The American Society of Mechanical Engineers

Reprinted From
DE-Vol. 55, Reliability, Stress Analysis, and Failure Prevention
Editor: Richard J. Schaller
Book No. G00816 - 1993

PARAMETER ESTIMATION TECHNIQUES BASED ON OPTIMIZING GOODNESS-OF-FIT STATISTICS FOR STRUCTURAL RELIABILITY

Alois Starlinger[1]
Structural Integrity Br.
NASA Lewis Research Center
Cleveland, Ohio

Stephen F. Duffy and Joseph L. Palko
Dept. of Civil Engineering
Cleveland State University
Cleveland, Ohio

ABSTRACT

New methods are presented that utilize the optimization of goodness-of-fit statistics in order to estimate Weibull parameters from failure data. It is assumed that the underlying population is characterized by a three-parameter Weibull distribution. Goodness-of-fit tests are based on the empirical distribution function (EDF). The EDF is a step function, calculated using failure data, and represents an approximation of the cumulative distribution function for the underlying population. Statistics (such as the Kolmogorov-Smirnov statistic and the Anderson-Darling statistic) measure the discrepancy between the EDF and the cumulative distribution function (CDF). These statistics are minimized with respect to the three Weibull parameters. Due to nonlinearities encountered in the minimization process, Powell's numerical optimization procedure is applied to obtain the optimum value of the EDF. Numerical examples show the applicability of these new estimation methods. The results are compared to the estimates obtained with Cooper's nonlinear regression algorithm.

INTRODUCTION

Typically, structural analysis techniques used to estimate the reliability of components fabricated from ceramic material systems (see Thomas and Wetherhold, 1991, and Palko et al., 1993) assume that the random strength parameters are characterized by a Weibull probability density function (PDF). This broad assumption, i.e., the use of a Weibull distribution as opposed to the use of other distributions such as a log-normal probability distribution necessitates some reflection. A wealth of experience indicates the Weibull distribution works well for monolithic ceramics. In fact, as Tracy et al. (1982) point out, if a structural component can be adequately modeled as a weakest link (or series) system, then the PDF of choice is the Weibull distribution. Alternatively, for parallel systems the log-normal distribution is appropriate. The structural analysis community has for the most part adopted the viewpoint (based on supporting experimental evidence) that monolithic ceramics behave in a weakest link fashion. However, very little failure data exists for laminated ceramic matrix composite (CMC) material systems, and definitely not enough to justify the use of a specific probability density function.

Accepting the use of a Weibull distribution for monolithic ceramics, the authors point out that several researchers (Margetson and Cooper, 1984 Duffy et al., 1993 and Foley et al., 1993) have presented data indicating certain monolithic ceramics exhibit threshold behavior. In addition, a threshold in the fiber direction of ceramic composites is intuitively plausible. The existence of a threshold for any type of ceramic material system should be approached with an open mind until a sufficient data base is assembled for a specific material system. If a material clearly exhibits zero-threshold behavior, and the underlying population can be characterized by the Weibull distribution, the very robust two parameter maximum

[1] National Research Council; Currently with Airex Composite Engineering

likelihood estimation algorithm is recommended (see ASTM Standard Practice C-1239). Alternatively, if the failure data suggests a threshold, then the estimation techniques presented here may apply.

In general, the objective of parameter estimation is to derive functions (or estimators) that yield, in some sense, optimized values of the underlying population parameters. Here the functional value of an estimator is a point estimate (in contrast to an interval estimate) of the true population parameter. The estimated values must be dependent on failure data. The values of point estimates computed from a number of samples obtained from a single population will vary from sample to sample. Thus the estimates can also be considered statistics. A sample is a collection (i.e., more than one) of observations taken from a well defined population, where a population represents the totality of observations about which statistical inferences are made. Here, the observations are the failure strengths of test specimens fabricated from ceramic material systems (where the system may be monolithic or composite).

As Stephens (1986) points out, the empirical distribution function (EDF) was originally developed as an aid in measuring the performance of a given parameter estimation technique. Statistics directly related to the EDF are appropriately referred to as goodness-of-fit statistics. In this article, goodness-of-fit statistics are utilized in directly computing parameter estimates, instead of the more traditional role of quantifying the performance of an estimator. Methods are proposed where parameter estimates are obtained by optimizing EDF statistics. Specifically, the first parameter estimation method minimizes the Kolmogorov-Smirnov goodness -of-fit statistic (D). A second estimation method that minimizes the Anderson-Darling goodness-of-fit statistic estimator (A^2) is also presented. The effectiveness of these estimation methods are studied by comparing results with the least-squares method originally developed by Cooper (1988), and later modified by Duffy et al. (1993).

