STATISTICA Program To Select 7 Run D-Optimal Design

Measure	7 RUNS	13 RUNS
Selected Runs	**1, 2, 6, 13, 1, 6, 13**	all
D-efficiency (%)	96.10	88.19
A-efficiency (%)	91.43	85.87
Max pred. var.	1.	.369

D-Optimal Design with 6 Forced Points

Measure	Forced Pts	7 RUNS	9 RUNS	12 RUNS	13 RUNS
Selected Runs (added runs in bold)	1, 2, 3, 4, 6, 7	1, 2, 3, 4, 6, 7, **13**	1, 2, 3, 4, 6, 7, **8, 11, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13, 6, 11**	all
D-efficiency (%)	65.60	85.45	91.86	94.74	88.19
A-efficiency (%)	40.	79.50	89.19	93.55	85.87
Max pred. var.	4.33	.91	.64	.391	.369

A-Optimal Design with 6 Forced Points

Number Runs	Forced Pts	7 RUNS	9 RUNS	12 RUNS	13 RUNS
Selected Runs (added runs in bold)	1, 2, 3, 4, 6, 7	1, 2, 3, 4, 6, 7, **11**	1, 2, 3, 4, 6, 7, **8, 11, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13, 6, 13**	all
D-efficiency (%)	65.60	85.45	91.86	94.74	88.19
A-efficiency (%)	40.	79.50	89.19	93.55	85.87
Max pred. var.	4.33	.91	.64	.391	.369

Comparison of Linear/Quad Effects & Inter. C*C A*B A*C B*C model

D-Optimal Design with 6 Forced Points

Number Runs	Forced Pts	7 RUNS[24]	9 RUNS	10 RUNS	12 RUNS	13 RUNS
Selected Runs (added runs in bold)	1, 2, 3, 4, 6, 7	--	1, 2, 3, 4, 6, 7, **8, 11, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13**	1, 2, 3, 4, 6, 7, **8, 9, 10, 11, 12, 13**	all
D-efficiency (%)	--	--	64.63	70.77	71.17	69.11
A-efficiency (%)	--	--	31.47	38.4	53.33	53.38
Max pred. var.	--	--	3.8	1. 25	.75	.745

Comparison of number of runs & using D vs A-Criterion for Linear Effects/Interacts A*B, A*C, B*C model.

D-Optimal Design with 6 Forced Points

Number Runs	Forced Pts	7 RUNS	9 RUNS	10 RUNS	12 RUNS	13 RUNS
Selected Runs (added runs in bold)	1, 2, 3, 4, 6, 7	--	1, 2, 3, 4, 6, 7, **8, 11, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13, 13, 11**	all
D-efficiency (%)	--	--	83.70	89.83	89.53	80.84
A-efficiency (%)	--	--	74.24	88.42	86.34	78.32
Max pred. var.	--	--	2.43	.875	.857	.708

A-Optimal Design with 6 Forced Points

Number Runs	Forced Pts	7 RUNS	9 RUNS	10 RUNS	12 RUNS	13 RUNS
Selected Runs (added runs in bold)	1, 2, 3, 4, 6, 7	--	1, 2, 3, 4, 6, 7, **8, 10, 12**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13, 6, 8**	all
D-efficiency (%)	--	--	74.16	89.83	89.53	80.84
A-efficiency (%)	--	--	60.22	88.42	86.34	78.32
Max pred. var.	--	--	2.5	.875	.8571429	.7083333

[24] No measures could be calculated for 7 runs because of the requested interactions in the model.

Comparison of forcing different points into D-Optimal Design using Linear Effects/Interacts (A*B, A*C, B*C)

Forced Points	Number Runs	8 RUNS	9 RUNS	10 RUNS	12 RUNS
8 Forced Points	D-efficiency (%)	52.98%	61.53%	70.72%	84.05%
6 Forced Points	D-efficiency (%)	64.59%	83.70%	89.83%	89.53%

Comparison between the number of runs & inclusion of different Interaction terms

D-Optimal Design With 6 Forced Points Model Incl.: Linear Effects/Interacts A*C, B*C

Number Runs	Forced Pts	9 RUNS	10 RUNS	12 RUNS	13 RUNS
Selected Runs (added runs in bold)	1, 2, 3, 4, 6, 7	1, 2, 3, 4, 6, 7, **8, 11, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13, 6, 11**	all
D-efficiency (%)	--	83.19	88.84	88.78	78.22
A-efficiency (%)	--	75.69	87.81	86.02	75.87
Max pred. var.	--	1.79	0.725	.681	.619

