50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference
17th
4 - 7 May 2009, Palm Springs, California

AIAA 2009-2256

Multiple Tail Median and Bootstrap Techniques for Conservative Reliability Estimates in Design Optimization

Palaniappan Ramu[1], Nam H. Kim[2] and Raphael T. Haftka[3]
[1] *Dept. Of Aerospace & Mechanical Engineering, University of Notre Dame,NotreDame,IN 46556*
[2,3]*Dept. of Mechanical & Aerospace Engineering, University of Florida, Gainesville, FL 32611, USA*

Abstract

Reliability based design with expensive computer models becomes computationally prohibitive when reliability is estimated by sampling methods. While designing for high reliability with few samples, techniques like tail models are widely used to extrapolate reliability levels from observed levels to unobserved levels. One such approach, the multiple tail median approach uses two classical tail modeling techniques and three additional extrapolation techniques in the performance space to find reliability estimates in the unobserved levels. The method provides the median as the best estimate and the range of the five methods as an estimate of the order of the magnitude of error in median. This work explores the usage of multiple tail median approach to estimate reliability in the framework of reliability-based design. Also, bootstrap technique is employed to obtain bounds on the samples and consequently to obtain a conservative estimate of reliability.

Keywords: **Reliability, Cumulative distribution function, Tail modeling, Monte carlo simulation**

Nomenclature

D	= Tip displacement. Eq. (15)
D_0	= Allowable deflection. Eq. (15)
E	= Young's Modulus. Eq. (15)
$F_G(g)$	= CDF of *G*, Eq. (6)
$F_G(u)$	= CDF of *G* at *u*. Eq.(6)

[1] Postdoctoral Research Associate, AIAA member
[2] Associate Professor, AIAA Member
[3] Distinguished Professor, AIAA Fellow

$F_u(z)$	= Conditional CDF, Eq. (4)
$\hat{F}_{\xi,\psi}(z)$	= Approximated conditional CDF, Eq. (3)
F_X	= Load in X direction. Eq. (14)
F_Y	= Load in Y direction. Eq. (14)
G	= Performance measure. Eq. (2)
G_d	= Displacement performance measure. Eq. (15)
G_p	= $(p\times N)^{th}$ quantile of G. Eq. (7)
G_s	= Stress performance measure. Eq. (14)
g_c	= Capacity. Eq. (8)
g_r	= Response. Eq. (8)
L	= Length. Eq. (15)
N	= Total number of samples
N_{ex}	= Number of exceedances (samples in tail region)
P_i	= Empirical CDF. Eq. (10)
p	= Probability
P_f	= Failure probability. Eq. (1)
R	= Yield strength. Eq. (14)
S_r	= reciprocal of conventional safety factor, Eq. (8)
t	= thickness. Eq. (14)
u	= threshold for samples assumed to lie in tail region, Eq.(2)
w	= width. Eq. (14)
z	= exceedance, Eq. (2)
β	= Reliability index. Eq. (1)
η	= mean(error in MTM)/range of the five estimates.
ξ	= Shape parameter, Eq. (3)
σ_{comp}	= Computed stress. Eq. (14)
Φ	= Standard normal cumulative distribution function (CDF)
ψ	= Scale parameter, Eq. (3)
Beta-LT	= Fit a linear polynomial to the tail data Inverse normal cumulative distribution function applied to the CDF of S_r. Eq. (11)
Beta-QH	= Fit a quadratic polynomial to half of the data. Inverse normal cumulative distribution function applied to the CDF of S_r. Eq. (12)
GPD	= Generalized Pareto Distribution
LnBeta-QT	= Fit a Quadratic polynomial to the tail data. Logarithmic transformation applied to the beta transformed CDF. Eq. (13)
ML	= Maximum Likelihood
MTM	= Multiple Tail Median
Reg	= Regression

1. Introduction

Simulation approaches like Monte Carlo Simulation (MCS) are widely used for reliability estimation though they are computationally expensive. MCS is preferable over analytical approaches because they can address multiple failure modes; handle complex performance functions, which is unlike analytical approaches. Reliability analysis is an iterative process and using MCS is computationally prohibitive. Researchers develop variants of MCS or other approximation methods like response surface or surrogate metamodels that replace the expensive simulations.

High reliability translates to small probability of failure, determined by the tails of the statistical distributions. Since the safety levels can vary by an order of magnitude with slight modifications in the tails of the response variables (Caers and Maes [1], 1998), the tails need to be modeled accurately. Application of crude MCS is not possible because of the computational expense. Therefore, extrapolation techniques can be used. Statistical techniques from extreme value theory (referred to as classical tail modeling techniques here) are available to perform this extrapolation. The basic idea in tail modeling techniques is to approximate the conditional cumulative distribution function (CDF) above a certain threshold by the Generalized Pareto Distribution (GPD) (Castillo [2], 1988). In order to do this, one needs to estimate the parameters of GPD. There are several competing methods available for parameter estimation. This paper uses the maximum likelihood and least square regression techniques.

