AIAA JOURNAL
Vol. 35, No. 5, May 1997

Statistical Experimentation Methods for Achieving Affordable Concurrent Systems Design

Wei Chen*
Clemson University, Clemson, South Carolina 29634-0921
and
Janet K. Allen,† Daniel P. Schrage,‡ and Farrokh Mistree§
Georgia Institute of Technology, Atlanta, Georgia 30332-0405

We describe an affordable method for designing in the conceptual stage using relatively high-fidelity concurrent systems analysis. Our method is rooted in three domains, namely, the design of experiments, the response surface methodology, and the compromise decision support problem. A sequential experimentation strategy as well as the heuristic rules for creating high-order response surface models are introduced as a cost-effective approach to applying the statistical experimentation methods in design of complex systems. A high-speed civil transport aircraft design is used as an example to illustrate the potential of our approach.

Nomenclature

d_i^+, d_i^-	= deviation variables in the compromise decision support problem
R	= regression coefficient
X	= system variables
Y	= performance variables

I. Introduction

OUR approach for design involving intensive computer simulations is shown in Fig. 1. First, there is the black box with inputs (system variables) and outputs (performance functions). The black box contains comprehensive analysis routines that are exercised to simulate various system responses. Because of the complexity of the system, a direct relationship between inputs and outputs is usually unknown and is analogous to a black box. The routines in the black box could be exercised blindly or systematically, resulting in cost-efficiency. We prefer a systematic approach and use design-of-experiment (DOE) principles to conduct a minimum ordered set of numerical experiments to generate response surfaces. The DOE techniques are formal techniques that support the design and analysis of experiments (Ref. 1, Chap. 1, and Ref. 2). Among various DOE techniques, the response surface methodology (RSM) (Ref. 3, Chap. 1, and Ref. 4) is a collection of statistical techniques that support the design of experiments and fitting response surface models. These response surface models are mathematical functions representing relationships between independent system variables X (factors), which are used to describe design decisions, and their dependent performance variables Y (responses), which describe the resulting system behavior. A second-order response surface model is shown in Fig. 1. A solution to the compromise decision support problem (DSP) represents values for a set of system variables that satisfy a set of constraints, goals, and bounds and minimizes the deviation between that which is desired G_i and that which is obtainable $A_i(x)$. The compromise DSP is unique in that it is a hybrid mathematical construct for formulating multiobjective design problems involving tradeoffs.[5] A detailed set of references to various applications of the compromise DSP is presented by Mistree et al.[6] The response surface models are used to approximately model the constraints and goals in the compromise DSP, thereby eliminating the need to run the comprehensive routines that would otherwise be part of the black box. In effect, we use the response surface models as fast analysis modules instead of some of the routines in the black box.

In designing complex systems, it has been shown that it is both effective and efficient to establish a direct relationship between the decision space and the performance space or space formed by state variables. As examples, RSM has been applied in the aircraft aerodynamic configuration design to select a set of design parameters that have the most significant impact on the system performance and achieve an optimal configuration based on the response surface model.[7] In structural design, RSM has been applied to approximate the mathematical model used for finite element analysis.[8] To obtain an accurate approximation, a variable complexity modeling method is proposed in which an approximate model is used in conjunction with a more accurate model.[9]

When applying the DOE approach, there is always a tradeoff between the number of experiments conducted and the accuracy of the response model. Because computer experiments differ from physical experiments in that there is no random error,[10,11] classic methods for the design and analysis of physical experiments (Ref. 1, Pt. 3 and Ref. 3, Chaps. 4 and 5) are no longer ideal for complex, deterministic computer models. We present a sequential experimentation strategy as well as the heuristic rules for creating high-order response surface models as a cost-effective approach to applying statistical experimentation methods in designing complex systems.

II. Tradeoff Between Number of Experiments and Accuracy

A. Sequential Experimentation Strategy

To improve both the adequacy of a response model and the experimental efficiency, a sequential experimentation scheme is proposed. This involves first performing low-order screening experiments over the entire possible design space to identify a region of interest and then performing secondary experiments to build on this region of interest a higher-order response model.

The sequential strategy is illustrated in Fig. 2. Given the number of factors, factor levels for discrete factors, bounds for continuous factors, and interesting range of each response, the procedure starts with screening experiments for fitting a low-order surface model and identifying the significance of different factors. Screening experiments are a class of designs used to study a large number of factors with a small number of experimental points. Many of today's

Received July 18, 1996; presented as Paper 96-4085 at the AIAA/USAF/NASA/ISSMO 6th Symposium on Multidisciplinary Analysis and Optimization, Bellevue, WA, Sept. 4–6, 1996; revision received Jan. 7, 1997; accepted for publication Jan. 13, 1997.

*Assistant Professor, Department of Mechanical Engineering. Member AIAA.

†Senior Research Scientist, Systems Realization Laboratory, George W. Woodruff School of Mechanical Engineering. Member AIAA.

‡Professor, Aerospace Systems Design Laboratory, School of Aerospace Engineering. Senior Member AIAA.

§Professor, Systems Realization Laboratory, George W. Woodruff School of Mechanical Engineering. Senior Member AIAA.

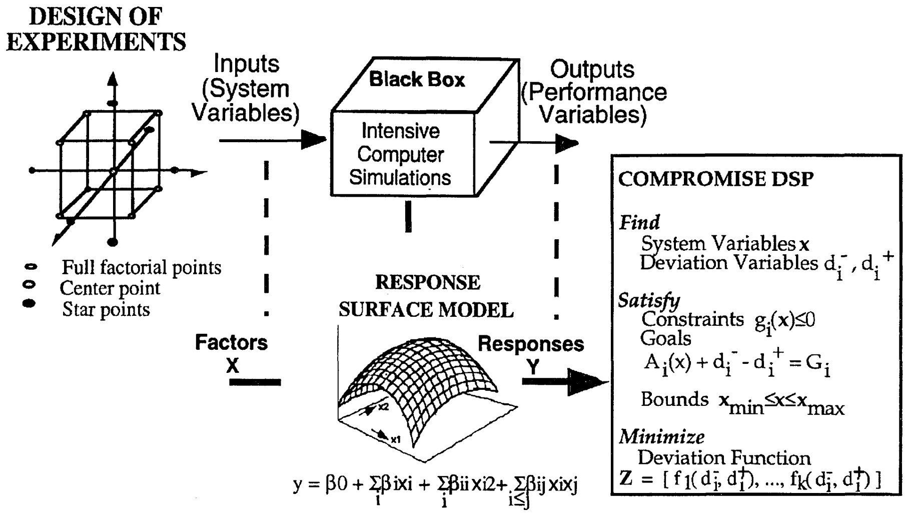


Fig. 1 Approach for design involving intensive computer simulations.

