AIAA JOURNAL
Vol. 38, No. 12, December 2000

Response Surface Approximations: Noise, Error Repair, and Modeling Errors

Melih Papila* and Raphael T. Haftka†
University of Florida, Gainesville, Florida 32611-6250

In the past few years there has been interest in using response surface techniques to create surrogates to computer simulations. Response surface techniques allow detection and correction of errors as well as filtering out numerical noise, but these techniques introduce additional modeling errors. Methods of reducing both noise and modeling errors are explored. It is also demonstrated that repairing designs with large errors is preferable to eliminating these designs from consideration. Response surface approximations for a high-speed civil transport wing weight equation created from the results of a large number of structural optimizations are used for demonstration. It is shown that the statistical tools available for response surface techniques are effective for error detection and for filtering out noise. Once the noise is reduced, modeling errors can also be reduced by increasing the order of the approximation.

I. Introduction

RESPONSE surface techniques are becoming important tools in design optimization based on noisy computational simulations. In multidisciplinary optimization, numerical noise is often a problem in addition to the difficulty of coupling simulations from different disciplines. Response surface (RS) techniques filter out numerical noise, provide a convenient representation of data from one discipline to another, and provide for easy interface with an optimizer.

For example, in preliminary aircraft design, structural weight equations[1] are often used to represent structural weight in configuration optimization. Such equations can be generated by RS techniques when new aircraft concepts are not modeled well by traditional weight equations. Balabanov et al.[2] developed quadratic RS for wing-bending material weight W_b (structural weight needed to resist bending) for a high-speed civil transport (HSCT) aircraft as a function of 29 configuration design variables. They constructed the approximation based on structural optimizations for thousands of configurations to improve on the flight optimization system (FLOPS)[3] general weight equations for transport aircraft.

Numerical noise is an important issue in RS construction. Numerical experiments may be noisy for several reasons, including discretization errors, incomplete convergence of iterative procedures, and roundoff errors. Giunta et al.[4] used RS approximations to filter out the noise in aerodynamic simulations in design optimization for HSCT. They reported improvement in designs obtained based on the smooth response surface compared to the original noisy simulations. Balabanov et al.[5] investigated the noisy behavior of wing-bending material weight of HSCT designs. They found that a large portion of the noise was due to incomplete optimization of the wing camber that affects the wing-bending material weight via the aerodynamic loads. Much of the remaining numerical noise was due to the structural optimization process itself. Venter and Haftka[6,7] also made use of response surface approximations for filtering noise in finite element analyses. In their case, the noise was associated with the dependence of discretization error on shape design variables.

Whereas low-amplitude random noise is filtered well by RS approximations, data with large errors, outliers, may cause significant loss of accuracy. Therefore, robust regression employs techniques for statistical methods to detect and remove or weigh down outliers, such as iteratively reweighted least-square (IRLS) fitting.[8] Once detected, outliers can be investigated further for possible mistakes. In addition to errors in data, modeling (bias) errors may also compromise the accuracy of approximation if the weight is not modeled well by the low-order polynomials typically used for RS approximations. Higher-order polynomials may be used to reduce this problem (e.g., Venter and Haftka[6,7]). However, with high levels of numerical noise, higher-order polynomials may not help much in reducing overall errors.

The objective of the present paper is to investigate methods of obtaining more accurate RS approximations by dealing with both noise and modeling errors. In particular, we investigate the use of IRLS procedures to detect erroneous simulations that need to be repaired. We also investigate how the amount of numerical noise in the simulations affects the utility of using higher-order terms in the response surface. RS approximations for bending-material weight of the HSCT wing are studied for demonstrating the interactions of these two types of error. Section II provides a summary of the methodology used in the paper. Section III describes the HSCT design problem. Section IV describes how outlier detection and repair allow us to improve response surface accuracy, and Sec. V offers concluding remarks.

II. Methodology: RS Approximations

RS approximate numerical or physical experimental data by an analytical expression that is usually a low-order polynomial. The methodology[9,10] assumes that the RS expression is exact and that the differences between the data and the RS (called residuals) are due to uncorrelated, normally distributed random noise of magnitude σ in the experiments. The RS approximation $\hat{y}$ is written in terms of n_b coefficients b_i and assumed shape functions ξ_i, usually monomials, as

$$\hat{y}(\boldsymbol{x}) = \sum_{i=1}^{n_b} b_i \xi_i(\boldsymbol{x}) \tag{1}$$

The difference (residual) between the data y_j for the jth point $\boldsymbol{x}_j$ and the estimate defined in Eq. (1) is given as

$$e_j = y_j - \hat{y}(\boldsymbol{x}_j) \tag{2}$$

The residual can now be written in matrix form for n_d data points,

$$\boldsymbol{e} = \boldsymbol{y} - \boldsymbol{X}\boldsymbol{b} \tag{3}$$

where $\boldsymbol{X}$ is the matrix whose component (i, j) is $\xi_j(\boldsymbol{x}_i)$. The coefficient vector $\boldsymbol{b}$ in Eq. (3) is solved for minimum residual vector in

Presented as Paper 99-1312 at the AIAA/ASME/ASCE/AHS/ASC 40th Structures, Structural Dynamics, and Materials Conference, St. Louis, MO, 12–15 April 1999; received 31 July 1999; revision received 21 October 1999; accepted for publication 4 May 2000;

*Graduate Research Assistant, Department of Aerospace Engineering, Mechanics and Engineering Science; papila@aero.ufl.edu. Student Member AIAA.