D-Optimal Design With 6 Forced Points Model Incl.: Linear Effects/Interacts A*B

Number Runs	Forced Pts	9 RUNS	10 RUNS	12 RUNS	13 RUNS
Selected Runs (added runs in bold)	1, 2, 3, 4, 6, 7	1, 2, 3, 4, 6, 7, **13, 11, 8**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13, 8, 11**	all
D-efficiency (%)	--	92.48	94.09	95.10	90.16
A-efficiency (%)	--	89.35	92.31	93.33	87.91
Max pred. var.	--	.846	.625	.524	.458

D-Optimal Design With 6 Forced Points Model Incl.: Linear Effects/Interacts A*C

Number Runs	Forced Pts	9 RUNS	10 RUNS	12 RUNS	13 RUNS
Selected Runs (added runs in bold)	1, 2, 3, 4, 6, 7	1, 2, 3, 4, 6, 7, **8, 11, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13, 11, 13**	all
D-efficiency (%)	54.79	87.16	90.72	91.16	82.07
A-efficiency (%)	31.25	82.62	89.55	89.01	79.58
Max pred. var.	6.83	1.07	.6	.574	.49

D-Optimal Design With 6 Forced Points Model Incl.: Linear Effects/Interacts B*C

Number Runs	Forced Pts	9 RUNS	10 RUNS	12 RUNS	13 RUNS
Selected Runs (added runs in bold)	1, 2, 3, 4, 6, 7	1, 2, 3, 4, 6, 7, **8, 11, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13**	1, 2, 3, 4, 6, 7, **8, 11, 12, 13, 6, 8**	all
D-efficiency (%)	54.79	87.16	90.72	91.16	82.07
A-efficiency (%)	21.28	82.62	89.55	89.01	79.58
Max pred. var.	9.	1.07	.6	.574	.49

DISTRIBUTION:

1	MS0501	J. Spooner, 2338
1	MS0516	G. W. Smith, 1564
1	MS0521	T. J. Young, 1567
25	MS0829	S. V. Crowder, 12323
1	MS0829	K. V. Diegert, 12323
1	MS0829	K. S. Marlman, 12323
1	MS0830	E. W. Collins, 12335
1	MS0830	M. F. Martin, 12335
1	MS0830	J. M. Sjulin, 12335
1	MS0957	R. C. Reuter, 1401
1	MS0960	J. Q. Searcy, 1400
1	MS9013	R. G. Miller, 2266
1	MS9014	T. R. Harrison, 2271
1	MS9014	W. G. Wilson, 2221
1	MS9033	C. L. Knapp, 2262
1	MS9034	D. J. Beyer, 2263
1	MS9035	R. D. Monson, 2265
1	MS9036	C. M. Hinckley, 2254
1	MS9108	E. T. Cull, 5139
1	MS9405	J. M. Hruby, 8230
1	MS9202	R. Bierbaum, 8116
1	MS9202	J. Handrock, 8116
1	MS9202	C. J. DeCarli, 8116 (contractor)
10	MS9202	A. J. Johnson, 8116
1	MS9420	L. A. West, 8200

1	MS9001	M. E. John, 8000;		
		Attn:	MS9002	P. N. Smith, 8500
			MS9003	D. R. Henson, 8900 (A)
			MS9005	R. C. Wayne, 2200
			MS9007	R. C. Wayne, 8400
			MS9031	P. E. Brewer, 8600
			MS9031	L. A. Hiles, 8800
			MS9054	W. J. McLean, 8300
			MS9405	M. T. Dyer, 8700

1	MS9004	M. E. John, 8100		
		Attn:	MS9056	J. Vitko, 8102
			MS9056	M. Lapp, 8102
			MS9103	G. Thomas, 8120
			MS9104	W. R. Bolton, 8120
			MS9105	H. H. Hirano, 8119
			MS9105	J. A. Lamph, 8118
			MS9201	L. D. Brandt, 8112
			MS9201	P. K. Falcone, 8114
			MS9202	R. E. Oetken, 8115
			MS9202	R. Rinne, 8100
			MS9214	L. M. Napolitano, 8130

3	MS9018	Central Technical Files, 8940-2
1	MS0899	Technical Library, 4916
1	MS9021	Technical Communications Department 8815/Technical Library, MS0899, 4916
1	MS9021	Technical Communications Department 8815 for OSTI