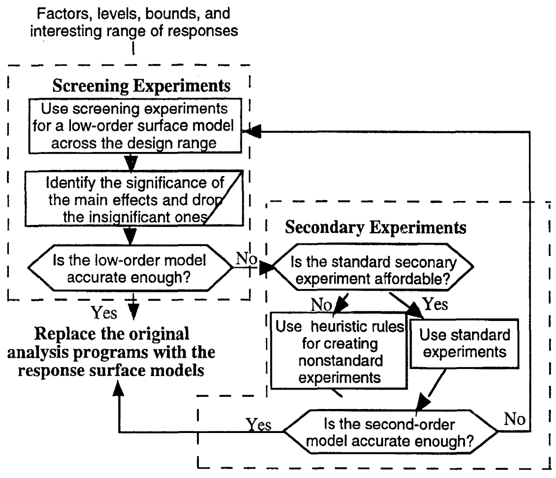


Fig. 2 Sequential experimentation strategy.

widely used screening designs were originally described by Plackett and Burman[12] for fitting a linear model. Box and Draper (Ref. 3, Chap. 15) summarize the designs requiring only a small number of runs. Among various fractional factorial experiments (FFEs), a two-level FFE with centerpoints is recommended for screening experiment design because of its efficiency and its insensitivity to response nonlinearity. Three- and higher-level FFEs usually yield little improvement, but result in great design and experimental complexity.

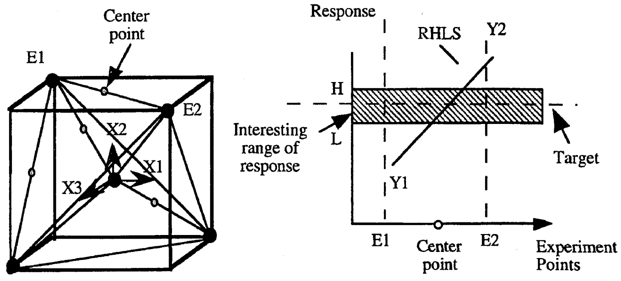
Results from the screening experiments are used to identify the significance of main effects (first-order single-factor effects) and to drop insignificant factors to reduce the size of the problem. We follow here the main-effect principle proposed by Lucas.[13] According to him, it is the empirical observation that main effects are more important than higher-order effects (whether they are two-factor interactions or quadratic effects). Hence, insignificant factors would be dropped with confidence based on only the main effects from screening experiments. For designs involving discrete system variables (design variables) with preferred levels, the results from screening experiments are used to identify the composite-basis design. The composite-basis design is a derivation of simulated designs and is used to embody the best characteristics of simulated or existing designs. Based on the contributions of factors at each individual level, the composite-basis design is a design with each control factor at its most favorable level, the level that is favorable for achieving the desired performance of responses. For discrete system variables, these favorable levels become the optimal solutions.

If the analysis of the results of the screening experiments and confirmation tests indicates that the initial model (e.g., a linear model) is accurate enough, the experiment is complete and the lower-order model, perhaps a linear model, is used to replace the original analysis routines. Otherwise, the experiment proceeds to the second stage, which involves a higher-order response surface model over the continuous system variables. A higher-order model can be determined either by standard experiments such as central composite design (CCD) or by using heuristic rules to create nonstandard experiments. The latter is recommended when the standard second-order experiments are no longer affordable. Affordability is related to both the number of factors and the time needed for each experiment. The introduction and illustration of heuristic rules for creating nonstandard experiments are presented in Secs. II.B and III.C, respectively. After the second-order response surface model is fitted, a designer checks to determine whether the accuracy of the model is satisfactory. If it is, the models are used to replace the original analysis routines (routines in the black box shown in Fig. 1). If it is not, alternative screening experiments must be used.

B. Heuristic Rules for Creating Nonstandard Secondary Experiments

A typical CCD consists of a complete or fraction of a first-order (2^n) factorial design, where parameter levels are coded to the usual -1 and $+1$ values and two star points on the axis of each design variable at a distance α from the center and one or a number of centerpoints are used (Fig. 1). For a design problem involving huge numbers of variables, even after reduction of design factors by screening experiments, it still may be too expensive to use the standard CCD for secondary experiments. For example, for 10 factors, the standard CCD using full factorial experiments as the factorial portion requires 1045 ($1 + 2n + 2^n$) experiments. However, to fit a second-order model, as long as the number of experiments is larger than the total degrees of freedom (DOF) of the regression equation, i.e., one plus the number of to-be-identified coefficients $2n + n(n-1)/2$, is acceptable. Based on DOF, the minimum number of experiments for 10 factors is 66. Using the terminology of statistics, the designs in which the number of experiments equals the number of coefficients to be estimated in the response surface models are called saturated designs. Saturated second-order designs include the Box–Draper design (Ref. 3, Chap. 15) and designs based on D-optimality.[14] A major limitation of a saturated design is the poor coverage of the region of interest. It usually works well in initial screening situations in which it is assumed that there will be a few important parameters. We develop heuristic rules for creating nonstandard experiments for fitting high-order models.

We consider a nonstandard experimental design as one that is based on screening experiments. To improve the accuracy of a regression model, it is important to find those good experimental points whose performance falls within the interesting response range. The target-oriented design (TOD) concept was developed to rate the direction in an experimental design in an effort to locate points in the design space that meet all of the constraints.[15] A similar concept is utilized in this work but is modified to develop heuristic



a) Design space b) RHLS

Fig. 3 Target-oriented design.

rules for creating nonstandard experiments. Modifications are made in the following two areas: the goodness measure and the way of finding the additional experimental points.

