AIAA-2000-4840
A Study of the Statistical Description of Errors from Structural Optimization

H. Kim, W.H. Mason, L.T. Watson and B. Grossman
Virginia Polytechnic Institute and State University
Blacksburg, VA

and

R.T. Haftka
University of Florida
Gainesville, FL

8th AIAA/USAF/NASA/ISSMO
Symposium on
Multidisciplinary Analysis and Optimization
6-8 September 2000 / Long Beach, CA

AIAA-2000-4840

A STUDY OF THE STATISTICAL DESCRIPTION OF ERRORS FROM STRUCTURAL OPTIMIZATION

Hongman Kim*, Raphael T. Haftka†, William H. Mason‡, Layne T. Watson§, and Bernard Grossman¶

Multidisciplinary Analysis and Design (*MAD*) *Center for Advanced Vehicles*
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061-0203

Abstract

The statistical distribution of errors due to incomplete convergence of optimization procedure was studied. The optimum wing structural weights of a high speed civil transport (HSCT) calculated by GENESIS structural optimization program contained substantial error. It was possible to reduce the error by tightening the convergence criteria to achieve high fidelity optimization. To estimate the overall uncertainly level of the low fidelity optimization, the optimization error was estimated at data points constituting an experimental design in the HSCT design space. Since the error from incomplete optimization is one-sided, the Weibull model distribution was fitted to the estimated error and achieved a good fit for low and mid fidelity data. Fitting the difference of two optimizations with different convergence criteria achieved a good fit for the low fidelity error more efficiently than a direct fit on the estimated error. The fitted error distributions may provide a guide for selection of convergence criteria, and will be useful in specifying distributions of uncertainty variables in robust optimization.

* Graduate Research Assistant, Department of Aerospace and Ocean Engineering, Student Member AIAA.
† Distinguished Professor, Department of Aerospace Engineering, Mechanics and Engineering Science, University of Florida, Gainesville, FL, Fellow AIAA.
‡ Professor, Department of Aerospace and Ocean Engineering, Associate Fellow AIAA.
§ Professor, Departments of Computer Science and Mathematics.
¶ Professor and Department Head, Department of Aerospace and Ocean Engineering, Associate Fellow AIAA.

1. Introduction

Response surface techniques were originally developed for fitting experimental data. They assume that the measurements they fit contain noise that is normally distributed and has zero mean. In the past decade, response surface techniques have gained popularity for fitting the results of computer simulations. These contain numerical noise due to discretization errors and due to iterative processes that are not fully converged.[1] It is well known that the characteristics of numerical noise are different from those of experimental noise. For example, repeated numerical simulations for the same data normally give the same results, while repeated physical experiments do not. This has forced some minor changes in response surface procedures. In particular, in the choice of points used for the fitting, known as *design of experiments*, the use of duplicate points is common when physical experiments are done, and this is avoided with numerical experiments.

Optimization is typically an iterative process, and is rarely allowed to converge to high precision due to computational cost considerations. Consequently, optimization results are usually a noisy function of the parameters of the problem. The errors due to incomplete convergence are normally one sided, because the solution is not as good as the true optimum. For example, in a minimization problem, the error may be expected to be positive. This implies that the mean of the error cannot be zero.

In multidisciplinary optimization (MDO), subsystems may be optimized, and the results of the optimization are fitted by a response surface as a function of the system design variables. For example, we have followed this procedure in configuration optimization of a high speed civil transport (HSCT), where we did structural optimizations to get optimum wing structural weight (W_s) and created response surfaces of optimum wing bending material weight (W_b) as a function of the configuration variables.[2] We found that the error in the optimum structural weight depends on the convergence parameters in the optimization procedure.

In previous work[3], we identified points with large optimization error by a standard robust statistics procedure and repaired them to improve the accuracy of the response surfaces. We next modified the identification procedure to take advantage of the one-sidedness of optimization errors.[4] In this framework, the statistical properties of the optimum may be useful for guiding the process of response surface fitting. In addition, such statistical properties can be of use when the overall uncertainty in the design process needs to be studied. The objective of this work is to collect information about uncertainty in the optimization process and to study its statistical properties.

2. Fit of Statistical Distribution of Optimization Error

Choice of Statistical Distribution for Optimization Error

Our goal is to model the error incurred by convergence problems in the optimization procedures by using statistical distribution functions. We start by selecting candidate model functions to fit the distribution.[5, 6] These functions usually have a few parameters that define the distribution. For example, the normal distribution has a mean and a standard deviation as the parameters. Table 1 describes the three distributions we have considered: normal, exponential, and Weibull distributions. Due to the bias in the error incurred by incomplete convergence of optimization, the exponential or Weibull models are expected to be better than the normal distribution. The exponential distribution has only one free parameter β. The Weibull distribution has two parameters, the shape parameter α, and the scale parameter β. Note that the Weibull distribution reduces to the exponential distribution when α=1. One common application of the exponential and Weibull models is for modeling of time to failure. So, the optimization error may be interpreted as analogous to time to failure. The Weibull model can be used to fit data for either increasing or decreasing failure rate depending on the shape parameter (See Appendix A).

Fit of the Distribution of Optimization Error to the Model Functions

We use maximum likelihood estimation (MLE) to fit the distribution to optimization error data.[5] Let x be a random variable whose probability density function (PDF) $f(x; \beta)$ is characterized by a single parameter β. Assuming that the sampled data x_i (i=1,…, n) are independently and identically distributed, the probability that the sample of size n consists of values in the small intervals, $x_i \le x \le x_i+\Delta x$, is given by the product.

$$f(x_1;\beta)\Delta x f(x_2;\beta)\Delta x \cdots f(x_n;\beta)\Delta x = l(\beta)(\Delta x)^n \qquad (1)$$

The x_i 's are fixed at the sample values, and $l(\beta)$ is called the likelihood function.

$$l(\beta) = \prod_{i=1}^{n} f(x_i;\beta) \qquad (2)$$

For example, when $f(x; \beta)$ is the exponential distribution function,

$$l(\beta) = \prod_{i=1}^{n} \frac{1}{\beta}\exp(-\frac{x_i}{\beta}). \qquad (3)$$

MLE seeks the parameter β that maximizes the likelihood function, or equivalently its logarithm. For the exponential distribution

$$\log(l) = -n\log(\beta) - \frac{1}{\beta}\sum_{i=1}^{n} x_i . \qquad (4)$$

From $\partial(\log(l))/\partial(\beta) = 0$, the MLE solution can be shown to be the mean of the data. However, for more general distributions MLE may require the solution of a system of nonlinear equations.

Applying MLE to Optimization Errors

When we need to fit optimization errors with a statistical distribution, we usually do not know the exact error because we do not have the true optimum. Instead, we normally have one set of results obtained with relatively lax convergence criteria and one set of results obtained with relatively tight convergence criteria. We refer to these as the low fidelity and high fidelity data. We denote the low fidelity optimum by W_l, the high fidelity optimum by W_h, and the true optimum by W_t. Then the optimization errors for the low and high fidelity data can be represented as random variables, s and t respectively.

$$\begin{aligned} s &= W_l - W_t \\ t &= W_h - W_t \end{aligned} \qquad (5)$$

For example, if our data follows the exponential distribution, s and t have the following probability distribution functions

$$\begin{aligned} g(s) &= \frac{1}{\beta_1}\exp(-\frac{s}{\beta_1}). \\ h(t) &= \frac{1}{\beta_2}\exp(-\frac{t}{\beta_2}) \end{aligned} \qquad (6)$$

The difference between the pair of W_l and W_h is equal to the difference between the errors, because the true optimum W_t is the same. That is,

$$x = s - t = (W_l - W_t) - (W_h - W_t) = W_l - W_h. \quad (7)$$

The probability density of the differences of optimum (in short, optimization differences) can be obtained as a convolution of the joint distribution[7] of $g(s)$ and $h(t)$, provided that s and t are independent each other.

$$f(x;\beta_1,\beta_2) = \begin{cases} \int_0^\infty g(s)h(s-x)ds = \frac{1}{\beta_1+\beta_2}\exp(\frac{x}{\beta_2}) & \text{if } x<0 \\ \int_x^\infty g(s)h(s-x)ds = \frac{1}{\beta_1+\beta_2}\exp(-\frac{x}{\beta_1}) & \text{if } x\geq 0 \end{cases}. \quad (8)$$

This is a continuous distribution of optimization difference in $x \in (-\infty, \infty)$. Now we can fit the optimization difference, x, to the model function (8) via maximum likelihood estimation. In the maximum likelihood approach we find parameters β_1 and β_2 which maximize the likelihood function $l(\beta_1, \beta_2)$.

$$l(\beta_1,\beta_2) = \prod_{i=1}^{n} f(x_i;\beta_1,\beta_2) \quad (9)$$

where f is defined in Equation (8). This is an unconstrained optimization problem in two variables. Equivalently we can maximize the log likelihood.