The multiple tail median approach (MTM) uses five different data fitting techniques. In addition to the GPD based techniques, it uses three alternate extrapolation techniques in the performance space. The first technique applies a nonlinear transformation to the CDF of the performance measure and approximates the tail of the transformed CDF using a linear polynomial fit to about top 10% of the data. The second technique approximates the upper half of the transformed CDF by a quadratic polynomial. The third technique applies a logarithmic transformation to the already transformed CDF and approximates the tail with a quadratic polynomial. It is to be noted that all five techniques do not approximate the functional expression of the model output; rather they approximate the tail of CDF. Thus, they do not need to be tailored to any functional form of the output. The MTM applies all the techniques simultaneously and use the median of the five estimates as the best estimate.

The MTM can be used for estimating reliability at target levels for each design point in a Design Of Experiment (DOE). Once the reliabilities are computed, a response surface is fitted to the estimates as a function of the design variables. This response function can further be used for constraint evaluation in a RBDO set up.

The paper is structured as follows. Classical tail modeling concepts and alternative extrapolation schemes and the proposed multiple tail median (MTM) approach is presented in

Section 2. MTM is applied to an engineering example in Section 3 and the RBDO using MTM is discussed in Section 4.

2. Classical Tail Modeling and Alternative Tail Extrapolation Schemes

The theory of tail models comprises a principle for model extrapolation based on the implementation of mathematical limits as finite level approximations. Since several advantages are reported by working in performance measure space (Ramu et al., [4], 2006), it is logical to attempt to perform tail modeling in the performance measure space to estimate quantities at unobserved levels. Reliability index and failure probability P_f are related as:

$$P_f = \Phi\left(-\beta\right) \tag{1}$$

where $\Phi(\bullet)$ is the CDF of the standard normal random variable.

In tail modeling, the interest is to address the excesses over a threshold. In these situations, the generalized pareto distribution (GPD) arises as the limiting distribution. The concept of GPD is presented in Figure 1. Let G be a performance measure which is random and u be a large threshold of G. The observations of G that exceed u are called exceedance, z, which is expressed as:

$$z = G - u \tag{2}$$

The conditional CDF $F_u(z)$ of the exceedance given that the data G is greater than the threshold u, is modeled fairly well by the GPD. Let approximation of $F_u(z)$ using GPD be denoted by $\hat{F}_{\xi,\psi}(z)$ where ξ and ψ are shape and scale parameters respectively. For a large enough u, the distribution function of $(G-u)$, conditional on $G > u$, is approximately written as (Coles [5], 2001):

$$\hat{F}_{\xi,\psi}\left(z\right) = \begin{cases} 1-\left\langle 1+\dfrac{\xi}{\psi} z \right\rangle_+^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1-\exp\left(-\dfrac{z}{\psi}\right) & \text{if } \xi = 0 \end{cases} \tag{3}$$

In Eq. (3), $\langle A \rangle_+ = \max(0, A)$ and $z > 0$. In GPD, ξ plays a key role in assessing the weight of the tail. Equation (3) can be justified as a limiting distribution as *u* increases (Coles [5], 2001, p. 75-6).

It is noted that conditional excess CDF $F_u(z)$ is related to the CDF of interest $F_G(g)$ through the following expression:

$$F_u(z) = \frac{F_G(g) - F_G(u)}{1 - F_G(u)} \tag{4}$$

From Eq. (4), the CDF of *G* can be expressed as:

$$F_G(g) = (1 - F_G(u))F_u(z) + F_G(u) \tag{5}$$

Substituting $F_u(z)$ from Eq. (2), Eq. (5) becomes:

$$F_G(g) = 1 - \left(1 - F_G(u)\right)\left\langle 1 + \frac{\xi}{\psi}(G - u)\right\rangle_+^{-\frac{1}{\xi}} \tag{6}$$

For simplicity of presentation, only the case of $\xi \neq 0$ is considered here. The shape and scale parameters can be estimated using either the maximum likelihood estimation or least-square regression method. Let *N* be the total number of samples and *p* be a probability level. Once we obtain estimates of the parameters as $\hat{\xi}$ and $\hat{\psi}$, it is possible to estimate the $(p\times N)^{\text{th}}$ quantile of *G* denoted as G_p by inverting Eq. (6):

$$\widehat{G_p} = \widehat{F^{-1}(p)} = u + \frac{\hat{\psi}}{\hat{\xi}}\left(\left(\frac{1-p}{1-F_G(u)}\right)^{-\hat{\xi}} - 1\right) \tag{7}$$

In structural applications, the performance measure is often defined as a difference between the capacity of a system g_c (e.g., allowable strength) and the response g_r (e.g., maximum stress). For the convenience of the following developments, we normalize the performance measure using the capacity. Thus, we have

$$G = \frac{g_c - g_r}{g_c} = 1 - S_r \tag{8}$$

where S_r is the reciprocal of the conventional safety factor. Failure occurs when $G > 0$, while the system is safe when $G < 0$. For the performance measure in the form of Eq. (8), we need to approximate the upper tail distribution.

