AIAA JOURNAL
Vol. 38, No. 5, May 2000

Partitioned, Multilevel Response Surfaces for Modeling Complex Systems

Patrick N. Koch,* Dimitri Mavris,† and Farrokh Mistree‡
Georgia Institute of Technology, Atlanta, Georgia 30332-0150

The most prevalent type of approximating functions employed for efficient engineering analysis and design integration are polynomial response surfaces. However, the construction of response surface approximations has been limited to problems with only a few variables, due to the number of analyses necessary to fit sufficiently accurate models. An approach is presented for partitioning response surfaces and constructing multilevel approximations for problems with larger numbers of variables. Using this approach, the (computer) experimentation necessary for fitting response surface models is reduced tremendously. A modified composite experimental design is also presented for the construction of response models that provide more consistently accurate predictions across the range of the design variables. The multilevel, partitioned response surface modeling and modified composite design approaches are demonstrated for the preliminary design of a commercial turbofan engine, an example problem defined in collaboration with Allison Engine Company, Rolls–Royce Aerospace Group.

Nomenclature

$\hat{y}$ = predicted response value
η = efficiency
$\hat{\mu}$ = mean response estimate
$\hat{\sigma}$ = standard deviation estimate

I. Overview of Problem

STATISTICAL techniques are widely used in multidisciplinary design to construct approximations of expensive computer analyses that are much more efficient to run, easier to integrate, and yield insight into the functional relationship between design variables $\boldsymbol{x}$ and performance responses $\boldsymbol{y}$. A recent review of statistical experimentation [design of experiments (DOE)] and approximation techniques for engineering design is given in Ref. 1. The most commonly used approximating functions are polynomial response surface equations. A recent review of several applications of response surface models in engineering design is also given in Ref. 1, and applications in structural design are presented in Ref. 2. In addition, many design methods have recently been developed incorporating response surface metamodeling[3] techniques.[4–6]

One significant limitation of response surface techniques is due to the problem of size for modeling large-scale, complex systems: the computational expense of experimentation associated with the combinatorial explosion in data points necessary for fitting models in a large number of variables. A detailed investigation into the problem of size is presented in Ref. 7 for the design of a high-speed civil transport. For complex systems, the number of variables affecting the system is often greater than desirable for response surface modeling with standard experiment designs (greater than 10 variables is often considered too large). In addition, if small or minimum size experiments are designed to accommodate modeling for large numbers of variables, the resulting models are frequently not of sufficient predictive accuracy (sparse experiments do not capture the trends sufficiently). The question being addressed in this paper, then, is the following: How can accurate-response surface models be fit efficiently for large numbers of design factors?

To address this question an approach for constructing multilevel, partitioned response surfaces is presented. Because this response modeling approach is implemented within the context of robust design, a review of approximation-based robust design and response surface metamodeling is first provided in Sec. II. The multilevel, partitioned response surface metamodeling approach is then presented in Sec. III and demonstrated in Sec. IV for the preliminary design of a commercial turbofan engine. Details for the robust design exploration formulation for this example problem are presented in Ref. 8; the focus in the current paper is on the multilevel response surface construction and validation of results obtained using these approximations. Closing remarks are presented in Sec. V.

II. Frame of Reference: Approximation-Based Robust Design

A. Robust Engineering Design

The basic approach to approximation-based robust design, shown in Fig. 1, is a generalization of response surface methodology,[9–11] Taguchi's parameter design,[12] and related recent engineering design approaches[4,6] in the literature. Note that although slight variations or a more detailed breakdown of this approach may exist, the basic approach of Fig. 1 is used for discussion purposes. The ultimate goal with approximation-based robust design is to arrive at improved or robust solutions efficiently. The factors (design variables) and responses (constraints, objectives) for a particular design problem provide the input or starting point for the approach of Fig. 1, and the solutions, improved or robust, become the outputs or results. To identify these solutions, this approach includes three sequential stages: screening, model building, and model exercising.

When the number of factors is too large for comprehensive exploration and/or when experimentation is expensive, screening experiments are used to reduce the set of factors to those that are most important to the response(s) being investigated. DOE is used to define the appropriate design analysis cases to evaluate the desired effects of the factors. If the factor set can be reduced, this reduced set provides the input for the second stage. In the model building stage, an experiment is run, and predictive metamodels are created to replace computationally expensive codes. If noise (uncontrollable) factors[13] are included, the mean and variance of each response must be estimated, and predictive models for both must be built. One approach for building these robustness metamodels is presented in Ref. 14, and three approaches are compared in Ref. 15. In the third stage (model exercising) the response models created

Presented as Paper 98-4958 at the AIAA/USAF/NASA/ISSMO 7th Symposium on Multidisciplinary Analysis and Optimization, St. Louis, MO, 2–4 September 1998; received 24 March 1999; revision received 30 August 1999; accepted for publication 1 September 1999.

*Graduate Research Fellow; currently Mechanical Engineer, Advanced Technologies and Applications, Engineous Software, Inc., 1800 Perimeter Park West, Suite 275, Morrisville, NC 27560; koch@engineous.com. Member AIAA.

†Assistant Professor, School of Aerospace Engineering. Senior Member AIAA.

‡Professor, G. W. Woodruff School of Mechanical Engineering; farrokh.mistree@me.gatech.edu. Associate Fellow AIAA.