Once we get the MLE estimates of the parameters, we have to check whether the data justifies our assumption on the distribution. The chi-square (χ^2) goodness of fit[5] is a formal test of hypothesis to check if a sample data follows a certain distribution. To compute the χ^2 test statistic, we need to subdivide the entire range of the fitted distribution into k adjacent intervals. When the sample size is n, the test statistic is calculated as

$$\chi^2 = \sum_{j=1}^{k} \frac{(N_j - np_j)^2}{np_j}. \quad (10)$$

N_j is the observed frequency in the j^{th} interval. p_j is the expected probability of the j^{th} interval from the fitted distribution and then np_j is the expected frequency in the j^{th} interval. If the fit is good, (N_j - np_j) will be small, and the test statistic χ^2 will also be small. Therefore we reject the null hypothesis that the data follows the fitted distribution if the test statistic χ^2 is large. The χ^2 test is a one-sided test and we used a 5% as the significance level. The p-value of the χ^2 test is defined as the probability that test statistic is greater than the calculated value when the fitted distribution is true. A p-value near 1 implies a good fit whereas a p-value near zero indicates a poor fit.

3. HSCT Design Optimization

Description of the HSCT Design Problem

The problem studied in this paper is a 250-passenger HSCT design with a 5500 n.mi. range and cruise Mach speed of 2.4. A general HSCT model[8] developed by the Multidisciplinary Analysis and Design (MAD) Center for Advanced Vehicles at Virginia Tech includes 29 configuration design variables. Of these, 26 describe the geometry, two the mission, and one the thrust (Table 2). The objective was to fit a response surface to the optimal wing bending material weight W_b from structural optimization as a function of aerodynamic design variables.[9] There are four load cases for the HSCT design study that represent different points on the flight envelope.

We studied a simplified version of the problem with five variables.[10] The five design variable case includes fuel weight, W_{fuel} and four wing shape parameters: root chord, c_{root}, tip chord, c_{tip}, inboard leading edge sweep angle, Λ_{ILE}, and the thickness to chord ratio for the airfoil, t/c. In this five variable case, fuselage, vertical tail, mission and thrust related parameters are kept unchanged at the baseline values. A design box is the range of design variables over which response surfaces are constructed. Table 2 shows the values and ranges of the design variables.

Structural Optimization

In this study, a structural optimization procedure based on finite element analysis using the GENESIS[11] program was applied. A finite element model developed by Balabanov[9] was used. The finite element model uses 40 design variables, including 26 variables to control skin panel thickness, 12 variables to control spar cap areas, and two for the rib cap areas. The HSCT codes calculate aerodynamic loads for each of four load cases, and a mesh generator by Balabanov calculates the FEM mesh and the applied load at the structural nodes, and creates the input for GENESIS. The structural optimization is performed for each aircraft configuration. The objective function is the total wing structural weight (W_s) and W_b is assumed to be 70% of the W_s. In previous papers, we calculated W_b by considering the skin elements that are not at minimum gauge. However, this procedure caused additional noise error besides the error due to incomplete optimization, which is our main concern in the paper. So in this work we used the objective function itself, W_s instead of W_b.

4. Results

Effects of Convergence Control Parameters on the Optimization Error

The optimization error depends on the convergence criteria, which are called control parameters in GENESIS. Control parameters in GENESIS can be categorized into move limit parameters, convergence criteria, and inner optimization control parameters. There are two loops in GENESIS. In the outer loop, an approximation for the optimization problem is generated and this approximate problem is passed to the inner loop of a gradient based optimization. After convergence of the approximate problem, a new approximation is constructed at the optimum of the approximate problem. The approximation and optimization is continued until no further change of design variables, called soft convergence, or no further change of the objective function, called hard convergence, occurs. The optimization parameters affect the performance of optimization, uncertainty in the optimization results, and the computational cost.

For the W_b response surface, we used a mixed experimental design[3] that included face centered central composite design and an orthogonal array designed to permit fitting quadratic and cubic polynomials. GENESIS structural optimizations using six different sets of parameters were performed on each of the 126 points. Table 3 lists the description of the parameters and the actual values used in this study. Case A2 employs the default parameters provided by GENESIS. Case A3 and Case A5 are the same as Case A2 except that the parameter ITRMOP was increased to 3 and 5, respectively. ITRMOP controls the convergence of the inner optimization. For the approximate optimization to converge, the inner loop convergence criterion for the objective function change must be satisfied ITRMOP consecutive times. The default value of ITRMOP is 2, and by increasing it, the inner iteration is forced to iterate further. In turn, this may force the outer loop to continue, because the so called soft convergence criteria −the change of design variables− are not met, which may have been satisfied if ITRMOP had been 2. This may result in significant improvement in the final optimum design. After extensive experimentation with the optimization parameters and help from the developers of GENESIS, we found that this parameter was the most important for improving the accuracy of the optimization for our problem. Case B2 employs tighter move limits and convergence criteria than Case A2. It reflects our first attempt to improve the optimization results. Case B3 and Case B5 are the same as Case B2 except that ITRMOP is 3 and 5, respectively. The value of all the control parameters for the six cases are given in Table 3.

In order to visualize the behavior of the errors, Figures 1 (a) and (b) show the W_s response for 21 HSCT designs along a line connecting two designs in the HSCT design space. It is clear from the figures that Case A2 and Case B2 have substantial errors. By increasing ITRMOP, the noise in W_s was substantially reduced and the gain with ITRMOP=5 over ITRMOP=3 was small. To calculate the error for each case, we need to know the true W_s, which, strictly speaking, cannot be known due to the iterative procedure inherent in the optimization. Here we estimate the true W_s by taking the best of the six GENESIS runs we already did using different parameters. Table 4 shows the performance of each set of GENESIS parameters for all 126 design points. The optimization error was calculated by comparing W_s to the best of the six W_s's available. Table 4 shows how many times each case produced the lowest, the best among the six W_s's for 126 points. Case A2 and Case B2 never found the best results. But with ITRMOP=5, Case A5 and Case B5 achieved the best W_s 58 and 52 times respectively, almost half of the data each, and together nearly 90 % of the data. Sometimes Case A3 and Case B3 came up with the best results, for 12 and 4 points, respectively.

For the default GENESIS parameter (Case A2), the mean error was 4.86 %. It is seen that tightening the convergence criteria for ITRMOP=2 (Case B2) actually had detrimental effect since the mean error increased to 5.63 %. By using ITRMOP=3 (Case A3 and Case B3), the mean errors were reduced to 0.546 % and 0.762 % respectively, about a tenth of the levels of the low fidelity errors. With Case A5 and Case B5, the mean error was very small, less than 0.2 % but this may be an underestimate because we did not perform even higher fidelity optimization. In terms of computational cost, the high fidelity optimization using ITRMOP=5 required more than twice the CPU time of the low fidelity optimization, ITRMOP=2.