†Distinguished Professor, Department of Aerospace Engineering, Mechanics and Engineering Science; haftka@ufl.edu. Fellow AIAA.

a least-square sense, and the remaining error vector $\boldsymbol{e}_r$ is found to satisfy

$$\boldsymbol{e}_r^T \boldsymbol{e}_r = \boldsymbol{y}^T \boldsymbol{y} - \boldsymbol{b}^T \boldsymbol{X}^T \boldsymbol{y} \tag{4}$$

where T denotes transpose. An unbiased estimate of the noise $\hat{\sigma}$ in the data is given as

$$\hat{\sigma} = \sqrt{\boldsymbol{e}_r^T \boldsymbol{e}_r \Big/ (n_d - n_b)} \tag{5}$$

If we fit the response surface with a large number of points and the true function is exactly modeled by Eq. (1), then $\hat{\sigma}$ represents error in the data that is corrected by the RS. That is, the RS will be more accurate than the data. However, with a finite number of data points, errors in the data cause errors in the coefficients, and that leads to a prediction error of the RS that depends on the location of the design point. The estimated standard error e_{es}, or square root of prediction variance, is an estimate of the prediction error in the response surface estimate at a design point. It is expressed as

$$e_{\mathrm{es}}(\boldsymbol{x}) = \hat{\sigma}\sqrt{(\boldsymbol{x}^m)^T (\boldsymbol{X}^T \boldsymbol{X})^{-1} (\boldsymbol{x}^m)} \tag{6}$$

where $\boldsymbol{x}^m$ is the vector of shape functions $\xi_j(\boldsymbol{x})$ used in Eq. (1). In addition, because Eq. (1) is only an approximation to the true function, $\hat{\sigma}$ will contain not only noise error, but also modeling error. Besides $\hat{\sigma}$, the quality of the approximation is often measured by the adjusted coefficient of multiple determination R_a^2

$$R_a^2 = 1 - \frac{\boldsymbol{e}_r^T \boldsymbol{e}_r \Big/ (n_d - n_b)}{\sum_{j=1}^{n_d} (y_j - \bar{y})^2 \Big/ (n_d - 1)} \tag{7}$$

where $\bar{y}$ is the average value of the response. An R_a^2 value larger than 0.9 is typically required for an adequate approximation.

To detect outliers, IRLS was used. IRLS procedures start with an initial response surface fitted to data and then low weights are assigned to points with large errors and the data refitted using a weighted least-square procedure. The process is repeated until convergence, and weights of points that do not fit the underlying model (outliers) tend to converge to small values. This effectively eliminates these points from the fitting process. The weight w assigned to a data point is given as[11]

$$w = \begin{cases} [1 - (|e/\hat{\sigma}|/B)^2]^2 & \text{if} \quad |e/\hat{\sigma}| \le B \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where e is the residual Eq. (2), $\hat{\sigma}$ is the estimate of noise Eq. (5), and B is tuning constant, usually $1 < B < 3$, and we used $B = 1.9$.

Coefficients b_i for IRLS approximation can be found by solving

$$\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X} \boldsymbol{b} = \boldsymbol{X}^T \boldsymbol{W} \boldsymbol{y} \tag{9}$$

where $\boldsymbol{W}$ is a diagonal weighting matrix using the weights from Eq. (8):

$$\boldsymbol{W} = \begin{bmatrix} w(e_1) & 0 & \cdots & 0 \\ 0 & w(e_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & w\left(e_{n_d}\right) \end{bmatrix} \tag{10}$$

When the weights are all one, the solution of Eq. (9) reduces to an ordinary least-square solution. However, in general Eqs. (8) and (9) are nonlinear equations, and the iterative IRLS procedure employs

$$\boldsymbol{b}^{(i+1)} = \boldsymbol{b}^{(i)} + \left[\boldsymbol{X}^T \boldsymbol{W}^{(i)} \boldsymbol{X}\right]^{-1} \boldsymbol{X}^T \boldsymbol{W}^{(i)} (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}^{(i)}) \tag{11}$$

Normally, the IRLS procedure is used only to eliminate or reduce the effect of errors by assigning them low weights.[8] Instead, we used the procedure for the purpose of identifying outliers for further evaluation.

For the present work, RS calculations were obtained using the JMP[12] software.

III. HSCT Wing-Bending Material Weight Problem

The example problem is a 250-passenger HSCT design with a 5500-n mile range and cruise Mach speed of 2.4. A general HSCT model[2,5,13] developed by the Multidisciplinary Analysis and Design (MAD) Center for Advanced Vehicles at Virginia Polytechnic Institute of State University includes 29 configuration design variables. Of these, 26 describe the geometry, 2 the mission, and 1 the thrust. Load cases for HSCT design studies are given in Table 1, where the first two reflect normal flight conditions, and the others represent severe and critical limit cases. Here, following Knill et al.,[13] we use two simplified versions of the problem with 5 and 10 configuration variables. The five-variable case includes fuel weight W_{fuel} and four wing-shape parameters: root chord c_{root}, tip chord c_{tip}, inboard leading edge (ILE) sweep angle Λ_{ILE}, and the thickness-to-chord ratio for the airfoil t/c. Five more variables, wing semispan $b/2$, outboard leading edge (OLE) sweep angle Λ_{OLE}, location of maximum thickness $(x/c)_{\max - t}$, leading edge (LE) radius parameter R_{LE}, and location of inboard nacelle, are added for the 10-variable problem. Figures 1a and 1b show a typical wing planform and the configuration variables for the five- and ten-variable cases, respectively, and Table 2 summarizes the configuration design variable ranges used for 5- and 10-variable cases. Fuselage, vertical tail, mission, and thrust-related parameters are kept unchanged.

To perform configuration optimization, an estimate of the weight of the aircraft is required. This weight is mostly estimated by traditional weight equations taken from the FLOPS[3] program. However, the weight of the structure needed to carry the bending loads on the wing is not well estimated by such equations because the HSCT carries a very large amount of fuel in the wing compared to subsonic transports. Consequently, a procedure was developed in Ref. 5 of estimating the bending material weight by structural optimization. The results of many structural optimizations are fitted with RS approximations as function of the configuration design variables and used in the configuration optimization. Our focus here is the uncertainty in these bending material weight RS approximations. In this study, a structural optimization procedure based on a finite element (FE) model developed by Balabanov et al.[2,5] with the GENESIS program[14] was used to estimate the bending material weight. The HSCT codes[2,5,13] calculate aerodynamic loads for the load cases given in Table 1. A mesh generator due to Balabanov et al.[2,5] creates the FE mesh distributes the aerodynamic and inertia loads onto the structural nodal points and generates input for GENESIS.