The accuracy of this approach depends on the choice of the threshold value *u.* Selection of threshold is a tradeoff between bias and variance. If the threshold selected is too low, then some data points belong to the central part of the distribution and do not provide a good approximation to the tails. On the other hand, if the threshold selected is too high, the data available for the tail approximation are too few and this might lead to excessive scatter in the final estimate. Boos [7] (1984) suggests that the ratio of N_{ex} (number of tail data) over N (total number data) should be 0.02 ($50 < N < 500$) and the ratio should be 0.1 for $500 < N < 1000$. Hasofer [8], (1996) suggests using $N_{ex} = 1.5\sqrt{N}$. Here, we use the 90% quantile as the threshold.

There are several methods such as maximum likelihood (MLE) and regression to estimate the parameters, $\hat{\xi}$ and $\hat{\psi}$. MLE is based on a likelihood function, which contains the unknown distribution parameters. The values of these parameters that maximize the likelihood function are the maximum likelihood estimators.The method of least squares minimizes the sum of the deviations squared (*least square error*) from a given set of data. The parameters are obtained by solving the following minimization problem

$$\underset{\xi,\psi}{Min} \sum_{i=n}^{N} \left(F_G(g_i) - P_i \right)^2 \tag{9}$$

where P_i is the empirical CDF and $F_G(g_i)$ is the CDF of G in Eq. (6). The empirical CDF is computed as:

$$P_i = \frac{i}{N+1}, \quad i = 1, \ldots, N \tag{10}$$

where N is the total number of samples. Least square regression requires no or minimal distributional assumptions.

In addition to the previous two classical tail modeling techniques, additional tail extrapolation techniques are proposed to estimate S_r, the reciprocal of the safety factor for low failure probability that is sufficient only to estimate S_r for substantially high failure probability (low reliability index). Failure probability can be transformed to reliability index by using Eq. (1). The same transformation is applied here to the CDF of S_r. The tail of the resulting transformed CDF is approximated by a linear polynomial in order to take advantage of fact that normally distributed S_r will be linearly related to the reliability index. This is expressed as:

$$\text{Beta-LT}: S_r = C_1 + C_2\beta \tag{11}$$

Since this approximation will not be accurate enough if S_r follows distributions very different from normal, the second technique approximates the relationship between S_r and reliability index from the mean to the maximum data (about half of the sample) using a quadratic polynomial and represented as:

$$\text{Beta-QH}: S_r = C_3 + C_4\beta + C_5\beta^2 \tag{12}$$

The third technique further applies a logarithmic transformation to the reliability index of tail data that tends to linearize the tail of the transformed CDF. This tail is approximated using a quadratic polynomial, which is expressed as:

$$\text{LnBeta-QT}: S_r = C_6 + C_7(\ln(\beta)) + C_8(\ln(\beta))^2 \tag{13}$$

Here C_i, $i = 1,\ldots, 8$, are the regression coefficients. The three transformations are described with the help of Figure 2. A data set of $N = 500$ with a mean of 10 and variance 9 following a lognormal distribution is used to illustrate the three techniques. In this paper we use least square regression to find the coefficients. However, MLE approach can also be used to find the coefficients.

The alternate extrapolation techniques and classical tail modeling techniques are conceptually the same. The major difference in perceiving the two classes is that the classical tail modeling techniques model the CDF of S_r, whereas the extrapolation schemes approximates the trend of S_r in terms of reliability index. The multiple tail median (MTM) approach applies the five techniques simultaneously and uses the median of the five estimates as a compromise best estimate. It is observed that the median is a more robust estimate than the mean, because the median is less sensitive to the outliers than the mean. The MTM was demonstrated on several statistical distributions and the performance compared in Ramu et al (2008).

3. Application of multiple tail median for reliability estimation of a cantilever beam

Consider the cantilevered beam design problem, shown in Figure 3 (Wu *et al.*, [11], 2001). The objective of the original problem is to minimize the weight or, equivalently, the cross sectional area, $A = w \cdot t$ subject to two reliability constraints, which require the reliability indices for strength and deflection constraints to be larger than three. The expressions of two performance measures are given as

Strength:
$$G_s = \frac{\sigma_{comp}}{R} = \frac{\left(\frac{600}{w^2 t} F_X + \frac{600}{wt^2} F_Y\right)}{R} \tag{14}$$

Tip Displacement:
$$G_d = \frac{D_O}{D} = \frac{D_O}{\frac{4L^3}{Ewt}\sqrt{\left(\frac{F_Y}{t^2}\right)^2 + \left(\frac{F_X}{w^2}\right)^2}} \tag{15}$$

where R is the yield strength, F_X and F_Y are the horizontal and vertical loads and w and t are the design parameters. L is the length and E is the elastic modulus. R, F_X, F_Y, and E are random in nature and are defined in Table 1. It is noted that the performance measures are expressed in a fashion such that failure occurs when G_s or G_d is greater than one. In this example, we consider system failure case with both failure modes. The optimal design variables taken from Qu and Haftka [12], (2004) for a system reliability case are presented in Table 2. The value of corresponding reliability index is three. The contribution of each failure mode is also presented in Table 2. Five hundred samples are generated, and for each sample, the critical S_r (maximum of the two) is computed. The conditional CDF of S_r can be approximated by classical techniques and the relationship between S_r and reliability index can also be approximated by the three alternative extrapolation techniques and by the MTM approach. These calculations are repeated for 1,000 different samples and the errors are compared in Tables 3 and 4. The accurate estimates of S_r are calculated using MCS of sample size 1e7. From Table 3, it is observed that the Beta-LT performs the best at all seven reliability indices followed by the Beta-QH as the second best. MTM consistently performed close to the second best estimate. Table 4 shows that the MTM error is closer to the best error than to the worst error.