In fact, the estimate of the true optimum by taking the best of six runs turned out to be very accurate, mostly thanks to the high fidelity of Case A5 and Case B5. This was tested by doing nine additional GENESIS runs on 30 randomly selected HSCT designs using Case B5 from different initial design variables. For these 30 configurations an average of 0.011% improvement was obtained. Therefore, we will denote the best of six W_s's as 'the estimate of true W_s', and accordingly the error of W_s with respect to this value will be denoted as 'estimated error', e_t, of optimization. Our concern about uncertainly will be mainly lying on the low (ITRMOP=2) and mid fidelity (ITRMOP=3) cases. For the high fidelity cases (ITRMOP=5), the error appears to be negligible.

Estimation of Statistical Distribution

Correlation between W_s and Error

Our approach is to determine the characteristics of optimization error for various optimization convergence setups through fitting a model function to the error data. The parameters of the model function are found by the maximum likelihood estimate (MLE). A basic assumption behind the MLE is that the data is independently and identically distributed. For the optimization data at hand, dependence of data may not be a problem, since the sampling was done by an experimental design without replicates. To check whether the error distribution is identical at different points, the estimated errors, e_t, were plotted against the estimated true W_s in Fig. 2.1 for the absolute error (*lb*.) and in Fig. 2.2 for the relative error (%). For example, with the absolute error of Case A2 in Fig. 2.1(a), the mean and variance of error appears to increase as W_s increases. A relatively high correlation coefficient of 0.5685 indicates the trend. Scaling the error by W_s may help to stabilize the increasing mean and variance along W_s. For the relative error of Case A2 in Fig. 2.2(a), the distribution became more even, and the correlation coefficient was reduced to 0.3210. One may observe similar effects of relative error over absolute error for other cases. The figures lead us to expect that if the structural optimization process ends with premature convergence, the absolute error would tend to be larger for heavy designs.

However, considering that highly influential points, such as the two outliers seen in Fig. 2.1 (b), tend to exaggerate the correlation coefficients, the assumption of homogeneous distribution may not be a bad assumption even for the absolute error. If the variation of distribution is important, the change of distribution may be modeled by the generalized linear model[12] (GLM). In GLM the mean of distribution is related to the regressor variables via linear model, while the fit in our approach assumes the mean of distribution is the same for different data points, which is a special case of GLM. Nonetheless, with the four distributions in Fig. 2.1, the fit of a distribution should provide useful information about the overall behavior of the error. So we fitted model distributions to both absolute and relative errors.

Direct Fit on Estimated Error

A Weibull model was fit to distribution of estimated error e_t, and the results are summarized in Table 5.1 for absolute error, and in Table 5.2 for relative error. The Weibull model has two parameters, a shape parameter, α, and a scale parameter β. According to the χ^2 test for the absolute error, the fit was good for Cases A2, B2, and B3. For Case A3, the fit was marginally rejected. For Case A5 and Case B5, the high fidelity cases with ITRMOP=5, the fits were not good at all. In fact, for those high fidelity errors, almost half of the data were zero. However, the magnitudes of the error are so small for the two high fidelity cases that the uncertainty may be negligible. The Weibull fit of the relative error was not as good as the fit for the absolute error. For instance, the p-value of the fit to the relative error of Case B2 was only 0.0327 compared to 0.8327 for the fit to the absolute error. This is unexpected since we observed that the distribution appears to be more homogeneous with relative error than with absolute error.

Recalling that the standard deviation is the same as the mean for the exponential distribution (when $\alpha = 1$ in Weibull), $\hat{\sigma}_{fit}$ greater than $\hat{\mu}_{fit}$ in Table 5.1 indicates that the errors from the structural optimization have greater scatter than the exponential model. From the fits, overall characteristics of the error can be estimated in terms of the mean ($\hat{\mu}_{fit}$), and standard deviation ($\hat{\sigma}_{fit}$) of the distribution. On the other hand, the mean and standard deviation can be estimated from e_t without fit by

$$\hat{\mu}_{data} = \bar{e}_t, \quad \hat{\sigma}_{data} = \sqrt{\frac{\sum_{i=1}^{n}((e_t)_i - \bar{e}_t)^2}{n-1}}. \tag{11}$$

In Table 5.1, $\hat{\mu}_{fit}$ and $\hat{\sigma}_{fit}$ were compared with $\hat{\mu}_{data}$ and $\hat{\sigma}_{data}$, and the agreement is good for low fidelity errors and is reasonable for the mid fidelity errors except for $\hat{\sigma}_{fit}$ of Case B3.

Tables 5.1 and 5.2 are obtained from 126 data. We expect that fewer points are required for a good fit. To find how many points are required, we calculated $\hat{\mu}_{fit}$ and $\hat{\sigma}_{fit}$ for 50 samples generated from a Weibull distribution with α = 0.7682 and β = 3915, which is the fitted value for e_t of Case B2. The mean (μ) and standard deviation (σ) of the distribution are 4572 *lbs*. and 6023 *lbs*., respectively. The mean of $\hat{\mu}_{fit}$ and 95 % confidence bounds of $\hat{\mu}_{fit}$ are drawn in Fig. 3 (a) as dash-dot lines. Similarly the mean of $\hat{\sigma}_{fit}$ and 95 % confidence bounds of $\hat{\sigma}_{fit}$ are shown in Fig. 3 (b). It appears that we need about 50 points to get a good fit with reasonable confidence bounds. Also, Fig. 3 (a) shows that $\hat{\mu}_{data}$ is very close to $\hat{\mu}_{fit}$, but Fig. 3 (b) shows that $\hat{\sigma}_{fit}$ has a narrower confidence interval than $\hat{\sigma}_{data}$.

Once we obtained a good fit, detailed information of the error distribution could be utilized. For example, we can calculate the probability that the error would exceed a certain level. In optimization for reliability, we

can use the fitted distribution in reliability calculations. For our data, it is not surprising that Weibull model fits errors from optimization failure well, since the Weibull distribution has its root in failure rate concepts.

The probability plot is a graphical tool to show the goodness of fit of distribution. We used the quantile-quantile plots[5], or so-called Q-Q plot. In the Q-Q plot, the percentiles of the data are plotted against the expected percentiles from the fit. If the fit is good, the scatter plot should appear as a straight line passing the origin with unit slope. Figure 4 shows the Q-Q plots for the Weibull fit for e_t of the low and mid fidelity optimizations. They show that the MLE resulted in a good fit for the majority of data, while more or less ignoring big error data. Figure 4 (c) indicates a particularly good fit, which is in agreement with the results of the χ^2 test. The histograms in Fig. 5 compare the shape of the error distribution to the Weibull fit. The optimization error is by definition non-negative and the probability for large error decreases rapidly to zero. One can see that the fits match well with the data for the low fidelity errors, Case A2 and Case B2, and the fits are reasonable for the mid-fidelity error cases, even for Case A3, although the fit was rejected by χ^2 test. This fact implies that the Weibull fit may be useful for Case A3 as well if we are mainly interested in the behavior of large error.

Indirect Approach Using Optimization Differences

In the previous section it was possible to estimate the low fidelity error accurately because high fidelity optimizations were available, such as Case A5 and Case B5. However, in many situations the high fidelity method may be computationally too expensive. In these situations, we may be able to fit the differences of W_s between two results from different sets of convergence parameters, to get the error distributions simultaneously. In our data, Case A3 and Case B2 represent our efforts to reduce the error in Case A2 by tightening the inner loop convergence criterion or outer loop criteria. If either optimization results by Case A3 or Case B2 are available along with Case A2, we can estimate the error distributions from the differences (See Eqs. 8 and 9). In the MLE procedure, the probability density of the difference, $f(x)$ in (8), was numerically integrated using Gaussian quadrature since there is no closed form solution for the Weibull model. Table 6 contains the correlation of e_t between the low and mid fidelity optimizations, Cases A2, A3, B2, and B3. The correlation coefficient between Cases A2 and A3 (= 0.2004), and Cases A2 and B2 (=0.1983), indicate that it may be reasonable to treat them as independent.