The original TOD concept is illustrated in Fig. 3. A direction is defined as a hyperline segment (HLS) between two experimental design points (full dark points in Fig. 3a). E_1 and E_2 are points used to explain our approach. Associated with each direction are response hyperline segments (RHLS) for each response, e.g., $(E_1, Y_1) \rightarrow (E_2, Y_2)$. A direction is considered good if each RHLS passes through its respective interesting response range. All good directions are rated and sorted on the basis of a measure of goodness. Then, a new experimental design is generated at the centerpoints of the good-direction (halfway) points and is used as an additional experiment to fit a second-order model.

In the original TOD concept, goodness is measured by how far the RHLS deviates from the target value (centerline between H and L). To be considered as a good direction, the RHLS has to fully intersect the interesting range of response (a band formed by H and L); this means that the endpoint values of each response (Y_1 and Y_2 on Fig. 3b) must straddle the appropriate target values. A limitation to this approach is that, often, the number of good directions is smaller than the number of experimental points needed. In that case, some supplementary points along bad directions must be used. Several other limitations are associated with the original strategy. First, not every response is required to have a target value. Responses that are used to evaluate the design constraints [$g_i(X) \leq 0$ in Fig. 1], if they are within the constraint limits, are acceptable. Using the constraint limit as the target value may lead the good design points, which are still in the wrong region. Second, adding centerpoints, i.e., zero after normalization, to endpoint values as additional experiments may make the design matrix singular. We modify the heuristic rules to account for the limitations discussed earlier. As for the measure of goodness, as long as the RHLS intersects the interesting range of response, it is considered to be a good direction. The interesting response range is defined by lower and upper bounds of the response. Therefore, the goodness measure is irrelevant to the response target value. To find additional experimental points, several points along a good direction can be selected as additional experimental points but none should be picked in the bad directions. Picking several points in a good direction instead of using only the centerpoint avoids design matrix singularity. Because a multiobjective design problem involves multiple responses, before counting good directions and locating additional points along the good directions, the experimental results need to be normalized and transformed into a single measurement. To reduce the risk associated with using heuristic rules, the interesting range of responses is extended by 20% from the original range specified in the problem statement. Applications of the sequential experimentation strategy and the heuristic rules are illustrated and verified in Sec. III.

III. Example: A High-Speed Civil Transport (HSCT) Design

A. Problem Statement

In the early stages of designing an HSCT aircraft, the problem statement could be given as follows. Given the mission requirements, the engine-cycle concept, and the generic HSCT baseline model, it is necessary to identify the appropriate number of passengers and flight range and to develop concurrently the airframe configuration and propulsion system top-level specifications that meet HSCT overall design requirements, including performance requirements, economic competitiveness, and environmental considerations as well as downstream design considerations.

In Fig. 4, the system descriptors of a compromise DSP represent the information flow in this construct. The "given" portion consists of the mission profile, the mixed-flow turbofan of the engine concept, and the baseline model, which was developed by the School of Aerospace Engineering at Georgia Institute of Technology. This model is used as a starting point for the concept exploration process. The compromise DSP is used to "find" the values of system variables (number of passengers, flight range, and top-level specifications of airframe and propulsion systems) to "satisfy" overall design requirements, including both system constraints, which are the "demands," and system goals, which are the "wishes." The overall design requirements are listed in Table 1. The demands (requirements that must be met) are modeled as constraints in the compromise DSP, whereas the wishes (requirements that we would like to satisfy to the extent possible) are modeled as goals.

For design evaluations, the Flight Optimization System (FLOPS)[16] is a concurrent systems analysis program in which engine-cycle analysis, overall vehicle synthesis, and mission analysis are integrated. FLOPS captures the interactions between multiple modules that are based on disciplinary analyses, such as

Table 1 Overall design requirements

Demands	*Limits*
Range of gross weight (GW)	791,667 lb $\leq$ GW $\leq$ 1,108,333 lb
Range of fuel weight (FUEMAX)	458,3333 lb $\leq$ FUEMAX $\leq$ 641,667 lb
Range of nitrous oxide emissions (TNOX)	5,833 lb $\leq$ TNOX $\leq$ 8,167 lb
Range of Productivity Index (PI)	71 kn $\leq$ PI $\leq$ 99 kn
Upper limit on landing field length (FARLAND)	FARLAND $\leq$ 11,000 ft
Upper limit on takeoff field length (FAROFF)	FAROFF $\leq$ 11,000 ft
Upper limit on approach speed (VAPP)	VAPP $\leq$ 140 kn
Range of specific fuel consumption (SFC) at cruise speed	1.298 $\leq$ SFC $\leq$ 1.382 lb/h per lb
Range of lift-over-drag (LOD)	6.87 $\leq$ LOD $\leq$ 7.53
Limit on compressor discharge temperature (CDT)	CDT $\leq$ 1710°R
Wishes	*Target (T)*
Minimize GW	TGW = 900,000 lb
Minimize FUEMAX	TFUEMAX = 480,000 lb
Minimize TNOX	TTNOX = 6,000 lb
Minimize SFC at cruise speed	TSFC = 1.34
Maximize LOD	TLOD = 7.4

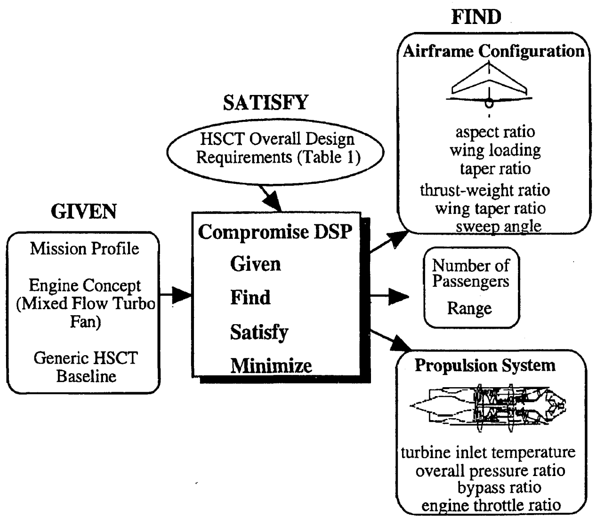


Fig. 4 HSCT design example.

aerodynamics, mission, propulsion, environment, and economics. Our intention here is not to advocate replacing FLOPS with response surface models for aircraft design; rather, FLOPS is used as an example of a computer-intensive analysis routine referred to in Sec. I.