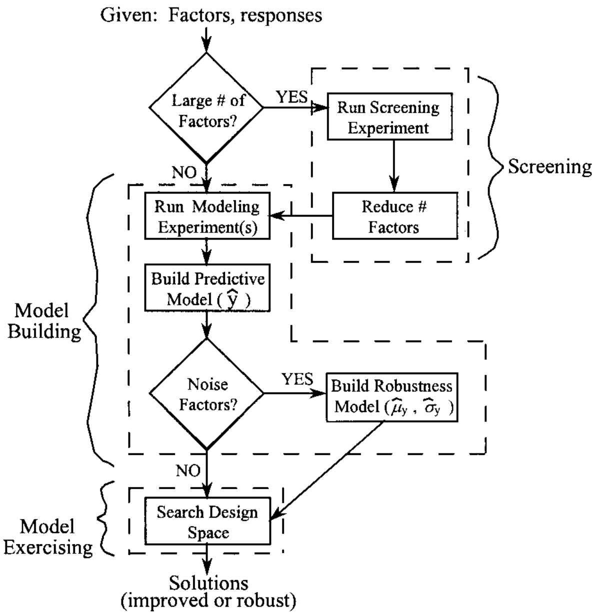


Fig. 1 Basic robust design approach.

in stage 2 are exercised to explore a design space efficiently while including uncertainty (noise) to identify robust solutions. The focus in this paper is on the model building stage of the approach of Fig. 1, particularly on response surface approximating functions.

B. Response Surface Metamodeling

The most widely used response surface approximating functions are simple low-order polynomials. If little curvature appears to exist, a predicted response value $\hat{y}$ can be obtained by fitting the first-order polynomial given in Eq. (1) (where x_i are the factors). If significant curvature exists, the second-order polynomial in Eq. (2), including all two-factor interactions, can be used. Thus,

$$\hat{y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i \tag{1}$$

$$\hat{y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i} \sum_{j>i} \beta_{ij} x_i x_j \tag{2}$$

The parameters (β terms) of the polynomials in Eqs. (1) and (2) are calculated using a least-squares regression analysis to fit these response surface approximations to existing data (empirical or generated from simulation/analysis routines). These approximations can then be used for prediction. A more complete discussion of response surfaces and least-squares fitting may be found in Ref. 11 and a discussion of experimental design in Ref. 16.

Among the various types of experimental designs for fitting a second-order response surface model and studying second-order effects, the central composite design (CCD) is probably the most widely used.[16] Central composite designs are made up of a two-level factorial or fractional of resolution at least V, $2k$ star points (where k is the number of factors) and at least one center point; these designs allow for the creation of a second-order surface with fewer points than would be required for a three-level full factorial design.

The two CCD experiments commonly used for computer experiments are shown in Fig. 2 for two factors; the factor ranges are generally normalized between -1 and 1. The central composite inscribed (CCI) experiment of Fig. 2a is offered to contain the experiment within the defined factor ranges. This composite experiment is a spherical design; the star points are defined by the factors ranges, and the factorial points are interior and defined by the combinations of $\pm 1/\alpha$. The value for α is normally chosen to ensure the design is either rotatable or spherical.[11] The predictive capability of a response surface fit using the CCI design, however, is unknown at the extremes of the factor ranges because no data points are taken at these extremes ([1, 1], [1, -1], [-1, 1], and [-1, -1] in Fig. 2a). In addition, the standard α value pushes the factorial points toward the

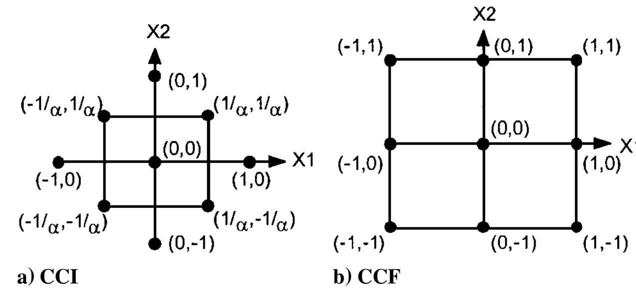


Fig. 2 Two forms of central composite experiment designs.

center with increasing factors ($1/\alpha$ is 0.707 for two factors and 0.297 for eight factors). Thus, the predictive accuracy of a response surface fit using the CCI design is even more questionable at the extremes with increasing factors. For this reason, this design should only be used if the region of interest is known to be spherical (interior points are of more interest than extreme points). If the extreme points are known to be important, the central composite face-centered (CCF) design shown in Fig. 2b is used. In this cuboidal design, the factorial points are pushed to the corner points defined by the factor extremes, keeping the star points at the factor extremes along each factor axis. With this configuration, however, other than the center point no interior data points are taken; thus only three factor levels, rather than five, are used (note that the CCF shown in Fig. 2b for two factors is identical to a three-level full factorial design; this is not the case for higher dimensions).

Unfortunately, with computer experiments, it is often not known in advance whether the interior points or extreme points are more important; a computer analysis code is often truly a black-box for which the input–output relationships are not clearly or completely understood. If the region where good solutions exist is known, the experiment should be designed around this region, and again it is unknown whether the CCI or CCF should be used.

III. Multilevel, Partitioned Response Surfaces

A. Constructing Partitioned Response Surfaces

An approach for fitting second-order polynomial response surface models in large numbers of design factors is described in this section. In this approach the factors and responses of a complex analysis code are grouped, and the response surface models themselves are partitioned to create multilevel models incorporating the effects of all factors.