The difference fit was performed on the pair (A2, A3), and then on the pair of (A2, B2). The first pair consists of a low fidelity and a mid fidelity optimization while the second one engages two low fidelity results. The results of the difference fit using the Weibull model are summarized in Table 7.1. According to the χ^2 test, the fit was rejected for the pair of (A2, B3), but was good for the pair of (A2, B2). In Table 7.2, the results from the difference fit were compared to e_t data. As results of difference fit using the Weibull model, we found the distribution parameters, α and β, for both cases in a pair simultaneously. The resultant fits can be compared to the e_t data for each case and the χ^2 tests indicated that the fit was good only for Case B2.

For the first pair, bad fits may have been expected because the fit on the difference itself was poor. However, in terms of mean and standard deviation, the prediction by indirect fit for Case A2 was in a reasonable range, -9.7% and -2.1% of discrepancies for $\hat{\mu}_{fit}$ and $\hat{\sigma}_{fit}$ with respect to the $\hat{\mu}_{data}$ and $\hat{\sigma}_{data}$ based on e_t. Figures 6.1 compare the direct and difference fits for the first pair in terms of cumulative frequencies. The difference fit for Case A2 is close to that of the direct fit as seen in Fig. 6.1 (a), although for Case A3 the difference fit does not match well the direct fit. Also, it would be meaningful to compare $\hat{\mu}_{fit}$ and $\hat{\sigma}_{fit}$ from the difference fit with the estimates using the data used in the fit. For example, for the pair of Cases A2 and A3, we can calculate an approximate true W_s as the best of W_s(A2) and W_s(A3). Accordingly the approximate true error (e_a) can be calculated with respect to the approximate true W_s, and then $\hat{\mu}_{data}$ and $\hat{\sigma}_{data}$ are calculated by applying Eq. (10) to e_a. It turned out, as noted from Table 7.2, that $\hat{\mu}_{fit}$ and $\hat{\sigma}_{fit}$ are better than $\hat{\mu}_{data}$ and $\hat{\sigma}_{data}$ based on e_a for both Case A2 and Case A3. However, the estimates for Case A3 have big discrepancies with respect to $\hat{\mu}_{data}$ and $\hat{\sigma}_{data}$ based on e_t.

For the second pair, Cases A2 and B2, the indirect fit gives much better estimates of mean and standard deviation than $\hat{\mu}_{data}$ and $\hat{\sigma}_{data}$ based on e_a, because e_a is far from true error because it is based entirely on the low fidelity data of Cases A2 and B2. For example, the $\hat{\mu}_{data}$ based on e_a has a –44.1% discrepancy compared to –13.7 % of $\hat{\mu}_{fit}$ for Case A2. Figures 6.2 show the prediction by indirect fit is in reasonable match with the direct fits for the second pair. If we compare the indirect fit for Case A2 between the first and the second pair, the first pair gives a little bit better fit than the second pair as seen in Fig. 6.1(a) and Fig. 6.2(a). However, it should be pointed out that the pair of (A2, B2) is computationally cheaper than the pair of (A2, A3), because it involves another low fidelity optimization, Case B2, instead of the mid fidelity optimization of Case A3. Overall we find that performing two sets of low fidelity optimizations allows us to

estimate well the mean and standard deviation of the error.

5. Concluding Remarks

We found that the error in the GENESIS structural optimization is greatly affected by the convergence parameters used in the optimization process. The most influential convergence criterion affecting the error of structural optimization was identified. As this criterion was tightened, the uncertainty was reduced and the originally low fidelity optimization procedure was improved to yield mid or high fidelity optimization. To find the statistical properties of optimizations of various fidelity, we ran structural optimizations with different convergence parameters on a set of experimental design points, originally devised for response surfaces of optimum wing structural weight of the high speed civil transport (HSCT).

Model statistical distributions were fitted to the optimization error via the maximum likelihood estimate. Because the error comes from failure of the minimization problem, the error was non-negative and one-sided. This fact led us to try the Weibull distribution. With several optimization results, including the high fidelity runs, available for each HSCT configuration design, it was possible to estimate the true optimum wing structural weight (W_s) by taking the best of them. Consequently, optimization errors were accurately estimated comparing W_s to the estimated true optimum.

Graphical tools and correlation tests were used to check the possibility of nonhomogeneous error distribution, and there was no strong variation of the distribution to doubt the usefulness of the fit. The Weibull model gave satisfactory results for the low and mid fidelity errors according to the χ^2 test. Q-Q plots and histograms were also utilized to check the goodness of fit. The error distribution has a sharp peak near the origin and rapidly decreased as the error increased. From the mathematical background behind the Weibull distribution, it was conjectured that the optimization errors have decreasing failure rate.

Also, an indirect approach of finding the error distributions was demonstrated. Instead of trying to find the true optimum via high fidelity optimization, the differences between two optima from different convergence criteria are fitted to the joint distribution of errors. A low fidelity error was fitted along with mid fidelity or another low fidelity optimization result. The difference fit approach showed good fits for low fidelity error. Also, this approach was more efficient for estimating the statistical properties because it requires only two different fidelity optimizations, without requiring estimates of true optima.

Acknowledgement

This work was supported by NSF grant DMI-997911.

References

1. 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 5th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Panama City, FL, Sept. 7-9, 1994, pp. 1117-1128.
2. Kaufman, M., Balabanov, V., Burgee, S. L., Giunta, A. A., Grossman, B., Haftka, R. T., Mason, W. H., and Watson, L. T., "Variable-Complexity Response Surface Approximations for Wing Structural Weight in HSCT Design," *Computational Mechanics*, 18(2), 1996, pp. 112-126.
3. Papila, M., and Haftka, R. T., "Uncertainty and Wing Structural Weight Approximations," AIAA 99-1312, 40th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, St. Louis, MO, April, 1999.
4. Kim, H., Papila, M., Mason, W. H., Haftka, R. T., and Watson, L. T, Grossman, B., "Detection and Correction of Poorly Converged Optimizations by Iteratively Reweighted Least Squares," AIAA 2000-1525, 41st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Atlanta, GA, April, 2000.
5. Law, A. M., and Kelton, W. D., *Simulation Modeling and Analysis*, McGraw Hill, New York, 1982.
6. Bury, K. V., *Statistical Distributions in Engineering*, Cambridge University Press, New York, 1999.
7. Solomon, F., *Probability and Stochastic Processes*, Prentice-Hall, New Jersey, 1987.
8. Giunta, A. A., Balabanov, V., Burgee, S., Grossman, B., Haftka, R. T., Mason, W. H., and Watson, L. T., "Multidisciplinary Optimisation of a Supersonic Transport Using Design of Experiments Theory and Response Surface Modelling," *Aeronautical Journal*, Vol. 101, No. 1008, 1997, pp. 347-356.
9. Balabanov, V., 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, 1999, pp. 308-315.
10. Knill, D. L., Giunta, A. A., Baker, C. A., Grossman, B., Mason, W. H., Haftka, R. T., and Watson, L. T., "Response Surface Models Combining Linear and Euler Aerodynamics for Supersonic Transport

Design," *Journal of Aircraft*, Vol. 36, No. 1, 1999, pp. 75-86.
11. VMA Engineering, *GENESIS User Manual*, Version 5.0, Colorado Springs, CO, 1998.
12. McCullough, P., and Nelder, J. A., *Generalized Linear Models*, 2nd ed., Chapman and Hall, New York, NY, 1989.
13. Hahn, G. J., and Shapiro, S. S., *Statistical Models in Engineering*, John Wiley and Sons, Inc., New York, 1994.