Figure 4 presents the box plot of η. There is a small percentage of cases when the range is not a conservative estimate of the error, but for most of the cases the range overestimates the error by factors of 2 to 10.

To show the reduction in computational requirement due to the tail extrapolations, an MCS study is performed. 100 repetitions of S_r estimates with 500,000 samples and the corresponding standard deviation are computed and presented in Table 5. At the reliability index of 4.2, the standard deviation in S_r estimate is 0.04, which is the same level with that from MTM using 500 samples. Therefore, for a same level of accuracy, the reduction in computational effort is about three orders of magnitude (500,000 to 500).

4. RBDO using MTM

RBDO is a two-stage methodology. The first stage evaluates the cost function, which is a function of design variables. The second stage estimates the reliability measure for constraint evaluation. Evolution of RBDO methods shows that researchers initially used failure probability in the constraint and because of the variation of huge orders of magnitude, they switched to reliability index. Though reliability index performed well for most problems, it had some singularity problems. Recently researchers showed that inverse measure (Qu and Haftka, 2004, Tu et al, 2000) can be successfully used for efficient RBDO. However, in the context of simulation methods to estimate these inverse measures, though the accuracy was increased and smoother convergence to optimal design was achieved, the computational expense remained high. Here, we propose to use the multiple tail model approach to estimate the inverse measure and use it in the evaluation of reliability constraint, in the RBDO framework.

One way to alleviate the expense involved in repeatedly accessing the computer models for reliability estimation is to construct response surface approximation and use them instead of the computer models. In addition, the noise problems that arise when using MCS with limited samples is also a motivation to use response surface approximations (RSA). RSA typically employ low order polynomials to approximate the inverse measures in terms of design variables to filter out noise and facilitate design optimization. These response surfaces are called design response surface (DRS) and are widely used in the RBDO (e.g. Sues et al. 1996). Here, we construct RSA of inverse measure reciprocal in terms of design variables based on a design of experiment. The inverse measure reciprocals are estimated at each design point in the DOE by Multiple Tail Model. Moreover, the principle of separable monte carlo is adopted to increase the accuracy of inverse measure reciprocal estimates.

Response Surface Approximation

MCS features several advantages such as easy implementation, robustness but large number of analyses is required to obtain a good estimate of reliability measure. In addition, it also produces noisy response and hence is difficult to use in optimization. RSA typically solve the two problems – simulation cost and noise from random sampling. However, in order to estimate the inverse measure at the design points in DOE, MCS requires large number of samples. Here, MTM approach is used to obviate the need for large number of analyses. Once the MTM estimates are obtained, response surface approximations are constructed in the design variable space, which is used in optimization.

RSA fit a closed form approximation to the limit state function to facilitate reliability analysis. RSA is very attractive because it helps avoiding the calls to expensive computer simulations like finite element analysis for response calculation. RSA usually fits low order polynomial to response in terms of random variables

$$\hat{g}(x) = Z(x)^T b \tag{16}$$

Where $\hat{g}(x)$ denotes the approximation to the limit state function $g(x)$. $Z(x)$ is the basis function vector that usually consists of monomials and b is the coefficient vector estimated by least square regression . RSA can be used in different ways. Qu and Haftka (2004) present a survey of modes in which RSA can be used as global RSA or local RSA. They construct a design response surface (DRS) of PSF in terms of design variables. At each design point in the DOE, they use 1e7 samples to estimate PSF. Here, we use 500 samples and MTM at each design point to estimate the inverse measure. Standard error metrics are used to assess the quality of response surface approximation.

RBDO of a cantilever beam

The cantilever beam example treated in previous section with a system failure mode is considered here. MTM is used to estimate inverse measures. The objective of the problem is to reduce the weight of the cantilever beam subject to different reliability levels. The reliability constraint is expressed in terms of PSF. Since PSF is tied to target failure probability, designs can be made with respect to a target failure probability or reliability index. Three different target levels are considered. MTM is used to estimate the inverse measure required for the reliability constraint evaluation. However, in order to avoid repeated calls to MTM at every design point the optimizer visits and to facilitate smooth convergence, a design response surface of inverse measure is constructed in terms of design variables width *w* and thickness *t*. This response surface is used in reliability estimation. The problem description is presented below:

$$
\begin{aligned}
&\min_{w,t} \ \text{Area} = w \times t \\
&\text{st} \quad \text{PSF}=\frac{1}{S_r} = f(w,t,X,Y,R) \geq 1
\end{aligned}
\tag{17}
$$

The range for the design response surface is presented in Table 6 is selected based on the mean based deterministic design, *w*=1.9574" and *t*=3.9149 (Qu and Haftka, 2004).