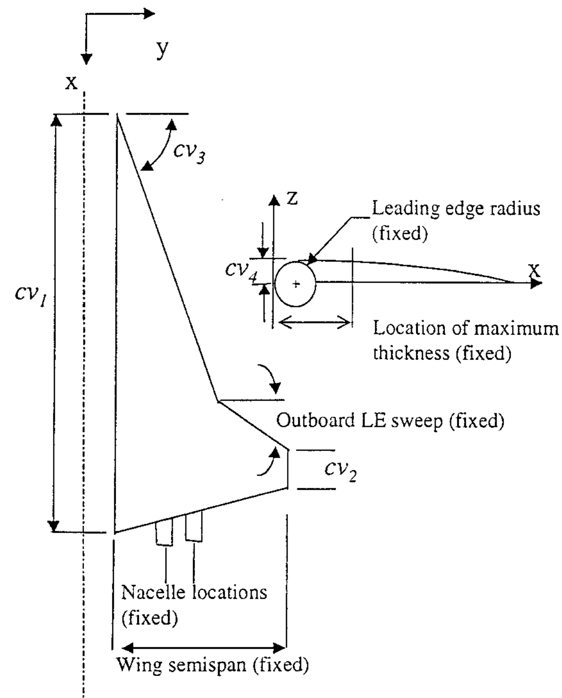
The FE model, shown in Fig. 2, uses 40 structural design variables, 26 to define skin panel thicknesses, 12 for spar cap areas, and 2 for the rib cap areas. The objective function is the total wing structural weight, and the wing-bending material weight W_b is calculated based on the values of the optimal structural design variables associated with bending resistance. This procedure generates numerical noise, and configurations of similar total structural weight can have significantly different bending material weights.

The optimization methods available in GENESIS are modified feasible directions (MFD), sequential linear programming, and sequential quadratic programming. Balabanov et al.[5] reported that none of the methods had definite advantage over the others in terms of numerical noise in their HSCT wing structural optimizations. Therefore, in this study the default optimization method, MFD, was used.

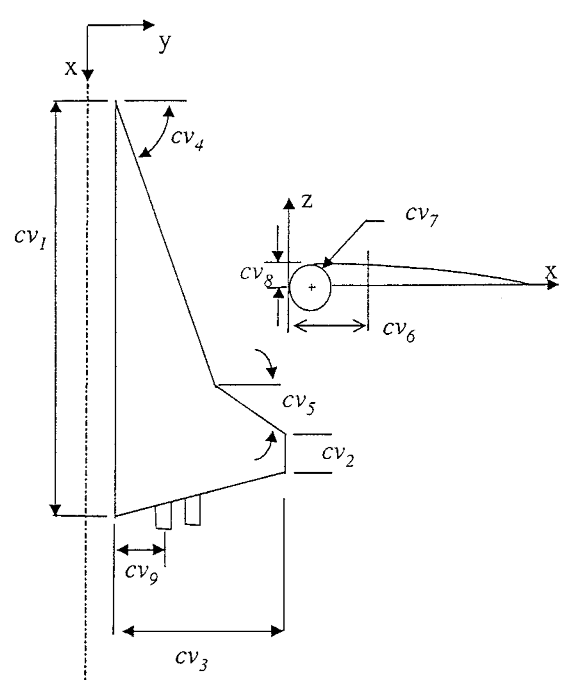
The numerical noise in the optimization depends on convergence parameters, called control parameters in GENESIS. These are categorized into move limit parameters, convergence criteria, and inner optimization control parameters. There are two loops in the GENESIS[14] as shown in Fig. 3. The outer loop performs detailed

Table 1 Load cases in HSCT design study

Load case	Mach number	Load factor	Altitude, ft
High-speed cruise	2.4	1.0	63,175
Transonic climb	1.2	1.0	29,670
Low-speed pullup	0.6	2.5	10,000
High-speed pullup	2.4	2.5	56,949
Taxiing	0.0	1.5	0



a) Case: 5 variables



b) Case: 10 variables

Fig. 1 HSCT wing planform and 5 and 10 configuration variables.

FE analysis of the structure and creates approximations of structural response to reduce the number of expensive FE analysis. The inner loop performs optimization based on the approximation with move limits imposed to restrict the optimization to the region where the approximation is valid. The approximation and optimization is continued until the change in the design variables is small enough (called soft convergence) or until the change in the objective function is small enough (called hard convergence). Convergence parameters affect the accuracy of optimization, quality of the optimization results, and the computational cost.

Three cases of control parameters are used in this study. Case 1 includes the default parameters provided by GENESIS. Case 2 is the same as case 1 except that a parameter, called ITRMOP, was increased from 2 to 5. Case 3, the settings used in previous HSCT

Table 2 Configuration design variables for HSCT with corresponding value ranges

Design variable	Number of configuration variables: 5	10
Root chord c_{root}, ft	150–190	150–190
Tip chord c_{tip}, ft	7–13	7–13
Wing semispan $b/2$, ft	74	58–78
ILE sweep, Λ_{ILE}, deg	67–76	67–76
OLE sweep Λ_{OLE}, deg	25	12–32
Location of max thickness $(x/c)_{\max-t}$, %	40	38–52
LE radius R_{LE}	2.5	2.1–4.1
Thickness to chord ratio at root $(t/c)_{\mathrm{root}}$, %	1.5–2.7	1.5–2.7
Inboard nacelle location y_{nacelle}, ft	20	10–35
Fuel weight W_{fuel}, lb	280,000–350,000	280,000–350,000

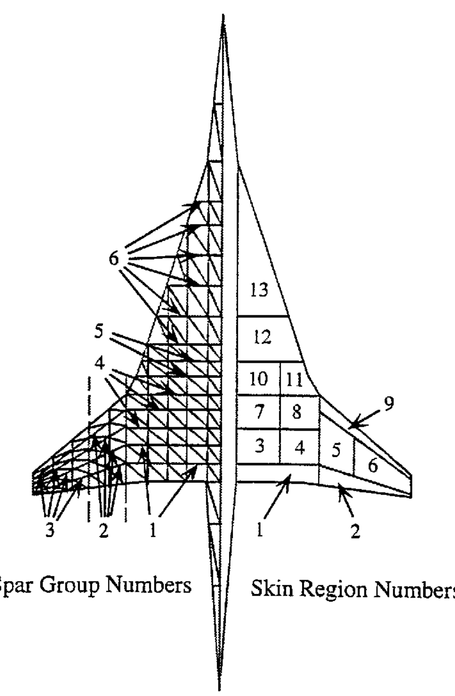


Fig. 2 HSCT finite element model and structural design variables.

studies,[2,5,15] employs tighter move limits and convergence criteria than case 1, with ITRMOP = 2. After extensive experimentation with the control parameters and help from the developers of GENESIS, we found that ITRMOP was the most important parameter for improving the accuracy of the optimization for our problem. It controls the convergence of the inner optimization. The inner-loop convergence criterion on objective function change must be satisfied ITRMOP consecutive times. We found that, for some configurations, complete convergence required ITRMOP = 5. In the following, we denote the bending material weight obtained with cases 1, 2, and 3 by W_{bd} (d for default), W_{bh} (h for high accuracy), and W_{bt} (t for tight move limits), respectively.