Table 1: Examples of model functions for continuous distribution

	Normal	Exponential	Weibull
Density $f(x)$	$\frac{1}{\sqrt{2\pi\sigma^2}}\exp(\frac{-(x-\mu)^2}{2\sigma^2})$	$\begin{cases}\frac{1}{\beta}\exp(-\frac{x}{\beta}) & if\ x \geq 0\\ 0 & otherwise\end{cases}$	$\begin{cases}\alpha\beta^{-\alpha}x^{\alpha-1}\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right) & if\ x \geq 0\\ 0 & otherwise\end{cases}$
Parameters	Location parameter μ Scale parameter σ	Scale parameter β	Shape parameter α Scale parameter β
Mean	μ	β	$\frac{\beta}{\alpha}\Gamma(\frac{1}{\alpha})$
Variance	σ^2	β^2	$\frac{\beta^2}{\alpha}\left\{2\Gamma(\frac{2}{\alpha})-\frac{1}{\alpha}\left[\Gamma(\frac{1}{\alpha})\right]^2\right\}$

Table 2: Configuration design variables for HSCT with corresponding configuration variable ranges of the simplified five variable problem

Design Variable	**Values and Ranges**
Planform Variables	
Root chord, c_{root}	150-190 *ft*
Tip chord, c_{tip}	7-13 *ft*
Wing semi span, *b/2*	74 *ft*
Length of inboard LE, s_{ILE}	132 *ft*
Inboard LE sweep, Λ_{ILE}	67^o– 76^o
Outboard le sweep, Λ_{OLE}	25^o
Length of inboard TE, s_{ITE}	Straight TE
Inboard TE sweep, Λ_{ITE}	Straight TE
Airfoil Variables	
Location of max. thickness, $(x/c)_{max\text{-}t}$	40 %
LE radius, R_{LE}	2.5
Thickness to chord ratio at root, $(t/c)_{root}$	1.5-2.7 %
Thickness to chord ratio LE break, $(t/c)_{break}$	$(t/c)_{break}=(t/c)_{root}$
Thickness to chord ratio at tip, $(t/c)_{tip}$	$(t/c)_{tip}=(t/c)_{root}$
Fuselage Variables	
Fuselage restraint 1 location, x_{fus1}	50 *ft*
Fuselage restraint 1 radius, r_{fus1}	5.2 *ft*
Fuselage restraint 2 location, x_{fus2}	100 *ft*
Fuselage restraint 2 radius, r_{fus2}	5.7 *ft*
Fuselage restraint 3 location, x_{fus3}	200 *ft*
Fuselage restraint 3 radius, r_{fus3}	5.9 *ft*
Fuselage restraint 4 location, x_{fus4}	250 *ft*
Fuselage restraint 4 radius, r_{fus4}	5.5 *ft*
Nacelle, Mission, and Empennage Variables	
Inboard nacelle location, $y_{nacelle}$	20 *ft*
Distance between nacelles, $\Delta y_{nacelle}$	6 *ft*
Fuel weight, W_{fuel}	350000-450000 *lb*
Starting cruise altitude	65000 *ft*
Cruise climb rate	100 *ft/min*
Vertical tail area	548 ft^2
Horizontal tail area	800 ft^2
Engine thrust	39000 *lb*.

Table 3: Optimization control parameters in GENESIS for six cases

Category	Name of Parameter	Description	**Case A2**	**Case A3**	**Case A5**	**Case B2**	**Case B3**	**Case B5**
Move Limits	DELP	Fractional change allowed for properties	0.5			0.5		
	DPMIN	Minimum move limit for properties	0.1			0.1E-4		
	DELX	Fractional change allowed for design variables	0.5			0.5		
	DXMIN	Minimum move limit for design variables	0.1			0.1E-4		
	REDUC1	To multiply all the move limits by this number if internal approximate problem is NOT doing well	0.5			0.5		
	REDUC2	To divide all the move limits by this number if internal approximate problem is doing well	0.75			0.75		
Outer loop convergence on objective function (Hard convergence)	CONV1	Relative convergence criteria on objective function	0.1E-2			0.1E-3		
	CONV2	Absolute convergence criteria on objective function	0.1E-2			0.1E-3		
	GMAX	Maximum constraint violation allowed at optimum	0.1E-3			0.1E-3		
Outer loop convergence on design variables (Soft convergence)	CONVCN	Relative criteria for change in design variables	0.1E-2			0.1E-4		
	CONVDV	Relative criteria for change in properties	0.1E-2			0.1E-4		
	CONVPR	Allowable change in the maximum constraint	0.1E-2			0.1E-4		
Inner loop convergence	ITRMOP	Number of consecutive iterations that must satisfy the relative or absolute convergence criteria before optimization is terminated in the approximate optimization problem	2	3	5	2	3	5

Table 4: Performance of structural optimization for various GENESIS parameter settings

	Case A2	**Case A3**	**Case A5**	**Case B2**	**Case B3**	**Case B5**
Number of points for which the best W_s was achieved	0	12	58	0	4	52
Mean of accurate true error (percentage to the mean *Ws*)	3931.0 *lb*. (4.860%)	441.4 *lb*. (0.546%)	161.8 *lb*. (0.200%)	4552.8 *lb*. (5.629%)	616.2 *lb*. (0.762%)	32.7 *lb*. (0.040%)
CPU time per GENESIS run (seconds)	78.1	117.6	156.7	61.4	109.0	143.3

Table 5.1: Results of direct fit of the Weibull model to e_t, absolute error

	Error A2	Error A3	Error A5	Error B2	Error B3	Error B5
$\hat{\mu}_{data}$ for e_t, *lb*.	3931	441.4	161.8	4553	616.2	32.7
$\hat{\mu}_{fit}$, *lb*. (discrepancy w.r.t $\hat{\mu}_{data}$ for e_t)	3850 (-2.1 %)	462.3 (4.7 %)	147.2 (-9.0 %)	4572 (0.4 %)	689.2 (11.8 %)	146.3 (347.4 %)
$\hat{\sigma}_{data}$ for e_t, *lb*.	7071	1408	966.6	5991	1086	73.8
$\hat{\sigma}_{fit}$, *lb*. (discrepancy w.r.t $\hat{\sigma}_{data}$ for e_t)	6894 (-2.5%)	1270 (-9.8 %)	1674 (73.2 %)	6023 (0.5 %)	1682 (54.9 %)	1695 (2197 %)
α (shape parameter)	0.5912	0.4348	0.2228	0.7682	0.4703	0.2214
β (scale parameter)	2510	172.2	2.870	3915	305.8	2.724
χ^2 statistic	12.25	14.79	224.3	3.524	9.397	190.0
p-value	0.0925	0.0387	0.0000	0.8327	0.2254	0.0000

Table 5.2: Results of direct fit of the Weibull model to e_t, relative error

	Error A2	Error A3	Error A5	Error B2	Error B3	Error B5
$\hat{\mu}_{data}$ for e_t, %	3.99	0.394	0.098	6.08	0.733	0.046
$\hat{\mu}_{fit}$, % (discrepancy w.r.t $\hat{\mu}_{data}$ for e_t)	3.96 (-0.8 %)	0.468 (18.8 %)	0.173 (76.5 %)	6.10 (0.3 %)	0.768 (4.8%)	0.222 (382.6 %)
$\hat{\sigma}_{data}$ for e_t, %	6.21	0.805	0.505	7.84	1.61	0.092
$\hat{\sigma}_{fit}$, % (discrepancy w.r.t $\hat{\sigma}_{data}$ for e_t)	6.07 (-2.3 %)	1.13 (40.4%)	2.11 (317.8 %)	8.09 (3.2 %)	1.74 (8.1 %)	2.60 (2726 %)
α (shape parameter)	0.6726	0.4730	0.2175	0.7636	0.4949	0.2209
β (scale parameter)	3.007	0.2102	0.0028	5.1977	0.3767	0.0041
χ^2 statistic	16.06	24.63	228.6	15.27	9.2381	200.5
p-value	0.0245	0.0009	0.0000	0.0327	0.2360	0.0000

Table 6: Correlation coefficients between e_t, absolute error

	Case A2	Case A3	Case B2	Case B3
Case A2	1	0.2004	0.1983	0.1559
Case A3		1	0.0361	0.3967
Case B2			1	0.1811
Case B3				1