Given a number of factors n and responses r associated with a particular complex analysis, the factors and responses are partitioned into two sets, for example. This partitioning is ideally based on domain knowledge; the factors and responses should be grouped based on knowledge about the problem and particular analysis, based on which factors are believed to affect particular responses directly. The exact partitioning is not necessarily critical, however, because each response is made a function of all factors; this is accomplished by concurrently constructing two-part partitioned response surfaces. For the partitioned factors and responses:

k = number of factors in set 1
$n-k$ = number of factors in set 2
s = responses of set 1, $\hat{y}_1$
$r-s$ = responses of set 2, $\hat{y}_2$

construct the partitioned response surface equations as follows. For responses of set 1,

$$\hat{y}_1 = \alpha_0 + \sum_{i=1}^{k} \alpha_i x_i + \sum_{i=1}^{k} \alpha_{ii} x_i^2 + \sum_{i} \sum_{j>i} \alpha_{ij} x_i x_j \tag{3}$$

where

$$\alpha_0 = \gamma_0 + \sum_{p=k+1}^{n} \gamma_p x_p + \sum_{p=k+1}^{n} \gamma_{pp} x_p^2 + \sum_{p} \sum_{q>p} \gamma_{pq} x_p x_q \tag{4}$$

For responses of set 2,

$$\hat{y}_2 = \beta_0 + \sum_{p=k+1}^{n} \beta_p x_p + \sum_{p=k+1}^{n} \beta_{pp} x_p^2 + \sum_{p} \sum_{q>p} \beta_{pq} x_p x_q \tag{5}$$

where

$$\beta_0 = \delta_0 + \sum_{i=1}^{k} \delta_i x_i + \sum_{i=1}^{k} \delta_{ii} x_i^2 + \sum_{i} \sum_{j>i} \delta_{ij} x_i x_j \tag{6}$$

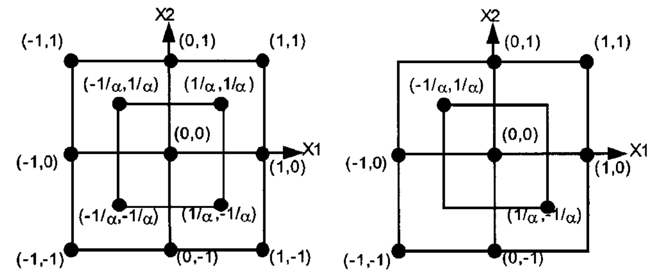
The first set of responses $\hat{y}_1$ are fit as a function of the k factors of set 1 [Eq. (3)], and the second set $\hat{y}_2$ is fit as a function of the $n-k$ factors of set 2 [Eq. (5)]. Two separate experiments are designed and run, ideally concurrently, to fit these two sets of response surfaces.

To capture the effects of the second set of factors (from $k+1$ to n) on the responses $\hat{y}_1$ modeled using the first set (from 1 to k), the mean term of these responses [α_0 in Eq. (3)] is fit as a function of the second set of factors; this is shown in Eq. (4). The original (fixed) mean term of Eq. (3) is thus replaced by the function in Eq. (4). The same is done for the second set of responses $\hat{y}_2$ using the first set of factors [Eq. (6)], and thus, the two-level response surfaces are created. The effect of modeling the mean term of each set of responses as a response surface itself is essentially to allow the intercept of the primary response surface [Eq. (3) or (5)] to move along the response, or y, axis with the secondary effects of the factors used to model the mean term. Note that in creating the second part of each set of response surfaces, the mean-term response models, no additional experimentation is necessary (two experiments are necessary for the four given models, not four experiments). Data gathered for fitting the response surfaces of Eq. (5) are used in fitting the mean-term response surfaces of Eq. (4); data gathered for fitting Eq. (3) are used in fitting Eq. (6). (Both sets of responses are monitored, and data are gathered during each experiment.)

The advantage of this approach is that the experimentation and model fitting expense is reduced tremendously. If, for example, 16 input factors are known to be important to the responses of a particular problem, fitting response models including all 16 factors using a full CCI or CCF design would require 65,569 computer analysis cases (note that using a resolution V fraction in the cube of the composite design would require fewer designs, as only 153 terms are to be estimated; however, note also that very small or minimum size designs can lead to models with poor predictive accuracy). If the 16 factors are partitioned into two sets of 8, and the responses are also grouped according to the two sets of factors, two experiments of 273 runs are required for the full CCD to fit the two-part response surface models of Eqs. (3–6). These two experiments can be run concurrently, and the experimentation expense to construct the models is cut by more than two orders of magnitude. Even without concurrency, with this partitioning, doubling the number of factors increases the necessary experimentation by a factor of two rather than quadratically [$2^k + 2^k$ rather than 2^{2k}, or $(2^k)^2$]. Alternately, a 64-trial 2^{8-2} design is resolution V; this design would require 81 runs for each experiment, reducing the expense of experimentation even further.

What is the limitation of this approach? The major assumption in this approach is that the interaction effects between the factors of each partitioned set are negligible or nonexistent. These interaction terms ($x_i x_p$ terms using the notation given earlier) are the only terms missing from the partitioned response surface models that would be present for response surfaces fit in all factors. The factors and responses must be partitioned such that these interaction terms are believed to be at least negligible. Many interactions are known to be nonexistent based on domain knowledge or can be shown to be nonexistent through experimentation. The factors and responses can then be partitioned so that only the interactions that make sense and are known to exist are included. If partitioning is difficult, however (if the dependency between subsets of factors is high), the omission of interaction terms between the sets can have a significant effect on the response models created.

In the approach just given, a set of factors and responses is partitioned into two groups, and the partitioned response surfaces then have two parts. This approach is not limited to two sets of factors and ranges, however; the factors and responses could be partitioned into more than two parts. Constructing the partitioned models would be increasingly more difficult to manage, but doable nonetheless. In this paper only two-level partitioned response surfaces are investigated. A modified composite design is used to construct these partitioned response models.



a) Combined CCI and CCF

b) Half-fraction, combined CCI/CCF

Fig. 3 Modified central composite experiments.