IV. Construction of RS Approximations

To improve the accuracy of the parameter estimates for the RS approximation, the HSCT configuration variables are scaled[9,10] to the range $(-1, +1)$. The root mean square error predictor $\hat{\sigma}$ [Eq. (5)], and R_a^2 [Eq. (7)] are used as measures of accuracy in the approximations. Note that $\hat{\sigma}$ includes both numerical noise, which is partly suppressed by the RS (so that the RS is more accurate than the data), as well as modeling error due to the low order of the RS. In the following, we try to estimate the two components. Compared to the original RS of Balabanov et al.,[5] the following measures are employed to improve accuracy: 1) application of higher-order models, 2) detection of outlier points by IRLS and their repair by

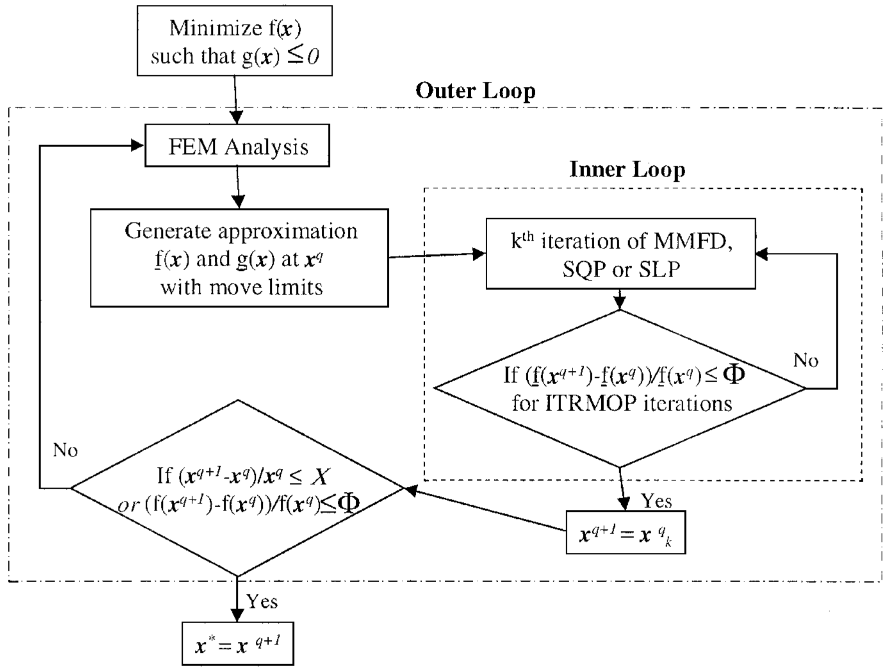


Fig. 3 Flowchart of the optimization in GENESIS (X, relative change criteria on design variables x; Φ, relative convergence criteria on objective function).

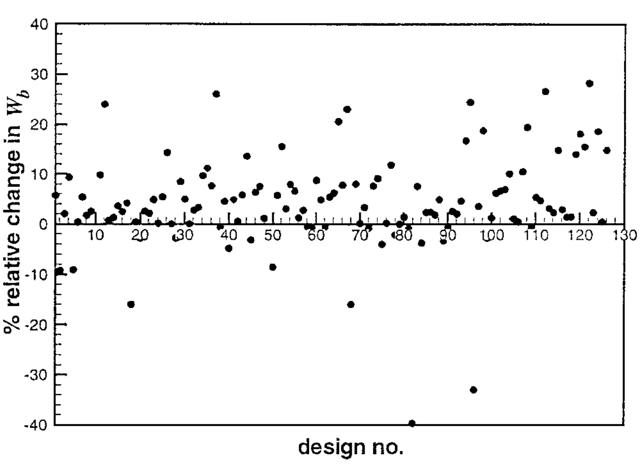


Fig. 4 Effect of overly tight optimization parameters: difference in optimization results between cases 1 (W_{bd}) and 3 (W_{bt}); 100 [(W_{bt} − W_{bd})/W_{bt}].

reoptimization using different convergence settings, and 3) removal of design points with excessively high weight.

A. Five Configuration Variables

Because we started with quadratic RS, a face-centered central composite design (FCCD) was used first. This design consists of a complete 2^5 factorial design (32 configurations), 2×5 (10) face-centered configurations, and one center configuration, for a total of 43 configurations. Next, to accommodate cubic approximations, another set of 43 design points was added, obtained by the FCCD with lower and upper limits of -0.75 and 0.75, respectively, so that this new design box is nested inside the original design box with a common center point. Then 27 configurations determined by D-optimality criterion[9,10,12] and 14 designs determined through an orthogonal array[15] were also added to the design space. Structural optimization results, W_{bd} and W_{bt}, were both found. Figure 4 shows how different convergence criteria affect the optimization results. Although W_{bt} uses tighter convergence criteria than the default W_{bd}, it yielded higher weights than W_{bd} for the majority of the design points. On further investigation, this turned out to be the effect of allowing very small move limits that produced premature convergence.

1. *Numerical Noise and Outlier Analysis*

The results obtained for wing-bending material weight by GENESIS for a given configuration differed by up to 40% depending on the optimization method, optimization settings, and even computer used.[16] For example, it can be seen from Fig. 4 that, although the bending material weight with default parameters W_{bd} are lower in general, they can be higher by up to 40% than values obtained with tight convergence parameters, W_{bt}. This indicates a substantial amount of noise and existence of outliers that can substantially degrade the accuracy of the RS approximation. RS approximations of full quadratic and cubic models as well as a cubic model with terms eliminated by stepwise regression were constructed. The first three rows of Table 3 summarize the results, which show that addition of the cubic terms did not improve significantly the accuracy of the RS. This indicates that the $\hat{\sigma}$ is dominated by noise.