A 16 design points Latin Hypercube DOE is used. In addition, 4 samples points in the corners of the DOE to avoid extrapolation errors are also considered. At each of the 20 design points, 500 samples of the random variables X, Y and R are generated to compute the reciprocal of system safety factor. At each design point, MTM is used to estimate the safety factor reciprocal at a required target reliability index. The MTM estimates at the design points are used to construct a cubic response surface of reciprocal of system safety factor in terms of design variables w and t. A cubic polynomial with two variables has 10 coefficients. Therefore 20 design points should be sufficient to construct the response surface. Once the RSA is constructed, it is used for reliability estimation in the RBDO. The error metrics of the RSA is presented in Table 7. In order to compare the optimal designs using MTM estimate,1e7 samples are used to estimate reciprocal of safety ratio at all the design points and they are used to construct a high fidelity response surface whose error metrics are also presented in Table 7. The PRESS-RMS error for the low fidelity response surface shows that it is not as good as the high fidelity response surface. Still, it is a fairly good response surface. The optimal designs obtained using both the response surfaces are compared in Table 8. The PSF computed using 1e7 samples at designs obtained from the low fidelity response surface is also presented in Table 8. At the optimal designs obtained at different target reliability indices, MTM can be used to predict errors in the estimate. The predicted mean of error and the actual mean of error with the mean of range for a single simulation are presented in Table 9. The mean of range can be obtained as a product of MTM and from our experience; the ratio of mean of error to mean of range is between 0.15 to 0.35. Thus, the bounds of the predicted mean of the error can be computed. These numbers are compared to the actual mean of error for that particular simulation and in all the 3 cases; the actual values were between the bounds predicted. These results further strengthen our conclusion that the mean of the range can be used as a good approximation of the mean of the error.

5. Bootstrap Technique for Conservative Reliability Estimates

When only a small number of samples are available, the bootstrap method can provide an efficient way of estimating the distribution of a statistical parameter (for example, the mean of a population) using the re-sampling technique (Efron,1982 and Chernick,1999). The idea is to create many sets of bootstrap samples by re-sampling with replacement from the original data. Then, the distribution of can be approximated by the empirical distribution of the parameter, estimate of θ computed from each set of the bootstrap samples. This method only requires the initial set of samples. Each re-sampling can be performed by randomly selecting data out of the initial samples. Since the re-sampling procedure allows selecting data with replacement, the statistical properties of the re-sampled data are different from that of the original data. This approach allows us to estimate the distribution of any statistical parameter without requiring additional data.

The standard error or confidence intervals of the statistical parameter can be estimated from the bootstrap distribution. However, the bootstrap method provides only an approximation of the true distribution because it depends on the values of the initial samples. In order to obtain reliable results, it is suggested that the size of the samples must be larger than 100 [Efron, 1982].

Here, the idea is to use the bootstrap technique to obtain confidence bounds on the tail estimates. The samples beyond the threshold is resampled many times. With every resampled sample, the inverse measures are obtained for the corresponding failure probability. Repeating this n times, provides one with n estimates of inverse measure for a particular probability. Now, percentile bounds of the inverse measure for each probability in the tail can be obtained. We propose to model the region beyond the current tail data and its bounds using MTM approach. This provides one with an estimate of inverse measure and corresponding bounds for a particular probability. In section 3 it was discussed that the range of the different methods can be used to approximate the error in the MTM estimate. But this approximation is very conservative often times. Instead, here we propose to use the upper bound from the bootstrap technique as a conservative measure (dictates higher probability of failure compared to the actual failure probability) of safety.

For the example problem presented in Section 4 is used here to demonstrate the obtaining conservative estimates using bootstrap technique. Table 10 shows the results for conservative estimates. At each reliability index, the ratio of the worst and best error over MTM and MTM upper bound errors are presented. The 75% values of the ratios show that the estimated obtained from the MTM-upper bound are conservative compared to the MTM error. In term of error, the MTM estimate might be better. But the error might be in a conservative or non-conservative sense. Here, conservativeness refers to a higher estimation of inverse measure. The difference between the actual and the estimate is negative if the estimate is conservative. Table 11 shows the conservativeness of using MTM upper bound. It can be observed that the errors are negative predominantly when using MTM upper bound. This shows that the upper bound obtained from bootstrap and further tail fitted, can be used as a conservative reliability estimate in the extrapolated zone.

6. Conclusions

Multiple Tail Median (MTM) approach can be successfully used for estimating reliability at unobserved levels using limited samples. The MTM approach not only provides the user with an estimate of reliability but also magnitude of its error. These reliability estimates can be used to construct a response surface in the design space, which can be used to evaluate the constraint in RBDO. This work showed that RBDO can be successfully performed using MTM. Also, it was shown that bootstrap techniques can be used to obtain conservative estimate of reliability.

References

[1] Caers J, Maes M. Identifying tails, bounds, and end-points of random variables. Structural Safety 1998; 20: 1-23.

[2] Castillo E. Extreme value theory in engineering. San Diego, USA: Academic Press; 1988.

[3] Goel T, Haftka RT, Shyy W, Queipo NV. Ensemble of multiple surrogates. Structural and Multidisciplinary Optimization 2006; 33(3).