B. Modified Composite Design

A modified composite design is shown in Fig. 3. This experiment combines the factorial portion of both the CCI and CCF experiments (Fig. 2) and is designed for the case in which it is unknown whether the interior or extreme points are most important. Two versions of this experiment are shown in Fig. 3, the full combination of the CCI and CCF (Fig. 3a) and a half-fraction in the factorial points of the combined CCI/CCF (Fig. 3b). With the full experiment of Fig. 3a, whereas for the two-factor experiment only four points are added over the standard CCI or CCF designs of Fig. 2, for an eight factor experiment 529 total points would be needed compared to the 273 for the standard CCI or CCF. As the number of factors increases, the factorial portion of the composite design (increasing exponentially, 2^k) dominates the experiment, essentially doubling its size. For the combined CCI/CCF of Fig. 3b, a half-fraction of each factorial portion is taken. It is common to take a fraction of the factorial points with a composite design to reduce the number of experiments [many more points are defined in a standard composite design than are necessary for fitting the standard second-order response surface of Eq. (2)]. The experiment of Fig. 3b, then, is the same size as the standard CCI or CCF, but allows data to be gathered both in the extremes and the interior points in an attempt to obtain better model fits.

To space the experiment points evenly within the design space, the α value for the modified composite design is taken to be 2, placing the interior factorial points at the selected combinations of ± 0.5 for all normalized factor ranges. Rotatability is based on statistical measures deemed inappropriate for deterministic computer experiments,[17] and thus rather than defining α to ensure this property, space filling[18] has been suggested to be appropriate to construct sufficiently accurate approximations. Evenly spaced design points can be more appropriate to capture the effects of deterministic computer analyses. For more detailed discussions of deterministic computer experiments, see Refs. 1 and 18.

The modified design is investigated in detail and compared to the standard CCI design in Ref. 19; more consistent accuracy of predictions across the ranges of design variables is observed with the modified design. (The errors of predictions do not vary as much across the range of the factors with the modified composite design as seen with the standard composite designs for the examples investigated in Ref. 19.) Other designed experiments or sampling techniques, Latin hypercube sampling, for example, may also perform favorably, but are not investigated in this paper. The design of Fig. 3b is employed with the approach for multilevel partitioned response surfaces of Sec. III.A in the next section.

IV. Example: Preliminary Design of a Commercial Turbofan Engine

The two-level partitioned response surface modeling approach is tested and demonstrated for the preliminary design of a commercial turbofan engine. The example presented is a 31-kN (7000-lb) thrust class turbine engine designed for regional and business jet applications. This example problem was defined in collaboration with Allison Engine Company, Rolls–Royce Aerospace Group, and is based on the existing Allison AE3007 engine. The focus in this example is on the engine thermodynamic cycle design and the basic configuration design (weight, overall dimensions, and stage numbers). Whereas the cycle design is independent of the configuration factors, the configuration responses are dependent on the cycle factors in addition to the configuration factors; thus two-level partitioned response surfaces are fit for the engine configuration design. These constructed response models are used to evaluate multiobjective tradeoffs; robust solutions are identified and response surface approximations are verified. Note that robust design techniques are incorporated since noise factors (not controllable by a designer) are included in this preliminary design example. The objective in seeking robust solutions is to minimize performance variation in addition to attempting to achieve performance targets.

A. Commercial Turbofan Problem Definition: Requirements and Design Factors

The primary concern for this commercial turbofan engine example is fuel consumption, followed by thrust, weight, and overall dimensions. The design point investigated in this study is a 12-km (40,000-ft) cruise at Mach 0.8 and 25°C (77°F) (off-design is not investigated in this example). The Rolls Royce software package Engine Maker was employed for the cycle and configuration analysis. It should be noted that Engine Maker is a simplified engine analysis code, requiring roughly only 1 s per execution on a Pentium personal computer and is thus used here for illustrative purposes (approximations may not be necessary for the formulated optimization problem). More complex engine analysis tools, however, require significantly larger computational effort, making classical optimization approaches prohibitively expensive.

The commercial turbofan design requirements and the Engine Maker model of the AE3007 engine (AE3007 EM) are presented in Table 1. The thermodynamic cycle performance is captured through the engine specific fuel consumption (SFC) and the net thrust (at the 12-km cruise). Both of these performance characteristics are normalized in Table 1 (based on the actual AE3007 values) for proprietary reasons; thus, they are nondimensionalized. The engine configuration is defined by the engine bare weight and the overall dimensions (length and diameter, where the maximum fan diameter is used as representative of the maximum engine diameter for comparison purposes).

The requirements of Table 1 are defined through both constraint (hard) values and goals (soft performance targets). Constraint values define feasibility, whereas goals represent desired targets (define the ideal scenario). The constraint values in Table 1 are set to identify solutions that perform at least as well as the Engine Maker model of the AE3007 engine, whereas the goal values reflect the desire to achieve improved performance. The last requirement of Table 1 (Phot/Pcold, the ratio of core exhaust and bypass exhaust pressures) is a constraint that ensures mixing for a mixed-flow turbofan engine. These requirements become the desired responses for which response surfaces are created.

The design factors for this study are listed in Tables 2 (cycle factors) and 3 (configuration factors). The ranges for experimentation for these factors are defined based on experience and based on the AE3007 values for these factors (not shown) and are, thus, specific to this problem. The 12 cycle factors of Table 2 are classified as 8 control factors and 4 noise factors (the component efficiencies). One key cycle factor, bypass ratio, is not included in the list of independent control factors. This factor was made a function of fan outer pressure ratio (FOPR), overall pressure ratio (OPR), and stator outlet temperature (SOT) to ensure feasible mixed flow turbofan cycles during automated experimentation and, thus, is not an independent factor here.