Of the 126 configurations, 13 configurations were identified as candidate outliers for repair by IRLS with w [Eq. (8)] smaller than 0.1. Because the data are result of the minimization problem, we consider the outliers as candidates for repair only when the data are larger than the IRLS fit. The outlier configurations were repaired by repeating the optimization with ITRMOP increased from 2 to 5 (high-accuracy optimization). This repair procedure costs one additional, more expensive optimization per outlier. On average, the repair required twice the CPU time spent with ITRMOP $= 2$, but it led to substantially lower wing-bending material weight W_b for 11 out of 13 candidate outliers. Then, new RS approximations were constructed from the repaired data. Table 3 includes also the approximation results after the first repair, indicating improved accuracy through the reduction in $\hat{\sigma}$. Also note that the effect of using cubic terms is now more pronounced. This indicates that the error due to noise was substantially higher than the modeling error before the repair of the data. The average correction for the 11 outliers was about 12,000 lb or about 17%. The average W_b for all 126 points was reduced from 51,104 lb to 50,020 or by about 2%. Because the root mean square error is dominated by the larger errors, it was reduced by 4–5%.

To illustrate further how the process affects the approximation fit to the structural optimization results, the GENESIS results vs the corresponding approximations are plotted for the 126 points in Fig. 5. The diagonal line in Figs. 5 is the ideal, corresponding to exact equality between the two values. Figure 5a shows the original RS process, and Fig. 5b shows the results of the standard IRLS process. That is, the 13 outliers are assigned very low weights, so that

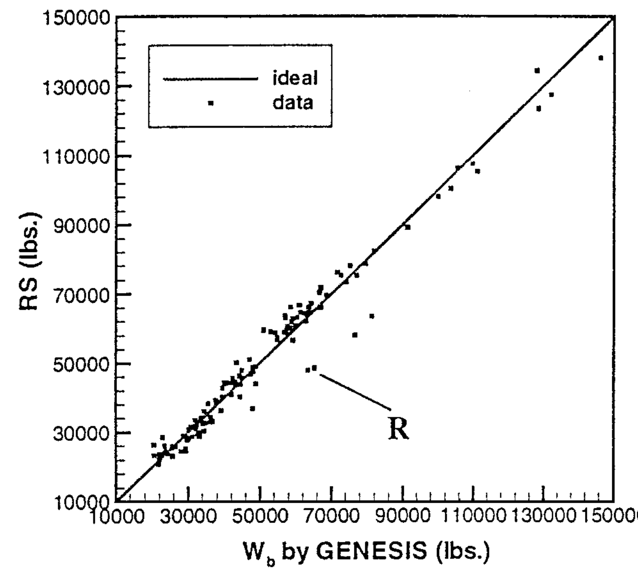
Table 3 Statistics for five-variable response surface approximations for wing bending material weight

RS approximation	$\hat{\sigma}$, lb	% of average W_b	R_a^2
126 Points; average W_b = 51,104 lb			
Full quadratic model	4874.6	9.54	0.9620
Full cubic model	4844.4	9.48	0.9624
Reduced cubic model	4329.2	8.47	0.9700
115 Points; average W_b = 45,229 lb			
Full quadratic model	4602.6	10.18	0.9166
Full cubic model	5005.3	11.07	0.9013
Reduced cubic model	4295.0	9.50	0.9273
126 Points: after outlier repair 1 ($13^a/13^b/11^c$); average W_b = 50,020 lb			
Full quadratic model	2758.9	5.52	0.9868
Full cubic model	2066.2	4.13	0.9926
Reduced cubic model	1925.8	3.85	0.9936
115 Points: after outlier repair 1 and reasonable design test; average W_b = 44,313 lb			
Full quadratic model	2054.3	4.64	0.9816
Full cubic model	1633.6	3.69	0.9884
Reduced cubic model	1500.8	3.39	0.9902
115 Points: after outlier repair 2 ($15^a/7^b/5^c$) and reasonable design test; average W_b = 44,125 lb			
Full quadratic model	1799.5	4.08	0.9856
Full cubic model	1408.0	3.19	0.9912
Reduced cubic model	1282.3	2.91	0.9927
115 Points: after outlier repair 3 ($16^a/3^b/2^c$) and reasonable design test; average W_b = 44,093 lb			
Full quadratic model	1762.3	4.00	0.9862
Full cubic model	1405.1	3.19	0.9912
Reduced cubic model	1274.3	2.89	0.9928
115 Points: min(W_{bd}, W_{bh})^d and reasonable design test; average W_b = 43,328 lb			
Full quadratic model	1722.9	3.98	0.9868
Full cubic model	933.0	2.15	0.9961
Reduced cubic model	855.1	1.97	0.9968

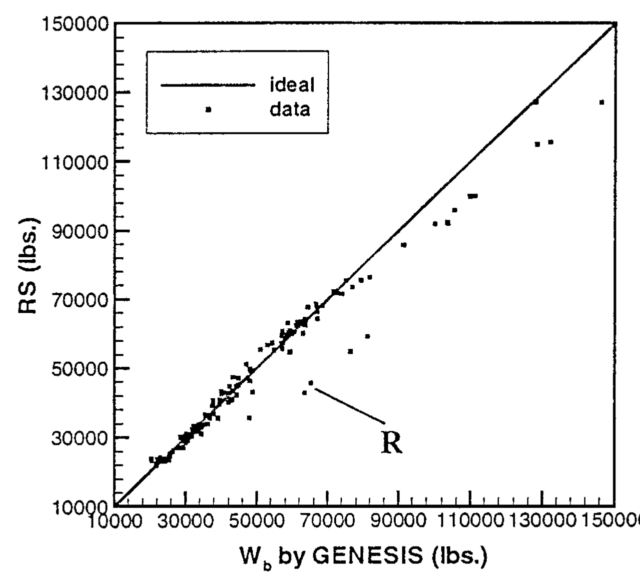
[a]Total number of outliers detected by IRLS.
[b]Number of outliers with weights above the RS.
[c]Number of outliers successfully repaired by ITRMOP = 5.
[d]W_{bd}, W_b with default parameters; W_{bh}, with ITRMOP = 5.

the approximation leaves them out. The 13 points are substantially below the diagonal line, indicating that GENESIS gives a much larger weight than that IRLS estimated. For example, the outlier marked as R in the Fig. 5a, the RS fit predicts W_b of 48,423 lb, compared to 65,185 lb from GENESIS. Comparing Figs. 5a and 5b, it is clear that the IRLS procedure moves the RS away from the outliers; for example, for outlier R, the RS now gives 45,578 lb. Figure 5c shows a comparison of the RS fitted through the repaired data. Even though the 13 points are included in the fit, they do not stand out as much after the repair. For example, for point R, the GENESIS result was reduced from 65,185 to 41,983 lb, whereas the new RS now predicts 44,606 lb. These results reflect the filtering capability of RS even when outliers exist in the data and also how repair and IRLS improve accuracy.