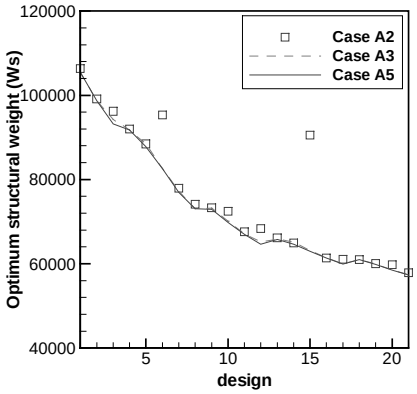
Table 7.1: Difference fit between the default (Case A2) and the cases with tightened convergence criteria

		Tightening inner loop convergence criterion	Tightening outer loop convergence criteria
Pair of W_s		$W_s(A2)$-$W_s(A3)$	$W_s(A2)$-$W_s(B2)$
Parameters of distribution of optimization error for default convergence criteria	α	0.553	0.509
	β	2105	1756
Parameters of distribution of optimization error for tightened convergence criteria	α	0.389	0.710
	β	34.3	3187
χ^2 statistic		25.1	7.33
p-value		0.0001	0.1970

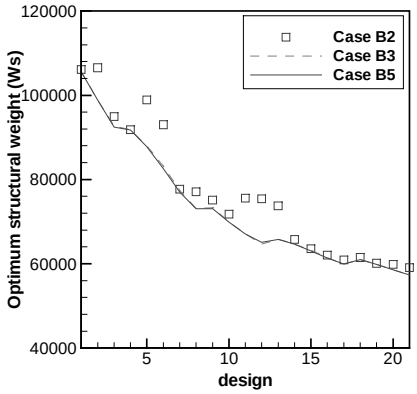
Table 7.2: Comparison of estimates of mean and standard deviation between indirect fit and estimation from data

	Tightening inner loop convergence criterion		Tightening outer loop convergence criteria	
Cases	Error A2	Error A3	Error A2	Error B2
$\hat{\mu}_{data}$ for e_t, *lb*.	3931	441.4	3931	4553
$\hat{\mu}_{fit}$ from difference fit, *lb*. (discrepancy w.r.t $\hat{\mu}_{data}$ for e_t)	3550 (-9.7%)	123.3 (-72.1%)	3394 (-13.7%)	3981 (-12.6%)
$\hat{\mu}_{data}$ for e_a, best of two, *lb*. (discrepancy w.r.t $\hat{\mu}_{data}$ for e_t)	3535 (-10.1%)	46.1 (-89.6%)	2199 (-44.1%)	2821 (-38.0%)
$\hat{\sigma}_{data}$ for e_t, *lb*.	7071.3	1407.8	7071.3	5991.0
$\hat{\sigma}_{fit}$ from difference fit, *lb*. (discrepancy w.r.t $\hat{\sigma}_{data}$ for e_t)	6922 (-2.1%)	406.2 (-71.1 %)	7393 (4.5%)	5729 (-4.4%)
$\hat{\sigma}_{data}$ for e_a, best of two, *lb*. (discrepancy w.r.t $\hat{\sigma}_{data}$ for e_t)	6895 (-2.5%)	349.6 (-75.2%)	5545 (-21.6%)	5082.9 (-15.2%)
χ^2 statistic*	26.4	89.9	17.8	7.65
p-value	0.0004	0.0000	0.0129	0.3644

* Based on values of e_t

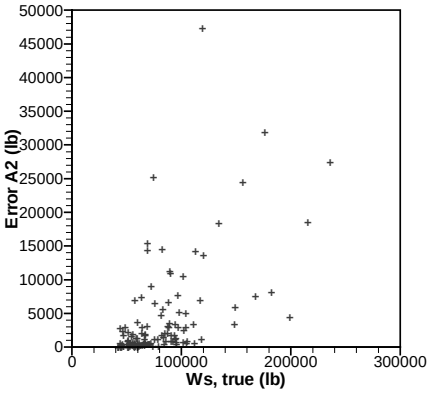


(a)

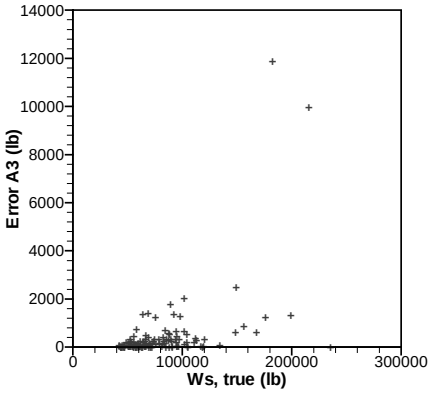


(b)

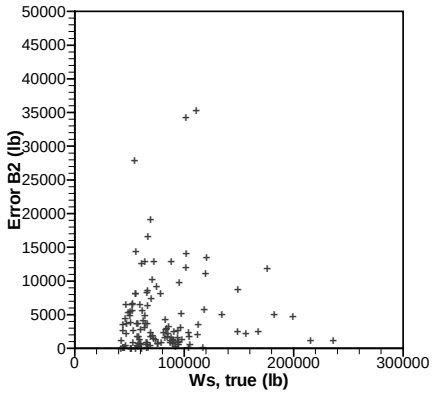
Figure 1: Optimum structural weight response along a design line for different GENESIS parameters, (a) Case A2, A3, and A5, (b) Case B2, B3, B5



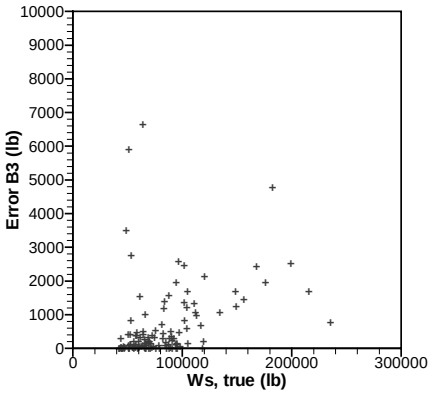
(a) Case A2 (correlation = 0.5685)



(b) Case A3 (correlation = 0.5089)

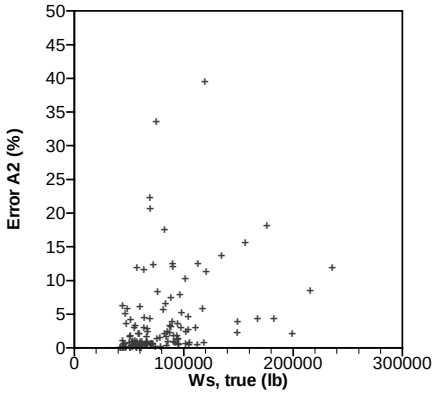


(c) Case B2 (correlation =0.0844)

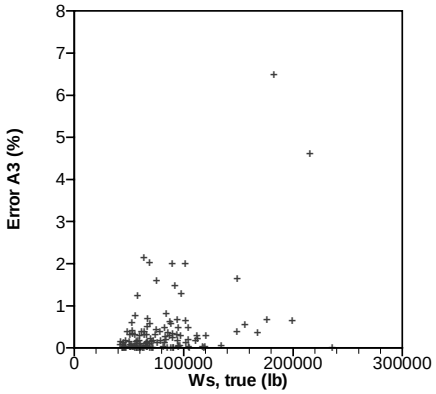


(d) Case B3 (correlation = 0.3588)

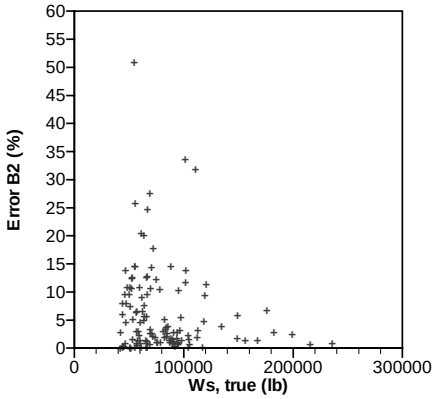
Figure 2.1: Plots of estimated error versus estimated true W_s (absolute error)