B. Screening Experiments

Our screening experiments are designed for 14 factors, including 12 control factors and 2 noise factors. (Noise factor is a term used in robust design.[17] It represents design parameters that are uncontrollable in a design process. The principles of robust design can be used to determine robust top-level design specifications that are insensitive to variations of noise factors.[18]) In Table 2, factors 1–12 are the to-be-determined control factors. Three levels of passenger number (factor 1) and range (factor 2) represent distinct concepts of aircraft in the early stages of designing an HSCT. Therefore, factors 1 and 2 are considered to be discrete and the other 10 are considered to be continuous variables. Factors 13 and 14, turbine inlet temperature and burner efficiency, are noise factors because they are uncertain downstream design considerations. For screening, a Plackett–Burman (P&B) experimental design is used. Seventeen experiments are conducted for the 14 factors with the factor levels at the lower and upper bounds and the centerpoint.

From the results of the P&B experiments, it is noted that the constraints of VAPP, FAROFF, and CDT are well satisfied. These are easy-to-satisfy overall design requirements and are eliminated from further study. The results of the screening experiments also are used to identify the composite basis design, the design that embodies the best characteristics of simulated or existing designs (see Sec. II.A). To choose the most favorable level for one factor, statistical analysis is employed first to calculate contributions of the factor (factor effects) at each individual level (Ref. 1, Pt. 3). In Fig. 5, the factor effects of NPT are plotted. The factor effects at a specific level are calculated by averaging the results of the response at that level. On the horizontal axis, -1, 0, and 1 represent the lower, middle, and upper levels, respectively. The vertical axis represents normalized effects based on the range specified in the overall design requirements. Note that the middle level of NPT (0, i.e., NPT = 300) is a favorable level for minimizing GW, FUEMAX, TNOX, FAROFF, and maximizing PI, and SFC. Therefore, 300 is chosen as the most favorable level

Table 2 Factors and levels for screening experiments

Factor No.	Factor Description	Lower (-1)	Middle (0)	Upper ($+1$)
1	Number of passenger (NPT)	250	300	350
2	Range (DESIGN), nm	5000	5500	6000
3	Aspect ratio (AR)	1.5	2.3	3.1
4	Thrust/weight ratio (THRUST)	0.28	0.38	0.48
5	Wing loading (SW), lb/ft^2	115	120	125
6	Taper ratio (TR)	0.06	0.07	0.08
7	Thickness/chord ratio (TCA)	0.02	0.03	0.04
8	Sweep angle (SWEEP), deg	55	60	75
9	Overall pressure ratio (OPRDES)	18	20	22
10	Fan pressure ratio (FPRDES)	4.0	4.3	4.6
11	Bypass ratio (BPRDES)	0.1	0.2	0.3
12	Engine throttle ratio (TTRDES)	1.02	1.06	1.1
13	Turbine entry temperature (TET)	3300	3400	3500
14	Burner efficiency (BURNEFF)	0.96	0.975	0.99

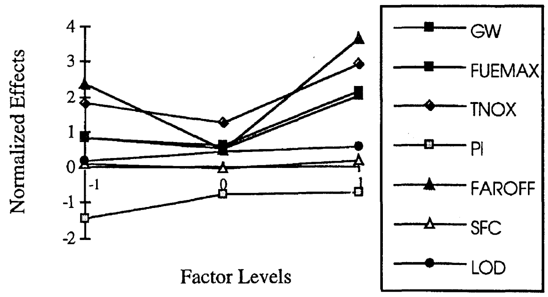


Fig. 5 Effects plot of NPT, number of passengers.

Table 3 Composite basis design

	Top-level design specifications	Level	Value
1	NPT	Middle	300
2	DESIGN, nm	Lower	5000
3	AR	Middle	2.3
4	THRUST	Middle	0.38
5	SW, lb/ ft^2	Middle	120
6	TR	Middle	0.07
7	TCA	Middle	0.03
8	SWEEP, deg	Middle	60
9	OPRDES	Middle	20
10	FPRDES	Middle	4.3
11	BPRDES	Middle	0.2
12	TTRDES	Middle	1.06

for the passenger number. Using the same strategy, it is found tha DRANGE = 5000 nm is the most favorable level. The strategy is re peated for choosing the appropriate level of the continuous top-leve design specifications. In Table 3, after tradeoffs are made in meetin the multiple overall design requirements the composite best design i given. Through statistical analysis of the screening experiments, it i also observed that factors 4, 5, and 6, i.e., TR, TCA, and SWEEP, ar not significant for all responses. They are considered as insignifican factors and are fixed at their middle levels in further experimentation

C. Secondary Experiments

From regression analysis of first-order models (linear models) it is observed that the regression coefficient R is far from 1. It in dicates that the linear model is not adequate for approximations Higher-order nonlinear response surface models thus are required To study main effects, two–two interaction and second-order effects at least 55 experiments, $\{1 + 2n + n(n-1)/2\}$, are necessary. Th heuristic rules discussed in Sec. II.B are applied to determine nev experiments based on the results of the 17 screening experiments, th interesting range of 7 responses, and their relative importance. Here it is assumed that the responses are equally important. Among th 136 ($17 \times 16/2$) HLS, 75 of them are identified as good direction intersecting the interesting range of response. One point on eacl good direction is picked for secondary experiments. Because 10 o the screening experiments also fall within the appropriate range o responses, we use a total of 85 nonstandard experiments for sec ondary experimentation. Response surface models based on the 8 nonstandard experiments are obtained. In all of these equations, th factors X_1–X_9 are normalized from -1 and $+1$. Equation (1) is th form of the response surface model, in which $n = 9$:

$$
\begin{aligned}
f(x_1, \ldots, x_n) \\
&= \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n \quad \text{linear terms} \\
&+ \gamma_1 x_1^2 + \cdots + \gamma_n x_n^2 \quad \text{quadratic terms} \\
&+ \beta_{12} x_1 x_2 + \cdots + \beta_{n,n-1} x_{n-1} x_n \quad \text{interaction terms}
\end{aligned} \tag{1}
$$