[4] Ramu P, Qu X, Youn BD, Haftka RT, Choi KK. Safety factors and inverse reliability measures. Int. J. reliability and safety 2006; 1: 187-05

[5] Coles S. An introduction to statistical modeling of extreme values. London, England: Springer-Verlag; 2001.

[6] McNeil AJ, Saladin T. The peaks over thresholds method for estimating high quantiles of loss distributions. In: Proceeding of XXVIIth international ASTIN Colloquim, Cairns, Australia, 1997; 23-43

[7] Boos D. Using extreme value theory to estimate large percentiles. Technometrics 1984; 26 (1): 33-39.

[8] Hasofer A. Non-parametric estimation of failure probabilities. In: Casciati F, Roberts B, editors. Mathematical models for structural reliability. CRC Press, Boca Raton, Florida, USA, p. 195-226.

[9] Beirlant J, Vynckier P, Teugels JL. Excess functions and estimation of extreme value index. Bernoulli, 1996; 2: 293-18.

[10] Bassi F, Embrechts P, Kafetzaki M. Risk management and quantile estimation. In: Adler RJ, et al., editors. A practical guide to heavy tails. Boston, Birkhaeuser, p. 111-30.

[11] Wu YT, Shin Y, Sues R, Cesare M. Safety factor based approach for probability–based design optimization. In: Proceedings of 42nd AIAA/ ASME/ ASCE/AHS/ASC structures, structural dynamics and materials conference, Seattle, WA, 2001. AIAA Paper 2001-1522.

[12] Qu X, Haftka RT. Reliability-based design optimization using probability sufficiency factor. Structural and Multidisciplinary Optimization 2004; 27(5), 314-25.

[13] Qu X, Haftka RT, Venkataraman S, Johnson TF. Deterministic and reliability-based design optimization of composite laminates for propellant tanks. AIAA Journal 2003; 41(10), 2029-36.

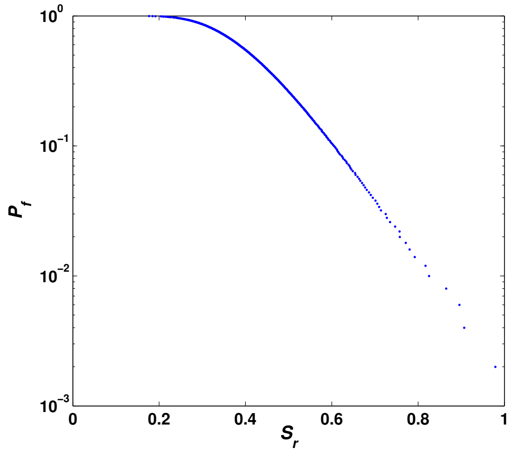
[14] P. Ramu, N. H. Kim, and R. T. Haftka, Multiple Tail Median Approach for High Reliability Estimation, Structural Safety, submitted, 2008

[15] Tu, J.; Choi, K. K.; Park, Y. H. 1999: A New Study on Reliability Based Design Optimization, *Journal of Mechanical Design,* ASME, Vol 121, No. 4, 557-564

[16] Efron, B., 1982, The Jackknife, the Bootstrap, and other Resampling Plans, SIAM, Philadelphia

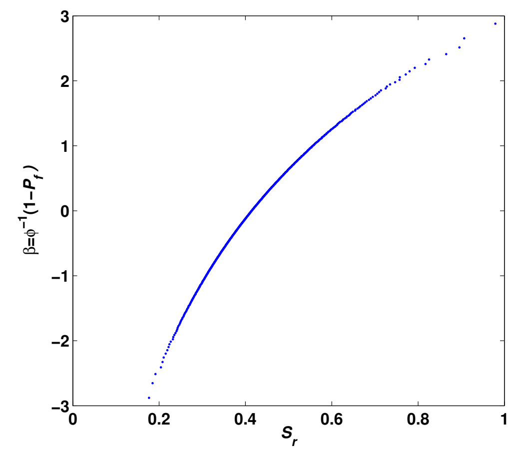
[17] Chernick, M.R., 1999, Bootstrap Methods, a Practitioner's guide, (Wiley series in probability and statistics) A Wiley-Interscience publication

Figures



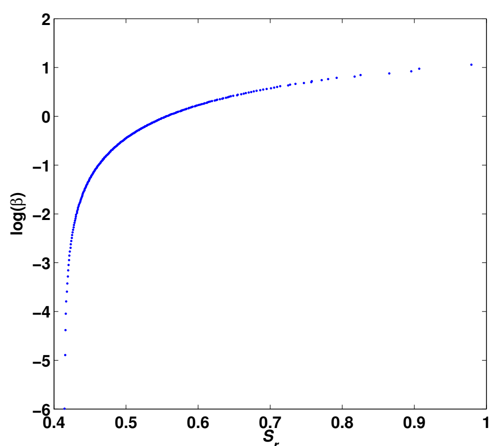
10^0
10^{-1}
10^{-2}
10^{-3}
P_f
0
0.2
0.4
0.6
0.8
1
S_r

(a)



3
2
1
0
−1
−2
−3
$\beta=\phi^{-1}(1-P_f)$
0
0.2
0.4
0.6
0.8
1
S_r

(b)



2
1
0
−1
−2
−3
−4
−5
−6
log(β)
0.4
0.5
0.6
0.7
0.8
0.9
1
S_r

(c)

Fig. 2. Transformation of the CDF of safety factor reciprocal (S_r). (a) CDF of S_r. (b) Inverse standard normal cumulative distribution function applied to the CDF (c) Logarithmic transformation applied to the reliability index.