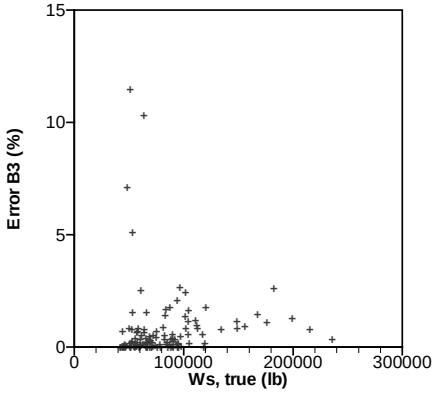
(a) Case A2 (correlation = 0.3210)



(b) Case A3 (correlation = 0.4326)

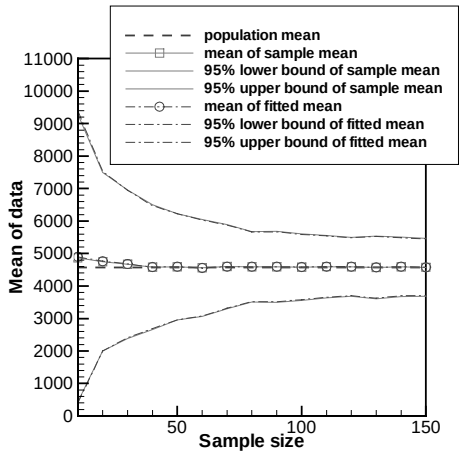


(c) Case B2 (correlation = -0.1321)

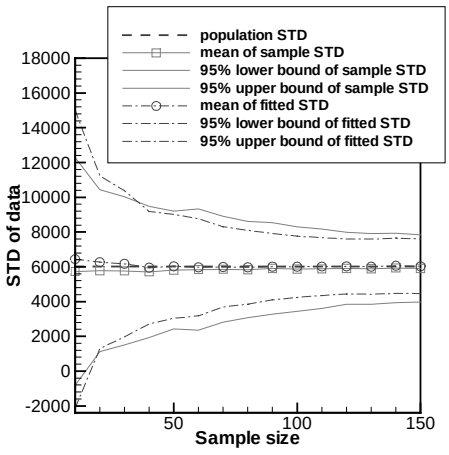


(d) Case B3 (correlation = 0.0412)

Figure 2.2: Plots of estimated error versus estimated true W_s (relative error)

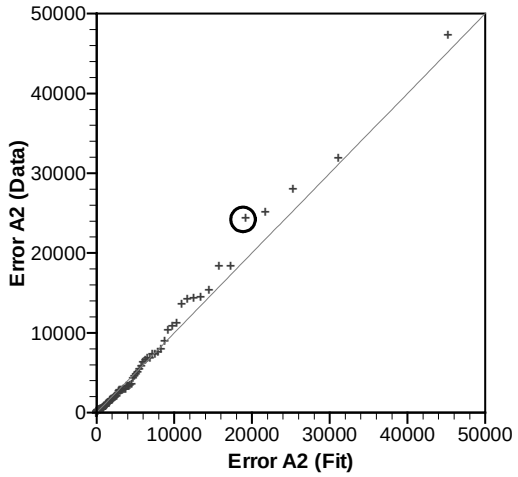


(a) $\hat{\mu}$

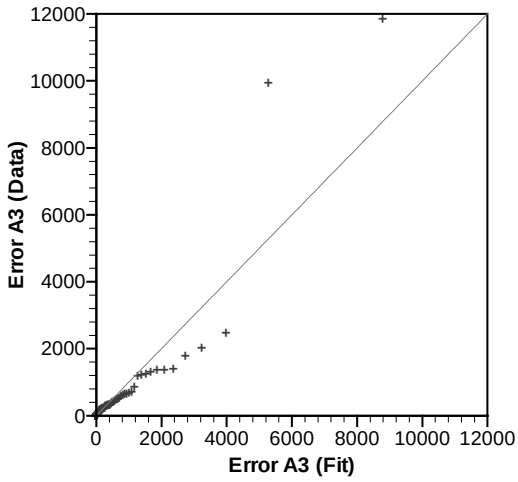


(b) $\hat{\sigma}$

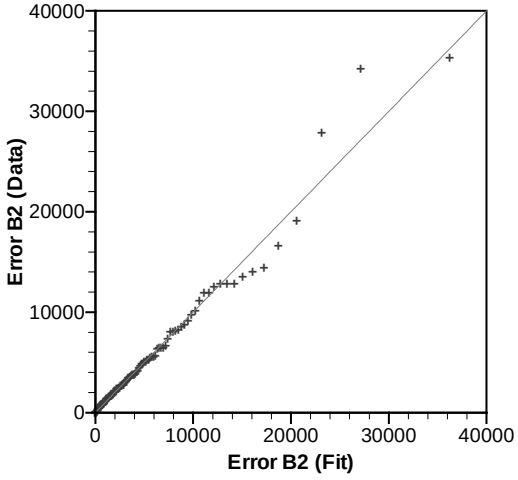
Figure 3: Comparison of $\hat{\mu}$ and $\hat{\sigma}$ of the Weibull distribution (α = 0.7682, β = 3915) between direct estimates from data and Weibull fit for various sizes of sample data



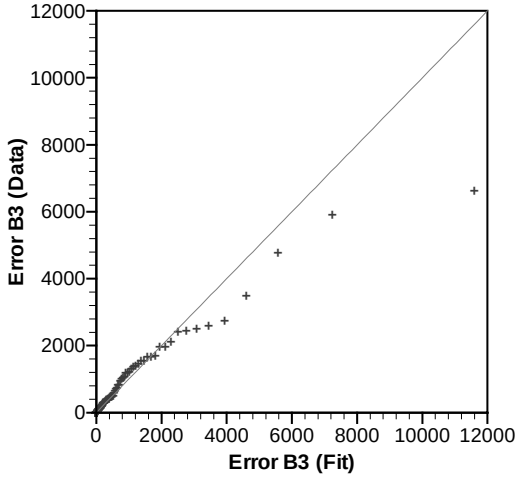
(a) Case A2



(b) Case A3

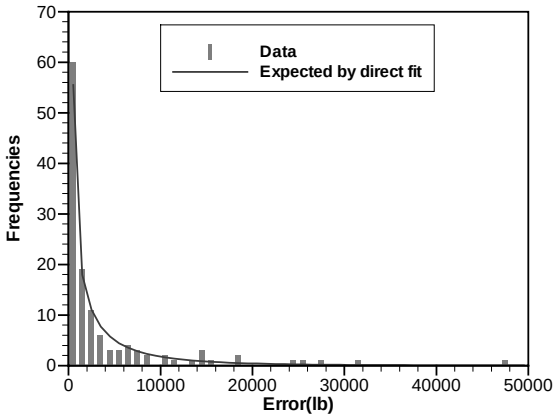


(c) Case B2

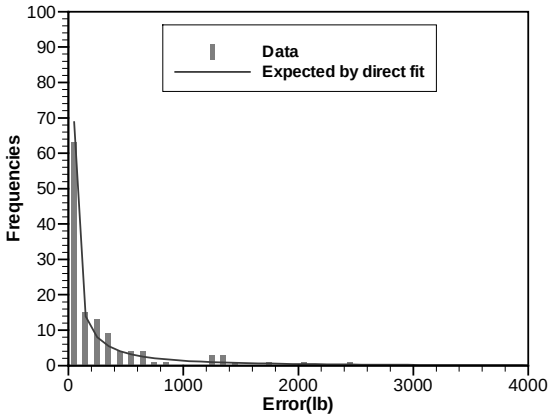


(d) Case B3

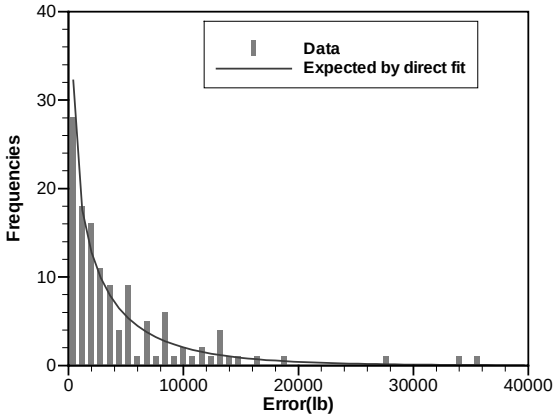
Figure 4: Q-Q plots of Weibull fit for distribution of the estimated error: points above the line indicate that the data is a distribution of a heavier tail than the fit. For example, the point marked in Case A2 indicates that the probability of having an error greater than 2000 *lb*. in the fitted distribution is equal to the probability of exceeding an error of 2400 *lb*. in the data.