From the regression analysis it is determined that the accuracy o response surface models is satisfactory. An example of analysis o variance (ANOVA) is given in Table 4 for the regression of gros weight (GW). Sum of squares (SOS) is a measure of the deviation o the experimental data from the mean value of the data. The total SO is contributed from both the SOS of regression and the SOS of resid uals. The mean square of error (MSE), square root of SOS/DOF, i a parameter that captures the influence of DOF. A lower MSE o residuals means a more significant regression and greater accurac of the model. The regression coefficient R is used for measurin the accuracy of curve fitting, considering all of these factors. Th closer R is to 1, the more significant is the regression and the mor accurate is the regression model. In Table 4, R^2 is used. Also callec the coefficient of determination, R^2 measures the proportion of th variation around the mean explained by the regression (SOS regres sion/SOS total); R^2_adj is the ratio of MSE regression/MSE total

The calculation of $R^{**}2$_press is very similar to that of $R^{**}2$, except that the predicted response used in the equation is obtained at a particular experimental point when all of the experimental points except this point are used in the least squares regression.

One benefit of using quadratic response surface models is that after normalization of the design factors, the coefficients of the quadratic equation directly indicate the significance of the first-order effects (linear terms), interaction effects (interaction terms), and second-order effects (quadratic terms) and provide more insight into the problem. In Fig. 6, the contributions of main effects to all performance variables are plotted.

On the basis of Fig. 6 and the plots of significance of interaction effects and second-order effects, we observe that the main effects dominate in the quadratic response surface, even though the curvature caused by the quadratic effects seems to be important. This observation matches well with Lucas's main-effect principle, Sec. II.A. Note the contribution of factor 2 (THRUST) is extremely significant in the group of main factors and second-order effects; factors 1 and 2 (AR and THRUST) contribute especially to the curvature of response surface models. These observations about interaction effects are useful; they help to investigate the possibility of reducing the noise effects when robust design considerations are introduced in a design model. A strong interaction between a control factor and a noise factor indicates that the control factor level could be adjusted to dampen the effect of the noise factor in causing performance variation.

D. Response Surface Models in Place of Computer-Intensive Analysis Routines

The mathematical formulation of a compromise DSP for determining the top-level HSCT design specifications (nmu = no meaningful units) is as follows.

Given:

Response Surface Models of GW(X), FUEMAX(X), PI(X), FAROFF(X), SFC(X), LOD(X), and TNOX(X)

NPT = 300 and DRANGE = 5000 nm (identified by screening experimentation).

Mission profile (altitude, range, reserve fuel, etc.)

Engine type = Mixflow turbofan

Generic HSCT baseline configuration

Overall design requirements (Table 1)

Noise factors, turbine entry temperature, and burner efficiency are at their midlevels, $X_8 = X_9 = 0$ (normalized)

Table 4 ANOVA for the regression of gross weight based on nonstandard experiments

Source of variation	SOS	DOF	MSE	Regression coefficients R^2	R^2_adj	R^2_press
Regression	17.369	54	0.56715	0.97813	0.93876	0.87675
Residuals	0.38841	30	0.11378			
Total	17.758	84	0.45978			

Find:

The system variables (top-level design specifications), $X_i, i = 1, 7$

$$\text{AR}, X_1 \quad \text{THRUST}, X_2 \quad \text{SW}, X_3 \quad \text{OPRDES}, X_4$$
$$\text{FPRDES}, X_5 \quad \text{BPRDES}, X_6 \quad \text{TTRDES}, X_7$$

The values of the deviation variables associated with goals $G(X)$:

$$\text{GW}: d_1^-, d_1^+ \qquad \text{FUEMAX}: d_2^-, d_2^+$$
$$\text{TNOX}: d_3^-, d_3^+ \qquad \text{SFC}: d_4^-, d_4^+ \qquad \text{LOD}: d_5^-, d_5^+$$

Satisfy:

The system constraints $C(X)$ as determined by FLOPS or response surface models:

$$791{,}667\,\text{lb} \le \text{GW}(X) \le 1{,}108{,}333\,\text{lb}$$
$$458{,}333\,\text{lb} \le \text{FUEMAX}(X) \le 641{,}667\,\text{lb}$$
$$5{,}833\,\text{lb} \le \text{TNOX}(X) \le 8{,}167\,\text{lb}$$
$$71\,\text{kn} \le \text{PI}(X) \le 99\,\text{kn}$$
$$\text{FAROFF}(X) \le 11{,}000\,\text{ft}$$
$$1.298\,\text{lb/hr/lb} \le \text{SFC}(X) \le 1.382\,\text{lb/hr/lb}$$
$$6.87 \le \text{LOD}(X) \le 7.53$$

The system goals $G(X)$ as determined by FLOPS or response surface models:

minimize GW:

$$\text{GW}(X)/900{,}000 + d_1^- - d_1^+ = 1.0\ [\text{nmu}]$$

minimize FUEMAX:

$$\text{FUEMAX}(X)/480{,}000 + d_2^- - d_2^+ = 1.0\ [\text{nmu}]$$

minimize TNOX:

$$\text{TNOX}(X)/6000 + d_3^- - d_3^+ = 1.0\ [\text{nmu}]$$

minimize SFC:

$$\text{SFC}(X)/1.34 + d_4^- - d_4^+ = 1.0\ [\text{nmu}]$$

maximize LOD:

$$\text{LOD}(X)/7.4 + d_5^- - d_5^+ = 1.0\ [\text{nmu}]$$

The bounds on the system variables (normalized) (actual bounds; see Table 1):

$$-1 \le \text{AR} \le 1 \qquad -1 \le \text{THRUST} \le 1 \qquad -1 \le \text{SW} \le 1$$
$$-1 \le \text{OPRDES} \le 1 \qquad -1 \le \text{FPRDES} \le 1$$
$$-1 \le \text{BPRDES} \le 1 \qquad -1 \le \text{TTRDES} \le 1$$