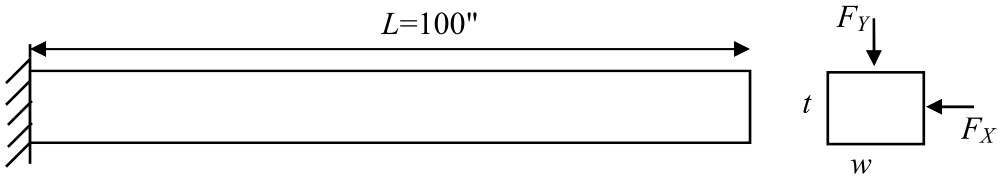


Fig. 3. Cantilever beam subjected to horizontal and vertical loads

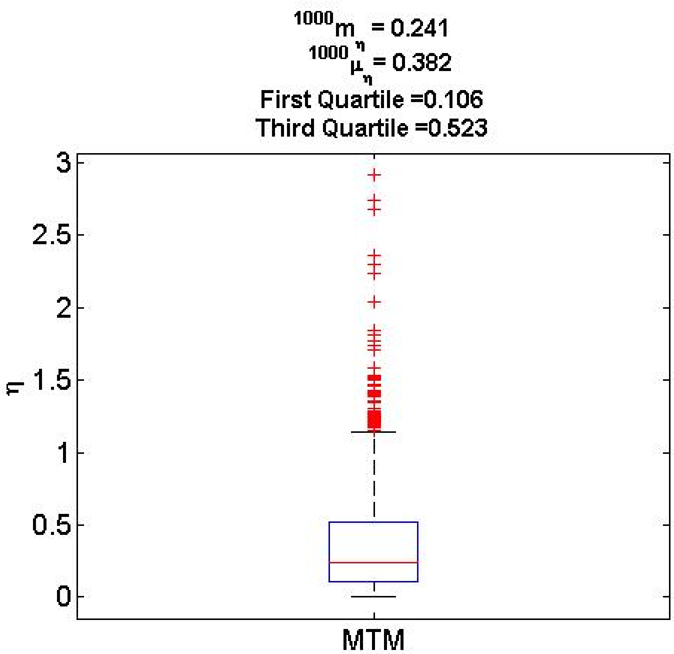


Fig. 4. Cantilever beam. Boxplot of η.

Tables

Table 1. Random variables for the cantilevered beam problem

Random Variable	F_X	F_Y	R	E
Distribution	Normal (500,100)lb	Normal (1000,100)lb	Normal (40000,2000) psi	Normal (29E6,1.45E6) psi

Table 2. Properties of the cantilever beam

L	100"
w	2.6041
t	3.6746
D_o	2.145
P_{f1}	0.00099
P_{f2}	0.00117
$P_{f1} \cap P_{f2}$	0.00016

P_{f1} - Failure probability in mode1. P_{f2} - Failure probability in mode 2.

Table 3. Cantilever Beam. Summary of errors in S_r for different techniques at different reliability indices (mean and median over 1000 repetitions of 500 samples)

Rel Index		3	3.2	3.4	3.6	3.8	4	4.2
GP-MLE	Mean	0.026	0.034	0.044	0.056	0.070	0.086	0.106
	Median	0.022	0.030	0.039	0.051	0.063	0.075	0.093
GP-Reg	Mean	0.056	0.080	0.113	0.159	0.225	0.323	0.478
	Median	0.036	0.048	0.062	0.078	0.095	0.110	0.132
LnBeta-QT	Mean	0.027	0.033	0.038	0.044	0.050	0.056	0.062
	Median	0.022	0.026	0.031	0.036	0.041	0.046	0.052
Beta-LT	Mean	0.019	0.021	0.024	0.026	0.028	0.031	0.034
	Median	0.016	0.018	0.020	0.022	0.024	0.027	0.030
Beta-QH	Mean	0.022	0.026	0.030	0.035	0.040	0.045	0.051
	Median	0.019	0.022	0.026	0.031	0.035	0.039	0.044
MTM	Mean	0.022	0.027	0.032	0.037	0.043	0.048	0.055
	Median	0.018	0.022	0.026	0.029	0.033	0.038	0.044

Table 4. Cantilever Beam. Summary of ratios of lowest and highest errors to MTM error (mean and median over 1000 repetitions of 500 samples)

Rel Index	Lowest Error/MTM Error				Highest Error/MTM Error			
	25%ile	Mean	Median	75%ile	25%ile	Mean	Median	75%ile
3	0.28	0.55	0.58	0.81	1.67	10.02	2.44	5.36
3.2	0.27	0.53	0.55	0.78	1.77	16.02	2.79	6.14

3.4	0.26	0.52	0.53	0.76	1.84	11.48	3.01	7.47
3.6	0.27	0.52	0.51	0.76	1.92	16.89	3.26	7.76
3.8	0.26	0.52	0.52	0.77	1.96	15.00	3.53	8.77
4	0.26	0.53	0.53	0.77	2.03	37.08	3.72	9.82
4.2	0.26	0.51	0.50	0.75	2.08	26.74	3.83	9.93

Table 5. S_r *estimates (without tail extrapolation) and standard deviation* at different reliability indices.