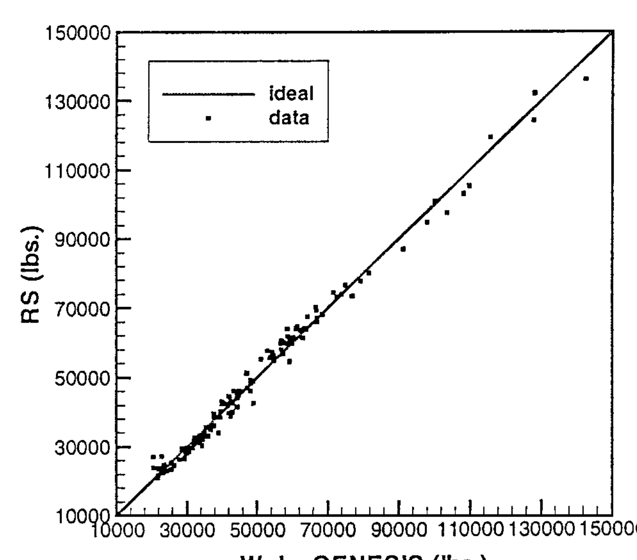
The RS fit and outlier repair reduce the noise errors whereas the use of cubic polynomials reduces the RS modeling errors. To reduce modeling errors further, we can use even higher-order polynomials, but that will require more optimizations and hence higher cost. Instead it is possible to reduce modeling errors by reducing the size of the design space. Previous studies[2,4,5] for HSCT showed that W_b for feasible or near-optimal designs is in the range of about 20,000–40,000 lb. Based on this experience, we decided to exclude designs with W_b larger than 80,000 lb from the RS. This approach of excluding design points with unreasonable results is known as the reasonable design space approach.[2] It improves the modeling accuracy of the RS by shrinking the region where the RS is fitted. After repair of the detected outliers, 11 designs were designated unreasonable, and new RS approximations were obtained by excluding them. In Table 3, without repair, the reasonable design space approach did not reduce the error; after repair, because excluding unreasonable designs reduces the modeling error, this approach helped most the quadratic approximation. We also checked on the usefulness of



a) RS approximation for original data

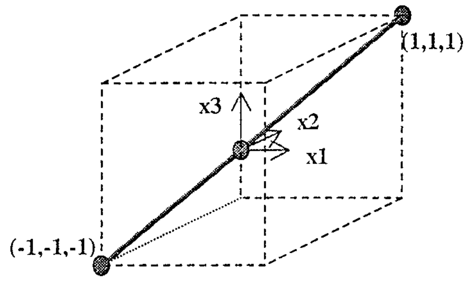


b) IRLS approximation and detected outliers

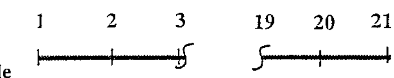


c) RS approximation for repaired data

Fig. 5 RS outliers and effect of repair.



a) Design line for three-dimensional design



variable							
1	-1.0	-0.9	-0.8	......	0.8	0.9	1.0
2	-1.0	-0.9	-0.8	......	0.8	0.9	1.0
3	-1.0	-0.9	-0.8	......	0.8	0.9	1.0
4	-1.0	-0.9	-0.8	......	0.8	0.9	1.0
5	-1.0	-0.9	-0.8	......	0.8	0.9	1.0

b) Design line for five configuration variables

Fig. 6 Extreme value design line.

conducting additional outlier detections. The results in Table 3 show significant improvement from the second round of repair, but very little from the third round.

2. *Repair vs IRLS*

To investigate how repaired RS approximation performs compared to RS leaving out the outliers, new sets of design points were created. A design line connecting the vertex of lower limits to the vertex of upper limits (where all design variables are all -1 and all $+1$, respectively) is obtained. A schematic representation of this extreme value design line is shown in Fig. 6a for a three-dimensional design box. Along the extreme value design line for the five configuration variables, 21 equally spaced points were used as shown in Fig. 6b. Another design line connecting two of the outliers from the second-round IRLS application was also generated similarly to the extreme value design line. Structural optimization results by both ITRMOP $=2$ and 5 were obtained at these 21 points of each design line as shown in Fig. 7. The minimum of the two cases for each point was considered as the true optimum. These GENESIS results were compared with the predictions by the approximations. Figure 7 indicates why repair of the outliers may be advisable rather than simply leaving them out. Removing the outliers from the data makes the approximation sensitive to the location in the design space. It worked well on the extreme value design line (Fig. 7a), but not as well as the repaired RS after repair for the other design line (Fig. 7b).

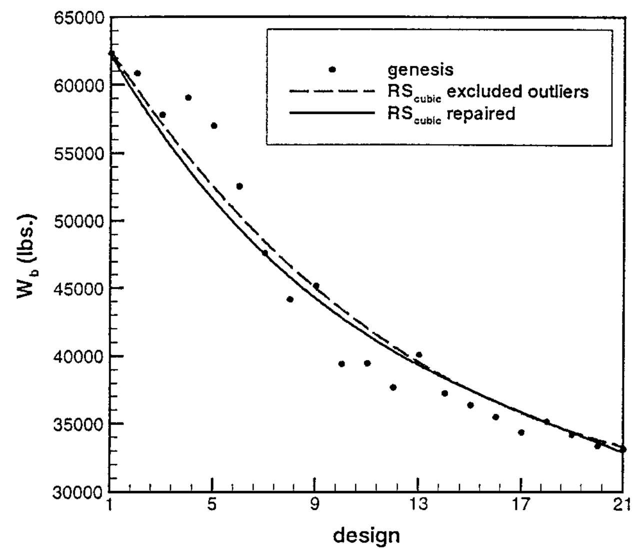
Table 4 shows a summary for average residual percentages from RS approximation constructed over the design space, but leaving out the outliers that is, IRLS, and RS approximation after outlier repair. For the 100 design points other than outliers used in constructing the approximations and for the 21 design points of the extreme value design line, the errors in both approximations are very close. The advantage of outlier repair vs the more standard elimination by IRLS is evident at the outlier points and on the design line between two outliers. On this line, the average error after repair is about 4%, but the IRLS approximation has an average error of about 10%.