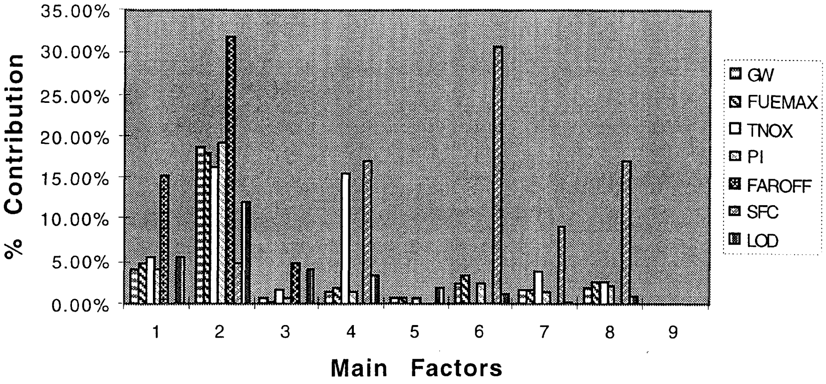


Fig. 6 Significance of main effects. System variables: 1, AR; 2, THRUST; 3, SW; 4, OPRDES; 5, FPRDES; 6, BPRDES; 7, TTRDES; 8, TET; 9, BURNEFF. (See Tables 1 and 2, respectively, for keys to abbreviations of performance and system variables.)

Minimize:

The sum of the deviation variables:

$$Z = \left[f_1(d_1^+),\ f_2(d_2^+),\ f_3(d_3^+),\ f_4(d_4^+),\ f_5(d_5^-)\right] \qquad (2)$$

The construct corresponds to Fig. 1. In the compromise DSP, each goal A_i has two associated deviation variables d_i^- and d_i^+ that indicate the extent of the deviation from the target G_i. The deviation function Z, which is to be minimized, can be formulated with either an Archimedean or a preemptive formulation.[5] To effect a solution on the basis of preference, goals are rank ordered into priority levels using the lexicographic minimum.[19]

The constraints and goals are derived from the overall design requirements (Table 1) with the elimination of well-satisfied constraints (Sec. III.A). These performance variables can be evaluated using either the original FLOPS program or the response surface models.

E. Verification

1. Verification of Heuristic Rules for Creating Nonstandard Secondary Experiments

The accuracy of the response surface models based on the nonstandard 85 experiments (see Sec. III.C) is compared with that of response surface models based on the standard CCD (531 experiments). One way of comparing accuracy is to compare the regression coefficients R in the ANOVA table to see which is closer to 1. We find that the R based on 85 nonstandard experiments is close, in some cases even closer to 1 than the R based on 531 experiments. The accuracy of the response surface models based on the nonstandard 85 experiments is verified further by comparing grid plots based on the response surface models with those based on FLOPS simulations. In Fig. 7, an example is given in the plot of GW over two critical control factors: AR and THRUST, across the design range: AR = [1.5–3.1] and THRUST = [0.28–0.48]. On the left is the accurate GW behavior obtained by FLOPS simulations at the grid points. On the right is the grid plot based on the response surface model created by nonstandard experiments (85 experiments). It is observed that the design behavior obtained from the response surface model based on nonstandard secondary experiments is close to reality (simulation program).

2. Verifications of Response Surface Models in Place of Computer-Intensive Analysis Routines

The compromise DSP given earlier is solved from different starting points by using both the original simulation program (FLOPS) and response surface models. The results are compared. Initial starting points are varied to examine the effects they have on the final solution of the compromise DSP when using both analysis modules. Three different starting points, the lower bounds, the middle values, and the upper bounds, of the seven design variables are used for these test runs.

From Fig. 8, note that the constraint violation quickly converges to zero when using the response surface model. However, when using FLOPS, the convergence heavily depends on the initial starting point, which indicates the possibility of local minima resulting from the highly nonlinear behavior of the design model. In two of the three runs using FLOPS, the constraint violation

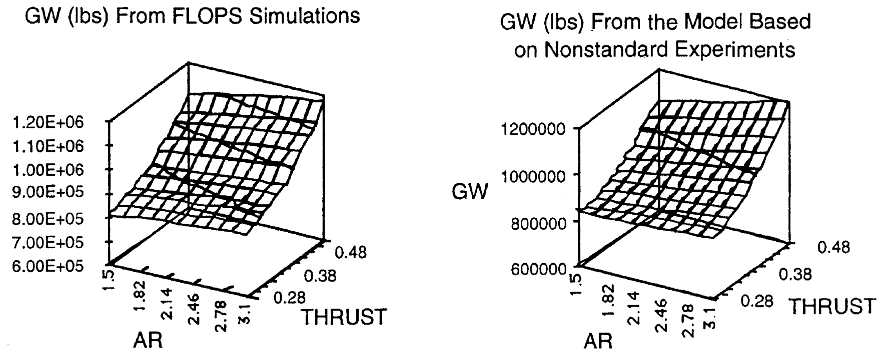
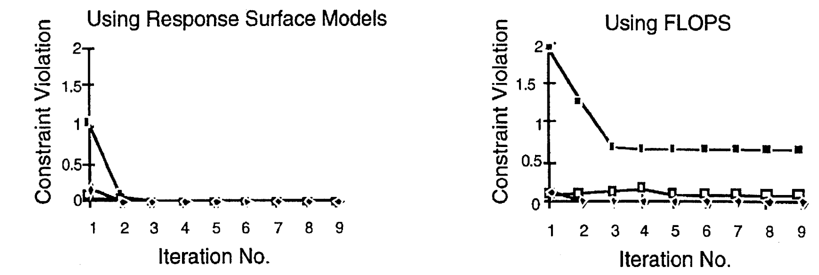
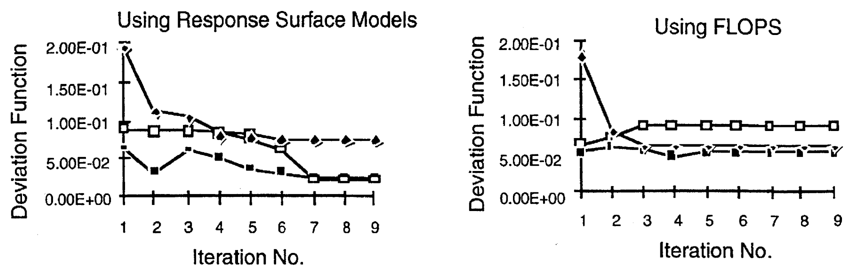