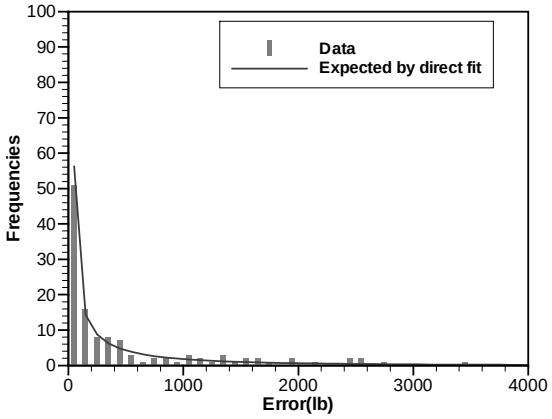
(a) Case A2



(b) Case A3

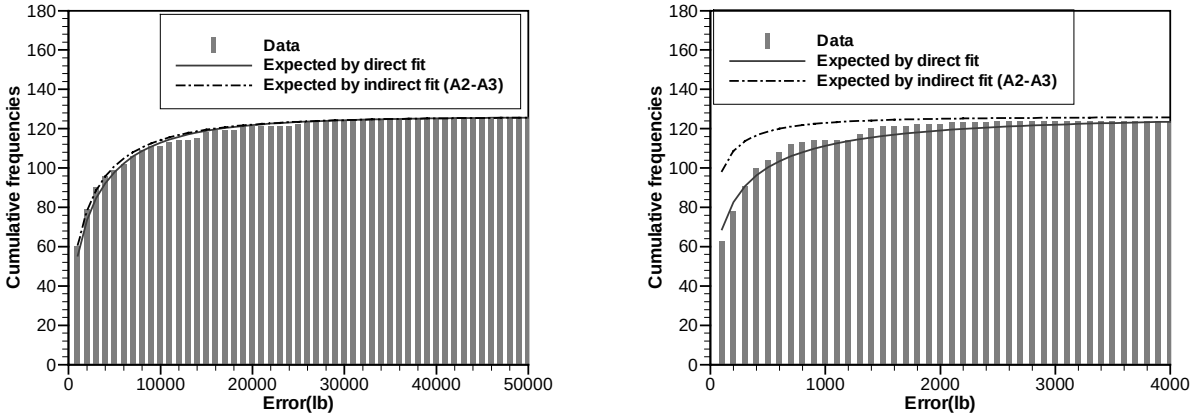


(c) Case B2



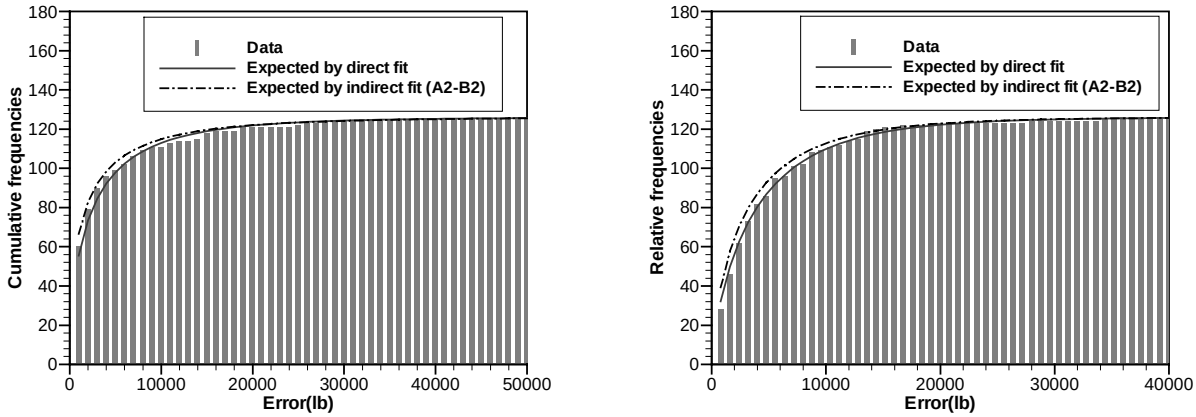
(d) Case B3

Figure 5: Comparison of histograms of e_t and direct fits of Weibull model



(a) Case A2 (b) Case A3

Figure 6.1: Comparison of cumulative frequencies between direct fit and indirect fit of Weibull model (Pair of Cases A2 and A3)



(a) Case A2 (b) Case B2

Figure 6.2: Comparison of cumulative frequencies between direct fit and indirect fit of Weibull model (Pair of Cases A2 and B2)

Appendix A. Failure rate of the Weibull distribution[13]

In the context of failure test problems such as measuring lifetime of light bulbs, it is useful to consider a function that gives the probability of failure during a very small time increment, assuming that no failure occurred before the time. This function known as the failure rate, or conditional failure function, is

$$\phi(t) = \frac{f(t)}{1 - F(t)}$$

where $f(t)$ and $F(t)$ are the probability density and distribution functions for the time to failure. Consequently, $\phi(t)dt$ can be interpreted as the conditional probability that failure will occur during the period between t and $(t+dt)$, on condition that there was no failure until time t.

For the Weibull distribution, the failure rate becomes

$$\phi(t) = \alpha\beta^{-\alpha} t^{\alpha-1}.$$

Figure A.1 shows the Weibull density functions with β =1 and various values of α. And the corresponding failure rates are shown in Fig. A.2. Note that when α = 1, failure rate is constant, and the Weibull model is reduced to the exponential distribution. For α <1, the failure rate decreases asymptotically to zero along t, but with α >1, the failure rate increases to infinity from zero. For example, a manufacturer of light bulbs may want to know the expected lifetime of their products by measuring the time to failure. Increasing failure rate implies that old light bulbs have greater chance of failure than newer ones, whereas decreasing failure rate indicates that old light bulbs are more unlikely to fail maybe because most of the problems occur at the beginning of the lifetime of light bulbs. Constant failure rate indicates that the probability of failure does not depend on how long the light bulbs have been in use.

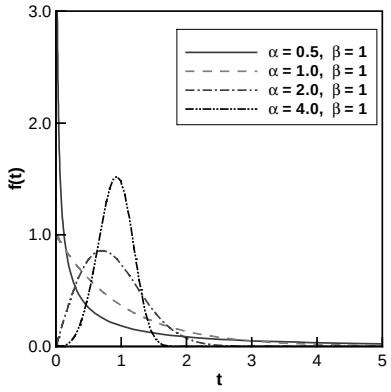


Figure A.1: Probability density functions for Weibull with α = 1 and various β

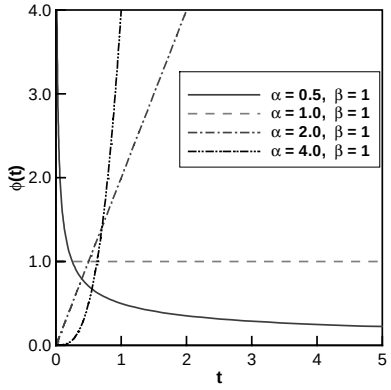


Figure A.2: Failure rate for Weibull distribution with α = 1 and various β