Table 1 Commercial turbofan requirements

Requirement	AE3007 EM	Constraint	Goal
Performance			
SFC (normalized)	1.035	≤ 1.030	0.960
Net thrust (normalized, 12 km, 25°C day)	990	≥ 990	1052
Configuration			
Bare engine weight, kg	647	≤ 646	612
Max length, m	2.10	≤ 2.13	1.83
Max fan diameter, m	0.975	≤ 0.975	0.940
Other constraint			
Phot/Pcold	0.880	0.94–1.0	

Table 2 Commercial turbofan cycle design factors and ranges

Factors	Low	High
Control		
Compression/expansion parameters		
Inlet flow, kg/s	31.75	47.63
FOPR	1.5	1.85
OPR	20	35
SOT, K	1366	1644
Pressure losses		
Transition duct pressure loss (DPP21), %	0.5	5
Diffuser/combustor pressure loss (DPP3), %	2	6
LPT[a] exit pressure loss (DPP5), %	0.5	3
Bypass duct pressure loss (DPP13), %	1	4
Noise		
Fan tip polytropic efficiency, η_{fan}	0.84	0.88
HPC[b] polytropic efficiency, η_{HPC}	0.86	0.90
HPT[c] isentropic efficiency, η_{HPT}	0.85	0.92
LPT isentropic efficiency, η_{LPT}	0.88	0.92

[a] LPT, Low-pressure turbine.
[b] HPC, High-pressure compressor.
[c] HPT, High-pressure turbine.

Table 3 Commercial turbofan configuration design factors and ranges

Control factors	Low	High
Fan inlet Mach number	0.5	0.7
Fan entry hub/tip ratio	0.2	0.4
Fan tip speed, m/s	366	564
Fan hub loading	1.0	2.0
HPC tip speed, m/s	274	518
HPC hub loading	0.9	2.0
LPT mean loading	1.5	2.5

The four component efficiencies of Table 2 are defined as noise factors for this study, not under the designer's control during preliminary design. These factors are viewed as technology-level factors at this stage of engine design (higher efficiencies represent higher component technology levels) used to evaluate the effects of changes in technology at the component level. The efficiencies are needed for cycle evaluation, but cannot be predicted until sufficient component level detail is available. The purpose of including these efficiencies as noise factors at the cycle level, then, is to reduce the effects on the cycle design of variation or uncertainty in component efficiency values.

It could be argued also that the five pressure loss factors of Table 2 are not control factors. The losses are related to the Mach number through the associated ducts and the size of these ducts. However, the way in which Engine Maker has been configured, these factors are needed as inputs for cycle and configuration design. Thus, they are included as control factors, primarily to monitor their effects on the engine design.

The seven configuration factors of Table 3 are all classified as control factors. Thus, 19 total factors are to be investigated with Engine Maker. These 19 factors are the results of previous screening experimentation[19]; each factor has been found to have a significant effect on at least one of the requirement responses of Table 1. Thus, the problem size cannot be reduced further by removing additional factors. However, 19 factors are not manageable for the construction of second-order response surfaces (524,327 analyses would be necessary for a full central composite designed experiment). Thus the two-level partitioned response surface approach Sec. II.A is appropriate here. The partitioning necessary for the construction of two-level response surfaces is easily defined for this problem; the requirements and factors have already been classified for the thermodynamic cycle design and the mechanical configuration design. Thus, the problem is partitioned into the 12 cycle factors (4 of which are noise factors) of Table 2 and the 7 configuration factors of Table 3, even though Engine Maker is used for both cycle and configuration analyses. This partitioning makes the problem significantly more manageable, and the experimentation for the cycle and the configuration can be done concurrently (on different machines). Yet the approach of Sec. III.A allows the appropriate responses to be modeled in all 19 factors.

B. Commercial Turbofan Two-Level Response Surface Construction

For the cycle design experimentation, a product array approach (crossed inner and outer arrays)[13] is implemented. The modified composite experiment of Fig. 3b is designed in the eight cycle control factors of Table 2, with a total of 273 experiments ($1/2 \bullet 2^8 + 1/2 \bullet 2^8$ factorial points, interior and exterior, $2 \bullet 8$ star points, and 1 center point). This experiment, used as the inner control array in the product array, is crossed with an outer noise array. For the noise array, the four noise factors of Table 2 are varied over their ranges using a half-fraction of a two-level factorial ($2^{4-1} = 8$) with an added center point for nine total cases. The total number of experiment cases for the cycle design subproblem when crossing the two arrays, then, is 2457 cases (273×9). Note that, if a full composite design were designed in all 12 cycle control and noise factors, 4121 cases would be necessary (fewer cases could be used if a fraction of resolution V is designed in the cube portion).

For each of the responses monitored during the cycle experimentation (the cycle responses as well as the configuration responses for fitting the second level of these partitioned response surfaces), the mean and standard deviation data are calculated for each run of the inner control array across the runs of the outer noise array. Response surface models for mean and standard deviation are then fit to these data; resulting model fits are summarized in Table 4. For each mean and standard deviation response, second- and third-order response surfaces are fit, and the best fit is chosen. (The modified composite experiment is a five-level experiment, and thus, the third-order terms can be added to the basic model of Eq. (3) or Eqs. (4–7); three-way interaction terms are not added simply because higher-order interactions are usually assumed negligible.) The order of fit and R^2 values for each mean and standard deviation response are given in Table 4. Recall that the response models fit for the configuration mean responses in the cycle factors are actually the second portion of these models, the intercept term models (see Sec. II.A); the primary models for the configuration responses are fit in the configuration factors of Table 3. Also, no model is fit for standard deviation of the fan diameter response because this response does not change with the noise factors; no deviation is observed.

In fitting models for prediction, R^2 values as close as possible to 1 are desirable. The fits are extremely good for all of the mean response models, with R^2 values greater than 0.99 in every case. The standard deviation models are also very good for SFC, thrust, fan diameter, and Phot/Pcold. The R^2 values for the standard deviation of weight and length, however, are lower (0.838 and 0.963, respectively). Fitting models for these responses is more difficult because they change in discrete jumps with stage numbers, and this difficulty is magnified in calculating and modeling the standard deviation. However, the approach in robust design is to reduce the predicted variance, rather than actually achieve strict targets. Also the standard deviations for weight and length are smaller (relatively) than the deviations of the other responses. With these relatively small standard deviations, these standard deviation approximations for the configuration responses are accepted.