Fig. 7 Comparison of response surface models created by nonstandard experiments and FLOPS simulations.

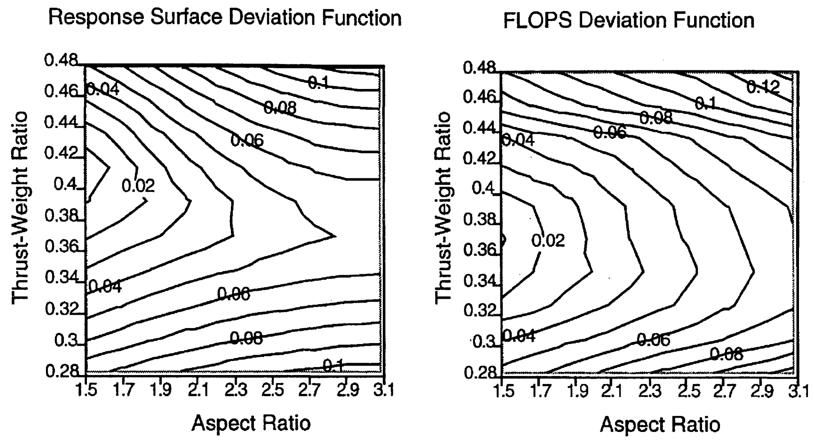


a) Convergence of constraint violation

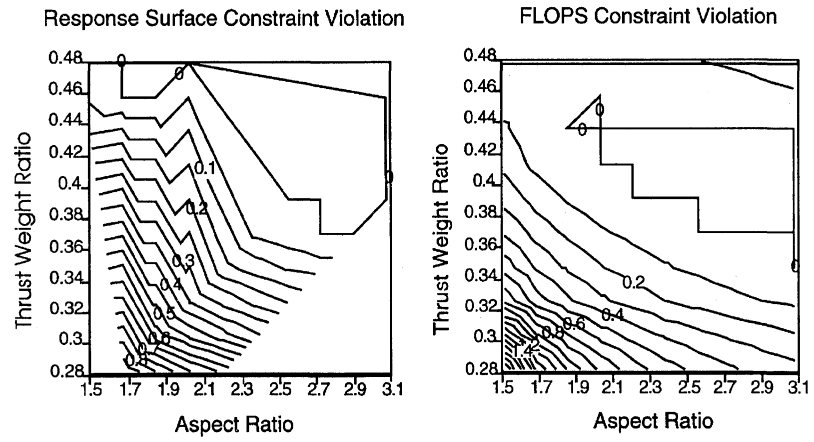


b) Convergence of deviation function

Fig. 8 Comparison of convergence of using response surface models and FLOPS: ■, lower; □, middle; and ♦, upper.



a) Comparison of values of deviation function



b) Comparison of constraint violations

Fig. 9 Comparison of constraint violations and the value of deviation function using response surface models and FLOPS.

does not converge to zero. This is not surprising because it has been observed from the grid plots that some of the responses based on the FLOPS program are erratic (nonsmooth), whereas the second-order response surface models are smooth functions. With a smooth function, the design point has a better chance of maneuvering around in the entire design space with a given solution algorithm, although a second-order model does not necessarily guarantee a global minimum solution from different starting points. Our study indicates that the use of response surface models might increase or, at least, does not eliminate the chance of finding a better solution.

How accurate are the results based on response surface models? Through confirmation tests, we find that the design solution achieved using response surface models is even better than that obtained using the FLOPS program. The deviation function value evaluated by FLOPS at the best design generated with the response surfaces is 0.056, which is better than 0.064, the smallest deviation function value obtained using FLOPS. When verifying the constraints, the previously eliminated well-satisfied constraints from the screening experiments also are examined. In Fig. 9, a comparison is made for constraint violation and deviation function value using response surface models and the FLOPS programs across a region of design space formed by two significant design variables, AR and THRUST. From Fig. 9, note that, in general, the patterns of contours based on response surface models and FLOPS programs are similar. This is reflected in similar topology in the decrease (increase) of constraint violation and the value of the deviation function. Note that, when a significantly higher degree of accuracy is desired, more elaborate experimental designs are needed for fitting the response surface model. Also note that the fidelity of contours depends on the number of simulation points. Here, the contour plots are quite ragged because the number of points (100) simulated across the design space is low.

IV. Conclusions

We report on our study of the application of statistical experimentation methods to concurrent systems design involving intensive computer simulations. We illustrate how statistical experimentation methods can be used to exercise the black box containing comprehensive analysis routines and create response surface models to replace them in solving the compromise DSP. Using an HSCT aircraft design as an example, we verify our approach and show how statistical experimentation methods can be used to improve design efficiency as well as effectively formalize design knowledge. Our major observations include the following.

1) Using the response surface models, not only can computational efficiency be greatly improved, but also the chance of finding a better result may be increased, or at least not eliminated (Sec. III.E).

2) The statistical methods provide an effective way to formalize the design knowledge in the early stages of design (Sec. III.B).

There are several advantages of the approach we present, as follows.