3. *RS vs Data*

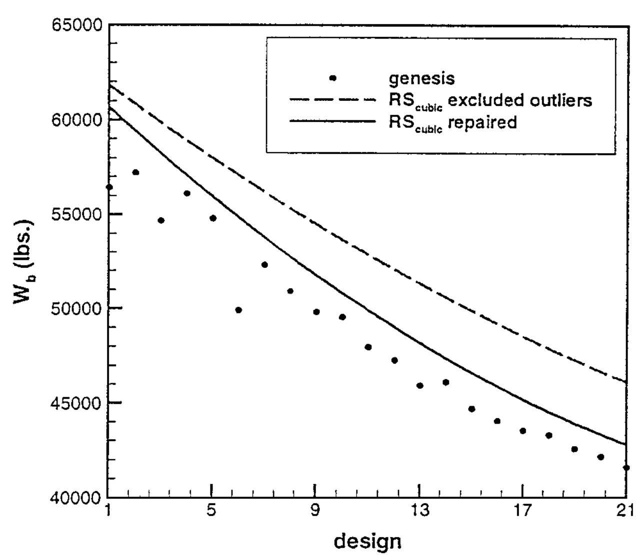
Finally, we checked for the remaining error in the data after correcting the detected outliers and how RS approximation handles that. All of the remaining design points (other than outliers) were also reoptimized by ITRMOP $=5$ for possible corrections. The last set of RS approximation summary in Table 3 presents results based on the minimum W_b from the default and high-accuracy optimizations. Even in this fully repaired data, there is still noise mainly due to the process of extracting the W_b from the structural weight as shown for the design line between two outliers in Fig. 8. This further correction for the remaining errors did not improve the quadratic approximation much, but the reduced noise helped the cubic approximations.

Table 4 Average residual percentage for IRLS fit and RS after outlier repair

	Average $\hat{e}_j$	
Design points	IRLS	RS
100 excluding outliers	1.65	1.79
15 outlier	7.12	3.72
21 on extreme value design line	3.47	3.55
21 on design line between 2 outliers	9.70	3.98



a) Extreme value design line



b) Design line connecting two outliers

Fig. 7 Comparison of RS by outlier repair and by outlier elimination over the design line (RS for 115 points; after repair 2 and reasonable design test; see Table 3).

Summarizing the results from Table 3, it appears that the most effective approach is to conduct a single round of repair, reducing $\hat{\sigma}$ to 3.4% of W_b. The RS filters out some of this error. Using Eq. (6), we found that the average prediction error at the 115 points is $0.57\hat{\sigma}$. This indicates that if the remaining error is pure noise, the RS filters out 43% of it. Altogether then, the use of the RS reduced the optimization error from about 9% to about 2%. The same error level can be achieved by optimizing every configuration with ITRMOP $=5$. However, this is more costly; in addition, without the use of the RS, we would not have known that we have a large error with the default ITRMOP $=2$.

Table 5 Statistics for 10-variable response surface approximations for wing bending material weight

RS approximation	$\hat{\sigma}$ (lb.)	% of average W_b	R_a^2
542 Points; average W_b = 42,969 lb			
Full quadratic model	4576.6	10.65	0.9667
Full cubic model	3472.2	8.08	0.9808
Reduced cubic model	3044.9	7.09	0.9853
492 Points; average W_b = 37,852 lb			
Full quadratic model	3308.6	8.74	0.9709
Full cubic model	2805.1	7.41	0.9791
Reduced cubic model	2418.5	6.39	0.9845
542 Points: after outlier repair 1 (36[a]/27[b]/23[c]); average W_b = 42,431 lb			
Full quadratic model	3310.4	7.80	0.9813
Full cubic model	2169.6	5.11	0.9919
Reduced cubic model	1895.9	4.47	0.9939
492 Points: after outlier repair 1 and reasonable design test; W_b = 37,486 lb			
Full quadratic model	2229.9	5.95	0.9860
Full cubic model	1551.2	4.14	0.9932
Reduced cubic model	1320.2	3.52	0.9951
492 Points: after outlier repair 2 (43[a]/17[b]/15[c]) and reasonable design test; W_b = 37,327 lb			
Full quadratic model	2061.0	5.52	0.9879
Full cubic model	1302.9	3.49	0.9952
Reduced cubic model	1121.8	3.01	0.9964
492 Points: after outlier repair 3 (43[a]/6[b]/1[c]) and reasonable design test; W_b = 37,320 lb			
Full quadratic model	2054.6	5.51	0.9879
Full cubic model	1285.6	3.44	0.9953
Reduced cubic model	1114.8	2.99	0.9965
492 Points: min(W_{bd}, W_{bh})[d] and reasonable design test; W_b = 36,781 lb			
Full quadratic model	2230.7	6.06	0.9854
Full cubic model	1228.0	3.34	0.9956
Reduced cubic model	1065.5	2.90	0.9967

[a]Total number of outliers detected by IRLS.
[b]Number of outliers with weights above the RS.
[c]Number of outliers successfully repaired by ITRMOP = 5.
[d]W_{bd}, W_b with default parameters; W_{bh}, with ITRMOP = 5.

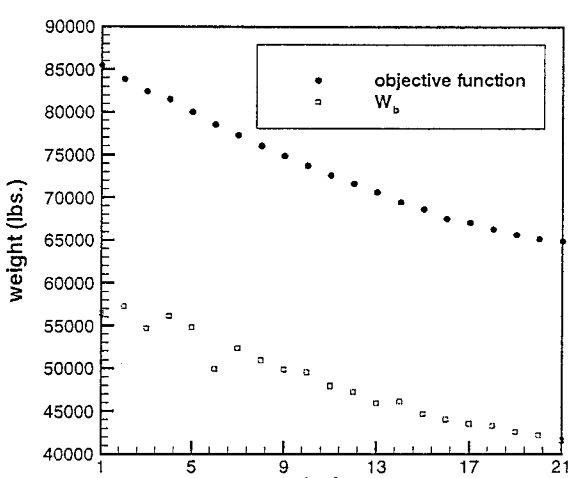


Fig. 8 Noisy W_b compared to objective function.

B. Ten Configuration Variables

Similar to the five-variable case, a FCCD was first used to represent the design space of 10 configuration variables. This design employs 1045 configurations, which is substantially more than the number of configurations needed to characterize a full quadratic model with 66 coefficients. To reduce that number, a reasonable design criterion was used as suggested and done by Balabanov et al.[2,5] The FLOPS program was used for an inexpensive estimate of the weight, and configurations with $W_b < 15{,}000$ lb or $W_b \geq 80{,}000$ lb were rejected. Also, using an approximate estimate of the range available with the maximum available fuel designs with range below 5000 n mile were rejected (the actual range requirement is 5500 n mile).