Table 4 Response surface fits for cycle and configuration experimentation

Experimentation	Order of fit	R^2 value
Cycle		
Performance		
Mean SFC	Third	0.998
Standard deviation SFC	Third	0.986
Mean thrust	Second	0.999
Standard deviation thrust	Second	0.991
Configuration (intercept term models)		
Weight (intercept term)	Second	0.998
Standard deviation weight	Third	0.838
Length (intercept term)	Third	0.997
Standard deviation length	Third	0.963
Fan diameter (intercept term)	Second	1.000
Other constraints		
Mean Phot/Pcold	Third	0.998
Standard deviation Phot/Pcold	Second	0.995
Configuration		
Weight	Third	0.929
Length	Third	0.991
Fan diameter	Second	1.000

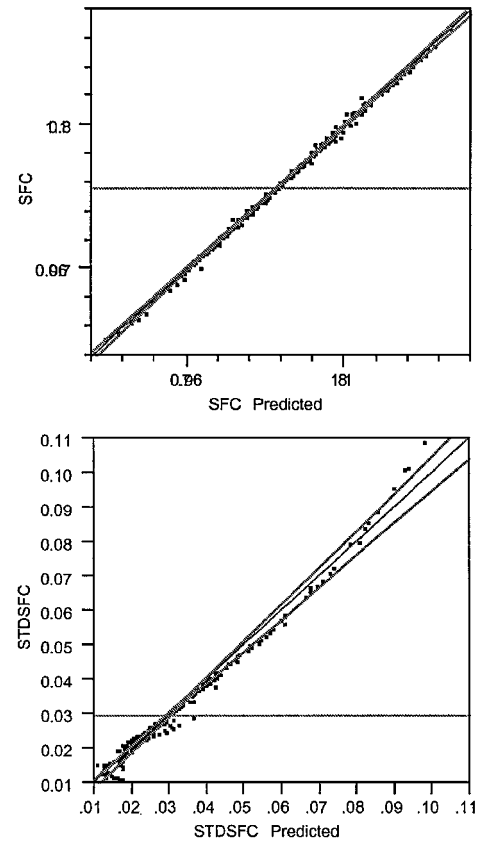


Fig. 4 SFC mean and standard deviation response fits (actual vs predicted).

In Fig. 4, the response model fits for the SFC responses (mean and standard deviation) of Table 4 are shown graphically as actual response data (experiment points) vs predicted response values. In these plots, the angled line represents the ideal fit (actual and predicted values being equal) around which the predicted data are scattered; the horizontal line represents the response mean value. As can be seen, the data are fit extremely well for mean SFC and very well for the standard deviation of SFC. Plots similar to those of Fig. 4 are presented in Ref. 19 for the remaining responses of Table 4.

For the configuration experimentation, the modified composite experiment of Fig. 3b is designed in the seven cycle control factors of Table 3. Because the configuration factors are all control factors, only the single modified composite experiment is needed, for a total of 143 cases. This experiment is run, and again both the configuration and cycle responses are monitored. For this configuration experimentation, however, the cycle responses (SFC, thrust, as well as Phot/Pcold) do not vary with the configuration parameters (identical values are obtained for these responses for all cases). Models are, thus, only fit for the configuration responses, and two-part partitioned response surfaces are not fit for the cycle responses.

The model fits obtained for the configuration responses are also shown in Table 4. Third-order fits are employed for weight and length. The R^2 values are very good for engine length and fan diameter (both over 0.99), but not quite as good for weight. The effects of the configuration factors on the number of stages for each component makes it more difficult to model the effects of these factors on the engine weight than for the cycle factors. Because the weight is a rough approximation at the system level, this model is accepted.

In this section, 14 response surface models have been constructed, 11 through the cycle experimentation and 3 through the configuration experimentation. The configuration responses have been fit as two-level partitioned response surfaces, with the first surface fit in the configuration factors and the second surface portion (mean term) for each configuration response fit in the cycle factors. Robustness response models (models for mean and standard deviation) have been constructed for all responses (with the exception of fan diameter). These response models are used to replace Engine Maker to allow efficient exploration of the preliminary design space to identify robust solutions.

C. Commercial Turbofan Response Surface-Based Robust Design Exploration

The response surface models constructed for this commercial turbofan example are employed for robust design space exploration based on the requirements of Table 1. A detailed description of the compromise Decision Support Problem (DSP)[20] formulation for this problem may be found in Refs. 8 and 19; in this paper, results obtained using the multilevel response surfaces are presented and evaluated.

Two design scenarios (goal achievement preferences) are evaluated for this example. In scenario 1, the mean response goals for SFC, thrust, weight, length, and fan diameter are included (no standard deviation goals) and are all weighted equally. This scenario represents the case in which robustness is not addressed and is used as a baseline for comparison with the robust solution. In scenario 2, the standard deviation goals are included in the formulation, and again equal weights are employed. This scenario represents the case in which both achieving performance targets (mean value) and reducing variation associated with the uncertain efficiencies are equally important.

The robust preliminary design exploration results obtained for these two design scenarios are given in Table 5. Also included in Table 5 are the cycle and configuration values for the Engine Maker model of the AE3007 engine (AE3007 EM). For both scenarios, feasible, converged solutions are obtained, and the performance and configuration values achieved (response surface predicted values) are at least slightly better than those of AE3007 EM (SFC, weight, length, and fan diameter are less in every case, and thrust is greater for all five scenarios). The cycle and configuration solutions behave as expected; when possible, the deviations from the goal targets are reduced (values closer to target). When robustness is added (standard deviation goals), mean performance is sacrificed slightly to improve robustness. Additional detail about these specific robust design exploration solutions, as well as additional robust design scenarios, can be found in Refs. 8 and 19. The focus in this paper is not on the numeric robust design results, except to demonstrate that the multilevel response surfaces can successfully be employed to identify robust regions of the design space and robust tradeoffs to be evaluated.