Rel Index	3	3.2	3.4	3.6	3.8	4	4.2
S_r	1.012	1.032	1.05	1.07	1.09	1.12	1.13
SD	0.003	0.004	0.01	0.01	0.01	0.02	0.04

*Mean of 100 repetitions of 5e5 samples each

Table 6. Range of design variables for the design response surface

System Variables	W	t
Range	1.5” to 3.0”	3.5” to 5.0”

Table 7. Error metrics for the high fidelity and low fidelity response surface

Samples	High Fidelity:1e7			Low Fidelity: 500		
Rel Index	3	3.6	4.2	3	3.6	4.2
RMS	0.02	0.03	0.04	0.03	0.07	0.09
RMS-Pre	0.02	0.04	0.05	0.04	0.09	0.14
Rsq	0.99	0.99	0.99	0.99	0.99	0.99
Rsq-Adj	0.99	0.99	0.99	0.99	0.99	0.99
Press-RMS	0.05	0.11	0.08	0.08	0.15	0.29
Rsq-Pred	0.99	0.99	0.99	0.99	0.98	0.95

Table 8. Optimal designs using high fidelity and low fidelity response surface

Samples	High Fidelity:1e7			Low Fidelity: 500			PSF*
Rel Index	w	t	A	w	t	A	
3	2.578	3.756	9.684	2.7498	3.5	9.6243	1.0008
3.6	2.611	3.765	9.831	2.6213	3.8083	9.9829	0.9976
4.2	2.949	3.5	10.325	2.9796	3.5	10.429	1.0112

* Estimates with 1e7 samples at optimal designs obtained using low fidelity response surface

Table 9. Bounds on the predicted mean of error using the mean of the range

Rel Index	3	3.6	4.2
Mean(Range)	0.117	0.109	0.356
Pred Mean(Err)	0.017-0.041	0.016-0.038	0.05-0.12
Act –Mean(Err)	0.026	0.023	0.109

Table 10. Cantilever Beam. Summary of Lowest and Largest error to MTM and MTM – Upper bound estimates

Rel Index	Lowest Error/MTM Error				Lowest Error/MTM-Upper Bound Error			
	25%ile	Mean	Median	75%ile	25%ile	Mean	Median	75%ile
3	0.27	0.53	0.52	0.77	0.09	0.25	0.85	0.53
3.2	0.24	0.50	0.51	0.75	0.11	0.25	7.01	0.54
3.4	0.24	0.49	0.50	0.72	0.11	0.28	0.59	0.55
3.6	0.25	0.50	0.50	0.71	0.11	0.27	2.34	0.57
3.8	0.24	0.51	0.51	0.73	0.14	0.30	0.65	0.56
4	0.26	0.50	0.52	0.78	0.14	0.29	0.65	0.56
4.2	0.26	0.49	0.51	0.76	0.13	0.30	0.76	0.58
	Highest Error/MTM Error				Highest Error/MTM-Upper Bound Error			
3	1.67	2.45	10.16	4.83	0.86	1.72	4.79	3.63
3.2	1.77	2.63	11.55	5.57	0.95	1.91	16.31	4.26
3.4	1.86	2.86	30.16	6.19	1.13	2.02	5.14	4.71
3.6	1.91	2.97	15.84	6.74	1.27	2.18	145.70	5.00
3.8	1.97	3.15	18.71	7.06	1.36	2.21	57.57	5.38
4	2.03	3.27	18.83	7.96	1.41	2.34	8.12	5.88
4.2	2.05	3.24	18.53	9.04	1.47	2.36	23.80	6.01

Table 11. Cantilever Beam. Summary of conservativeness using MTM and MTM upper bound estimates

	Error using MTM estimate				Error using MTM-Upper Bound Estimate			
	25%ile	Mean	Median	75%ile	25%ile	Mean	Median	75%ile
3	-0.022	-0.002	-0.004	0.015	-0.049	-0.022	-0.025	0.001
3.2	-0.024	0.001	-0.004	0.018	-0.053	-0.021	-0.025	0.008
3.4	-0.026	0.003	-0.002	0.025	-0.055	-0.019	-0.022	0.016
3.6	-0.026	0.006	0.000	0.031	-0.057	-0.017	-0.019	0.025
3.8	-0.026	0.008	0.002	0.038	-0.060	-0.013	-0.016	0.036
4	-0.029	0.009	0.003	0.045	-0.063	-0.011	-0.014	0.046
4.2	-0.028	0.015	0.009	0.057	-0.061	-0.004	-0.006	0.061