1) The sequential experimentation strategy for function approximation introduced in Sec. II.A gives several advantages. The use of first-order designs allows the design solution to move in the general direction of the "optimum" without the requirement to construct a second-order model. This provides the foundation for determining composite basis design illustrated in Sec. III.B. If the first-order design proves to be insufficient, a second-order design can always be constructed by augmenting the first-order model (Sec. III.C). The flexibility of fitting several types of models to the same data set also allows the user to experiment.

2) As introduced in Sec. II.B and illustrated in Sec. III.C, heuristic rules for creating nonstandard experiments can reduce significantly the number of secondary experiments for problems with a large number of design variables. Because the rules are based on the information of the interesting ranges of responses and the results

from screening experiments, compared to the standard experiments or recommended saturated designs in the literature, a better tradeoff between accuracy and efficiency is achieved.

We use the compromise DSP as an example of design programs for concurrent systems design to demonstrate our approach. Our approach can be extended to cases in which any type of design optimization programs is used. It is important to note the limitations of the statistical experimentation methods. A major concern is that when the performance is highly nonlinear and a second-order model is not adequate for approximations, transformations of the data are required or alternative approximation methods such as neural networks are needed to get a more accurate approximation. Usually, a large number of experiments is required to ensure accuracy when using these alternative approaches.

Acknowledgments

Wei Chen was supported by the G.W. Woodruff School of Mechanical Engineering at the Georgia Institute of Technology during her graduate study. We gratefully acknowledge National Science Foundation Grant DMI 9420405 and NASA Grant NAG-1-1564. We thank Rudi Cartuyvels (rc@imsnit.be) from LMS Numerical Technologies, Belgium, for providing us with a copy of NORMAN® to automate the process of creating response surface models. We thank the Aerospace Systems Design Laboratory at Georgia Institute of Technology for providing the HSCT baseline model.

References

[1]Box, G. E. P., Hunter, W. G., and Hunter, J. S., *Statistics for Experimenters*, Wiley, New York, 1978.

[2]Montgomery, D. C., *Design and Analysis of Experiments*, Wiley, New York, 1991, Chap. 1.

[3]Box, G. E. P., and Draper, N. R., *Empirical Model-Building and Response Surfaces*, Wiley, New York, 1987.

[4]Khuri, A., and Cornell, J. A., *Response Surfaces: Designs and Analysis*, Marcel Dekker, New York, 1987, Chap. 1.

[5]Mistree, F., Hughes, O. F., and Bras, B. A., "The Compromise Decision Support Problem and the Adaptive Linear Programming Algorithm," *Structural Optimization: Status and Promise*, edited by M. P. Kamat, Vol. 150, Progress in Astronautics and Aeronautics, AIAA, Washington, DC, 1993, pp. 247–286.

[6]Mistree, F., Smith, W. F., Bras, B., Allen, J. K., and Muster, D., "Decision-Based Design: A Contemporary Paradigm for Ship Design," *Transactions, Society of Naval Architects and Marine Engineers*, Society of Naval Architects and Marine Engineers, Jersey City, NJ, Vol. 98, 1990, pp. 565–597.

[7]Engelund, W. C., Douglas, O. S., Lepsch, R. A., and McMillian, M. M., "Aerodynamic Configuration Design Using Response Surface Methodology Analysis," AIAA Paper 93-3967, 1993.

[8]Schoofs, A. J. G., "Experimental Design and Structural Optimization," Ph.D. Dissertation, Dept. of Mechanical Engineering, Technische Univ. Eindhovern, The Netherlands, 1987.

[9]Gangadharan, S. N., Haftka, R. T., and Fiocca, Y. I., "Variable-Complexity-Modeling Structural Optimization Using Response Surface Methodology," AIAA Paper 95-1164, April 1995.

[10]Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P., "Design and Analysis of Computer Experiments," *Statistical Science*, Vol. 4, No. 4, 1989, pp. 409–435.

[11]Welch, W. J., Yu, T. K., Kang, S. M., and Wu, J., "Computer Experiments for Quality Control by Parameter Design," *Journal of Quality Technology*, Vol. 22, 1990, pp. 15–22.

[12]Plackett, R. L., and Burman, J. P., "The Design of Optimum Multifactorial Experiments," *Biometrika*, Vol. 33, Pt. 4, 1946, pp. 305–325.

[13]Lucas, J. M., "Using Response Surface Methodology to Achieve a Robust Process," *45th Annual Quality Congress Transactions*, Milwaukee, WI, Vol. 45, 1991, pp. 383–392.

[14]St. John, R. C., "D-Optimality for Regression Designs: A Review," *Technometrics*, Vol. 17, No. 1, 1975, pp. 15–23.

[15]Cartuyvels, R., and Dupas, L. H., "NORMAN/DEBORA: A Powerful CAD-Integrating Automatic Sequencing System Including Advanced DOE/RSM Techniques for Various Engineering Optimization Problems," *JSPE-IFIP WG 5.3 DIISM'93 Workshop (Design of Information Infrastructure Systems for Manufacturing)*, Tokyo, Japan, 1993.

[16]McCullers, L. A., "Flight Optimization System, User's Guide," NASA Langley Research Center, Hampton, VA, 1993.

[17]Phadke, M. S., *Quality Engineering Using Robust Design*, Prentice–Hall, Englewood Cliffs, NJ, 1989, Chap. 1.

[18]Chen, W., Allen, J. K., Mistree, F., and Tsui, K.-L., "Integration of Response Surface Method with the Compromise DSP in Developing a General Robust Design Procedure," *Advances in Design Automation*, edited by S. Azarm, H. Eschenauer, W. Gilmore, J. McCarthy, and M. Yoshimura, ASME DE-Vol. 82-2, American Society of Mechanical Engineers, New York, 1995, pp. 485–492.

[19]Ignizio, J. P., "Multiobjective Mathematical Programming Via the MULTIPLEX Model and Algorithm," *European Journal of Operational Research*, Vol. 22, No. 3, 1985, pp. 338–346.

A. D. Belegundu
Associate Editor