To evaluate the accuracy of the response surface predictions, the factor values of Table 5 are entered back into Engine Maker. For this evaluation, 1000 random combinations of the four noise factors (component efficiencies) are run for each scenario; the mean and standard deviation are then calculated for these 1000 runs. The errors of the predicted solutions of Table 5, when compared to the actual Engine Maker values, are given in Table 6. The negative error values reflect a response result that is underpredicted (value from response surface is less than actual value).

In Table 6 the errors for mean SFC, thrust, fan diameter, and the Phot/Pcold mixing constraint are less than 1% for both scenarios. As expected, the errors for mean weight and length are higher than those of the other mean responses (recall that these responses change discretely with component stage numbers). The errors for mean weight, however, are surprisingly very low, less than 2%. These errors are sufficiently small and acceptable. The errors for mean engine length, however, are much larger (8–13%). Fitting a continuous surface to these data leads to less accuracy than for the continuous responses. However, the engine length is overpredicted in every case. The actual engine length for these scenarios are all less than the goal target as desired. Because this requirement is not a primary concern, and does not lead to problems, the accuracy of the model is accepted for this demonstration.

The errors for the standard deviation predictions are significantly larger, ranging from 12% to over 200%. This result is consistent with the results of an investigation into noise modeling for approximation-based robust design presented in Ref. 15: Standard deviation is much more difficult to model than mean response; sufficient data points are needed to capture the standard deviation of a response. Here, the standard deviation data are gathered across the nine cases of the outer, noise array. If a full factorial were used for the noise array to gather more data, the size of the experiment would double; using a central composite design for the noise array would

Table 6 Response surface errors for results of Table 5

Engine maker	AE3007 EM	% error: Scenario 1, Overall	% error: Scenario 2, Overall, robust
SFC (normalized)	1.035	−0.398	−0.199
Std. deviation SFC	——	72.140	−40.684
Thrust (normalized)	990	−0.628	0.554
Std. deviation thrust	——	63.941	77.025
Weight, kg	647	−0.710	1.815
Std. deviation weight	——	−24.112	−13.226
Length, m	2.10	13.680	8.108
Std. deviation length	——	226.63	123.21
Fan diameter, m	0.975	0.305	0.055
Phot/Pcold	0.880	−0.299	0.100

Table 5 Commercial turbofan cycle and configuration robust preliminary design exploration results

Inputs	Design scenario 1, Overall	Design scenario 2, Overall, robust	Outputs	AE3007 EM	Design scenario 1, Overall	Design scenario 2, Overall, robust
Inlet flow, kg/s	41.27	39.96	SFC, normalized	1.035	1.002	1.005
FOPR	1.67	1.67	Std. dev. SFC, 1b/h/1b	——	0.0152	0.0111
OPR	20.30	20.00	Thrust (normalized)	990	1039.50	1014.97
SOT, K	1553	1640	Std. dev. thrust, N	——	143.6	130.2
Transition duct loss, %	0.79	3.45	Weight, kg	647	621.8	615.8
Diffuser/combustor loss, %	3.37	2.00	Std. dev. weight, kg	——	0.00368	0.00318
LPT[a] exit pressure loss, %	1.08	2.05	Length, m	2.10	1.80	1.83
Bypass duct loss, %	1.00	1.01	Std. dev. length, m	——	0.00841	0.00762
Fan inlet Mach number	0.70	0.69	Fan diameter, m	0.975	0.917	0.917
Fan entry hub/tip ratio	0.20	0.26	Phot/Pcold	0.880	0.999	1.001
Fan tip speed, m/s	563.9	497.7				
Fan hub loading	2.000	1.235				
HPC[b] tip speed, m/s	518.2	358.2				
HPC hub loading	0.900	1.163				
LPT mean loading	2.161	1.981				

[a]LPT, Low-pressure turbine. [b]HPC, High-pressure compressor.

basically triple the total number of cases run. Rather than increasing the experiment size, however, it is important to view the trends of the standard deviation predictions. Even though the standard deviation predictions have high errors, the trends have been captured. The robust solution of scenario 2 has actual standard deviation values less than the values for the AE3007 EM model (the standard deviations are reduced through this robust preliminary design exploration). Interestingly, even the standard deviations of scenario 1 (for which robustness is not included in the deviation function) are lower than those of the AE3007 EM model (scenario 1 is more robust than AE3007 EM). Because both mean performance and robustness are improved over the AE3007 EM model (without using this engine as a starting point for robust preliminary design exploration), the solutions of Table 5 are accepted.

For the commercial turbofan engine preliminary design example investigated here, second- and third-order response surfaces were sufficient to approximate the analysis codes. Alternate approximation approaches are currently being investigated for more nonlinear analyses.[17,18,21,22]

V. Conclusion

The multilevel, partitioned response surface modeling approach of Sec. III has been successfully demonstrated in Sec. IV for the commercial turbofan preliminary design example problem. Using the two-level partitioned response surfaces, the configuration design responses are modeled in all 15 control factors (7 configuration factors and 8 cycle factors). With this approach, partitioning the factors and responses based on the configuration and cycle design reduces the necessary experimentation tremendously. If a single full composite experiment was to design in all 15 control factors and was crossed with the 4-factor noise array, 295,191 analysis cases would have been required. With the partitioned response surface approach and concurrent cycle/configuration experimentation, the experimentation is limited by the cycle experimentation (the crossed inner and outer arrays, 2457 cases). Experimentation expense is, thus, reduced by more than two orders of magnitude. By further increasing concurrency, running the cycle inner, control array concurrently (distributed computing) for each of the 9 cases of the noise array, the experimentation time can be reduced to the computational expense of the 273 cases of a single cycle control array.