Of the original 1045 designs, 696 were found reasonable, and the D-optimality criterion in JMP[12] was used to select 292 configurations. For a cubic model approximation, the required number of data points is substantially higher (286 coefficients to be characterized). In addition, the number of required levels for a cubic is at least 4, so that even a full FCCD is inadequate. To satisfy the level requirement, an orthogonal array of 250 configurations of five levels was obtained[15] added for a total of 542 points. Full quadratic, full cubic, and reduced cubic (by mixed-mode stepwise regression) models were used to approximate the wing-bending material weight. The accuracy of the approximations is summarized in the first three rows of Table 5. Comparing the first and third rows in Tables 3 and 5, we note a somewhat higher improvement going from quadratic to cubic polynomials. That is an indication that level of the error due to noise and modeling error is closer for 10-variable case.

The steps described for the 5-variable case for outlier determination were also followed for the 10-variable case. By the use of IRLS approximation, 27 of the configurations were detected as candidate outliers. Reoptimization by ITRMOP = 5 resulted in lower W_b for 23 of the 27 designs. Table 5 presents the approximation results after the first repair. Compared to Table 3, we note that repair made less difference for the 10-variable case. This may reflect the lower percentage of outlier points. A reasonableness check based on GENESIS calculations and two more subsequent IRLS and repair applications were also performed as shown in Table 5.

Increasing the number of variables is expected to increase the importance of the modeling error. Indeed, Table 5 shows that the improvement due to the cubic terms is more pronounced than in the five-variable case (Table 3). Overall, Table 5 shows that a single round of repair reduced the optimization error from about 7 to 3.5%. The average prediction error at the 492 points is found to be $0.58\hat{\sigma}$, indicating a potential further reduction of 42% in the noise to about 2%.

V. Conclusions

The use of response surface techniques for handling noisy data was demonstrated for 5- and 10-variable examples of an HSCT. It was shown that an IRLS procedure can effectively identify outlier design points where the structural optimization produced overheavy designs due to premature convergence. Repairing the outliers was shown to be superior to a standard IRLS procedure that simply eliminates them from consideration. Elimination of unreasonable designs with high wing-bending material weights also improved the accuracy. By reducing the noise in the data, these two approaches allowed us to benefit from higher-order (cubic) models. In combination, these procedures helped to reduce the error in the optimization from about 9 to about 3%. The noise-filtering effect of the RS was estimated to reduce the error to about 2%.

Acknowledgments

This work has been supported by NASA Grants NAG1-2000 and NAG1-2177.

References

[1]Torenbeek, E., "Development and Application of a Comprehensive, Design-Sensitive Weight Prediction Method for Wing Structures of Transport Category Aircraft," Delft Univ. of Technology, Rept. LR-693, Delft, The Netherlands, Sept. 1992.

[2]Balabanov, V. O., Giunta, A. A., Golovidov, O., Grossman, B., Mason, W. H., Watson, L. T., and Haftka, R. T., "Reasonable Design Space Approach to Response Surface Approximation," *Journal of Aircraft*, Vol. 36, No. 1, 1999, pp. 308–315.

[3]McCullers, L. A., "Flight Optimization System Release 5.92 User's Guide," NASA Langley Research Center, 1997.

[4]Giunta, A. A., Dudley, J. M., Narducci, R., Grossman, B, Haftka, R. T., Mason, W. H., and Watson, L. T., "Noisy Aerodynamic Response and Smooth Approximations in HSCT Design," *Proceedings of the 5th AIAA/NASA/ISSMO Multidisciplinary Analysis and Optimization Symposium*, AIAA, Washington, DC, 1994, pp. 1117–1128.

[5]Balabanov, V., Kaufman, M., Knill, D. L., Haim, D., Golovidov, O., Giunta, A. A., Haftka, R. T., Grossman, B., Mason W. H., and Watson, L. T., "Dependence of Optimal Structural Weight on Aerodynamic Shape

for a High Speed Civil Transport," *Proceedings of the 6th AIAA/NASA/USAF Symposium on Multidisciplinary Analysis and Optimization*, AIAA, Reston, VA, 1996, pp. 599–612.

[6]Venter, G., and Haftka, R. T., "Using Response Surface Methodology in Fuzzy Set Based Design Optimization," *Proceedings of the 39th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, AIAA, Reston, VA, 1998, pp. 641–651.

[7]Venter, G., and Haftka R. T., "Response Surface Approximations for Cost Optimization of Dropped Ply Composite Laminates with Uncertainty," *Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, AIAA, Reston, VA, 1998, pp. 1145–1163.

[8]Holland, P. W., and Welsch, R. E., "Robust Regression Using Iteratively Reweighted Least-Squares," *Communications in Statistics: Theory and Methods*, Vol. 6, No. 9, 1977, pp. 813–827.

[9]Myers, R. H., and Montgomery, D. C., *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, Wiley, New York, 1995, Chaps. 1–3, 7.

[10]Khuri, A. I., and Cornell, J. A., *Response Surfaces: Designs and Analyses*, Marcel Dekker, New York, 1996, Chaps. 1–4.

[11]Beaton, A. E., and Tukey, J. W., "The Fitting of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data," *Technometrics*, Vol. 16, No. 2, 1974, pp. 147–185.

[12]JMP Statistics and Graphics and User's Guides, Ver. 3, SAS Inst., Inc., Cary, NC, Jan. 1998.

[13]Knill, D. L., Giunta, A. A., Baker, C. A., Grossman, B., Mason, W. H., Haftka, R. T., and Watson, L. T., "Response Surface Methods Combining Linear and Euler Aerodynamics for Supersonic Transport Design," *Journal of Aircraft*, Vol. 36, No. 1, 1999, pp. 75–86.

[14]GENESIS User Manual, Ver. 4.0, VMA Engineering, Colorado Springs, CO, Oct. 1997.

[15]Owen, A., "Lattice Sampling Revisited: Monte Carlo Variance of Means over Randomized Orthogonal Arrays," *Annals of Statistics*, Vol. 22, No. 2, 1994, pp. 930–945.

[16]Papila, M., and Haftka, R. T., "Uncertainty and Wing Structural Weight Approximations," *Proceedings of the 40th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, AIAA, Reston, VA, 1999, pp. 988–1002.

A. Chattopadhyay
Associate Editor