The multilevel partitioned response surface metamodeling approach presented here is appropriate and useful for large-scale complex design problems in which 1) the number of variables prohibits standard response surface experimentation and modeling, 2) the problem can be easily partitioned and interactions between factors of partitioned sets can be neglected, and 3) second- or third-order polynomial response surfaces are adequate to model the factor/response relationships of a complex analysis code(s), over the desired factor ranges, with sufficient accuracy.

The limitation of response surface approximations that has not been discussed here is associated with the third condition. For highly nonlinear analyses, alternate approximation approaches must be investigated for efficient analysis and design space exploration. One approach in the current literature, kriging,[17] is being tested for engineering applications[18,21] and compared to response surface approximations,[22] and it shows promise as a viable approximation option for nonlinear problems.

Acknowledgments

Financial support from National Science Foundation Grant DMI-96-12365 is gratefully acknowledged. Cost of computer time was underwritten by the Systems Realization Laboratory, Georgia Institute of Technology. This research was conducted while Patrick Koch was a doctoral research fellow at Georgia Institute of Technology. Allison Engine Company, Rolls–Royce Aerospace Group, particularly David Eames, is gratefully acknowledged for use of the Engine Maker software and for assistance in defining the commercial turbofan engine preliminary design example.

References

[1]Simpson, T. W., Peplinski, J., Koch, P. N., and Allen, J. K., "On the Use of Statistics in Design and the Implications for Deterministic Computer Experiments," American Society of Mechanical Engineers, Paper DETC97/DTM3881, 1997.

[2]Barthelemy, J.-F. M., and Haftka, R. T., "Approximation Concepts for Optimum Structural Design—A Review," *Structural Optimization*, Vol. 5, 1993, pp. 129–144.

[3]Kleijnen, J. P. C., *Statistical Tools for Simulation Practitioners*, Marcel Dekker, New York, 1987.

[4]Chen, W., Allen, J. K., Mavris, D. N., and Mistree, F., "A Concept Exploration Method for Determining Robust Top-Level Specifications," *Engineering Optimization*, Vol. 26, No. 2, 1996, pp. 137–158.

[5]Kaufman, M., Balabanov, V., Burgee, S. L., Giunta, A. A., Grossman, B., Mason, W. H., and Watson, L. T., "Variable-Complexity Response Surface Approximations for Wing Structural Weight in HSCT Design," AIAA Paper 96-0089, Jan. 1996.

[6]Mavris, D., Bandte, O., and Schrage, D., "Application of Probabilistic Methods for the Determination of an Economically Robust HSCT Configuration," *Proceedings of the 6th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, AIAA, Reston, VA, 1996, pp. 968–978.

[7]Koch, P. N., Simpson, T. W., Allen, J. K., and Mistree, F., "Statistical Approximations for Multidisciplinary Design Optimization: The Problem of Size," *Journal of Aircraft*, Vol. 36, No. 1, 1999, pp. 275–286.

[8]Koch, P. N., Mavris, D., and Mistree, F., "Multi-Level, Partitioned Response Surfaces for Modeling Complex Systems," *Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, AIAA, Reston, VA, 1998, pp. 1954–1968.

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

[10]Khuri, A., and Cornell, J. A., *Response Surfaces: Design and Analysis*, Marcel Dekker, New York, 1987.

[11]Myers, R. H., and Montgomery, D. C., *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, Series in Probability and Statistics, Wiley, New York, 1995.

[12]Byrne, D. M., and Taguchi, S., "The Taguchi Approach to Parameter Design," *40th Annual Quality Congress Transactions*, American Society for Quality Control, Milwaukee, WI, 1987, pp. 19–26.

[13]Phadke, M. S., *Quality Engineering using Robust Design*, Prentice Hall, Englewood Cliffs, NJ, 1989.

[14]Chen, W., Allen, J. K., Tsui, K.-L., and Mistree, F., "A Procedure for Robust Design: Minimizing Variations Caused by Noise Factors and Control Factors," *Journal of Mechanical Design*, Vol. 118, No. 4, 1996, pp. 478–485.

[15]Koch, P. N., Mavris, D., Allen, J. K., and Mistree, F., "Modeling Noise in Approximation-Based Robust Design: A Comparison and Critical Discussion," *Advances in Design*, American Society of Mechanical Engineers, Paper DETC97/DAC-5588, 1988.

[16]Montgomery, D. C., *Design and Analysis of Experiments*, 4th ed., Wiley, New York, 1997.

[17]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.

[18]Booker, A. J., "Case Studies in Design and Analysis of Computer Experiments," *Joint Statistical Meeting: Proceedings of the Section on Physical and Engineering Sciences*, American Statistical Association, Alexandria, VA, 1996.

[19]Koch, P. N., "Hierarchical Modeling and Robust Synthesis for the Preliminary Design of Large Scale Complex Systems," Ph.D. Dissertation, George W. Woodruff School of Mechanical Engineering, Georgia Inst. of Technology, Atlanta, GA, Dec. 1997.

[20]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, AIAA, Washington, DC, 1993, pp. 247–286.

[21]Simpson, T. W., Allen, J. K., and Mistree, F., "Spatial Correlation Metamodels for Global Approximation in Structural Design Optimization," American Society of Mechanical Engineers, Paper DETC98/DAC-5613.

[22]Simpson, T. W., "Comparison of Response Surface and Kriging Models in the Multidisciplinary Design of an Aerospike Nozzle," NASA CR-1998-206935, ICASE Rept. 98-16, Inst. for Computer Applications in Science and Engineering, Hampton, VA, Feb. 1998.

A. Messac
Associate Editor