GOODNESS-OF-FIT STATISTICS

The EDF is a step function, denoted here as $F_N(x)$, that is dependent on the number and individual values of failure observations within a sample. The function serves as an approximation of the cumulative distribution function for the underlying population. Statistics associated with the EDF, such as the Kolmogorov-Smirnov statistic and the Anderson-Darling statistic are measures of the discrepancy between the EDF and the cumulative distribution function (CDF), which is identified as $F(x)$. Thus a decision regarding the type of CDF (or PDF) must be made a priori in order to calculate either EDF statistic. Traditionally, the EDF statistics have been employed to assess the relative merits in choosing a particular CDF. Focusing attention on the Weibull PDF, the three parameter function has the form

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x-\gamma}{\beta}\right)^{(\alpha-1)} exp\left\{-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right\} \qquad (1)$$

for a continuous random variable x, when $0 \le \gamma \le x$, and

$$f(x) = 0 \qquad (2)$$

for $x \le \gamma$. The Weibull CDF is given by the expression

$$F(x) = 1 - exp\left\{-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right\} \qquad (3)$$

for $x > \gamma$, and

$$F(x) = 0 \qquad (4)$$

for $x \le \gamma$. Here α is the Weibull modulus or shape parameter, β is the material scale parameter, and γ is the threshold parameter. β can be described as the Weibull characteristic strength of a specimen with unit volume loaded in uniform uniaxial tension. The parameter β has units of stress * (volume)$^{(1/\alpha)}$, α is dimensionless, and γhas the units of stress. The estimates for α and β are restricted to non-negative values, and estimates of γ are restricted to non-negative values.

The first goodness-of-fit statistic discussed is the Kolmogorov-Smirnov (KS) statistic. This goodness-of-fit statistic (denoted as D) belongs to the supremum class of EDF statistics, and is defined as

$$D = sup\ |F_N(x) - F(x)| \qquad (5)$$

$$= max\ (D^+, D^-)$$

where

$$D^+ = sup\ \{F_N(x) - F(x)\} \qquad (6)$$

$$D^- = sup\ \{F(x) - F_N(x)\} \qquad (7)$$

Here D is a measure of the largest difference (i.e., the supremum) in functional value between the EDF and the

CDF. To facilitate computations, notation adopted by Stephens is followed where

$$Z_i = F(x_i) \tag{8}$$

is used to denote the value of the CDF for an individual failure datum, x_i. By arranging the Z_i values in ascending order such that

$$Z_1 < Z_2 < \cdots < Z_N \tag{9}$$

where N is the number of specimens in a sample, suitable formulas for the KS statistic D^+ and D^- can be derived using Z_i, i.e.,

$$D^+ = \max_i \left\{ \frac{i}{N - Z_i} \right\} \quad for\ 1 \le i \le N \tag{10}$$

$$D^- = \max_i \left\{ Z_i - \frac{i-1}{N} \right\} \quad for\ 1 \le i \le N \tag{11}$$

When applying the concepts above to strength data of ceramic materials, insertion of Eq. 3 into Eq. 8 yields

$$Z_i = 1 - exp\left[- \left(\frac{\sigma_i - \gamma}{\beta} \right)^{\alpha} \right] \tag{12}$$

Here σ_i (which replaces x_i in Eq. 3) is the maximum stress at failure for each test specimen. If estimated values of α, β, and γ were available, the KS statistic would be obtained from Eqs. 10 and 11. Typically, maximum likelihood techniques and linear regression methods have been employed to determine estimated values of α, β, and γ. Alternatively, the authors propose to directly minimize the KS statistic with respect to the parameters α, β, and γ. Powell's optimization method (discussed in the next section) is applied to obtain the minimum value of this statistic. The results, which correspond to the minimum value of D, are estimates of the three Weibull parameters (i.e. $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\gamma}$). Utilizing Eqs. 3 and 8 assumes that the test specimen geometry is a unit volume and the specimen is subjected to a uniaxial tensile stress. To circumvent this restriction, the expression

$$Z_i = 1 - exp\left(- \left[V_T \left(\frac{\sigma_i - \tilde{\gamma}}{\tilde{\beta}} \right)^{\tilde{\alpha}} \right] \right) \tag{13}$$

is substituted for tensile specimens where all failures occur within the volume (V_T) of the gage section. Here $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\gamma}$ represent estimated values of the underlying population parameters.

Two basic failure populations were admitted in the formulations presented here, i.e., failures attributed to surface flaws and those due to volume flaws. This traditional approach of grouping failure origins into volume and surface flaws is an artifact from parameter estimation techniques developed for monolithic ceramics. Due to the lack of experimental data, this division (which must be based on fractographic analysis) may, or may not be appropriate for ceramic composites. At the present time, maintaining uniform densities throughout the bulk of a ceramic composite material is a major impediment that restricts the widespread commercialization of ceramic composites. Therefore, it is anticipated that the majority of failures will initiate within the volume of a ceramic composite. However, this may change as processing techniques are improved. If failures occur along the surface of the tensile specimen, the expression

$$Z_i = 1 - exp\left(- \left[A_T \left(\frac{\sigma_i - \tilde{\gamma}}{\tilde{\beta}} \right)^{\tilde{\alpha}} \right] \right) \tag{14}$$

is used where A_T is the surface area of the gage section for the tensile specimen.

Since the individual failure data (σ_i) represent the failure strength of a given ceramic test specimen, the estimators presented here were formulated for two widely used test configurations: the four-point bend test and the uniaxial tensile test (which was discussed above). Currently, the four-point bend-bar is the more popular test geometry used in strength tests of ceramic materials. When failures occur within the volume of a bend-bar specimen, the expression for Z_i takes the form

$$Z_i = 1 - exp\left\{ - \frac{V_B}{2(\tilde{\alpha} + 1)} \left(\frac{\sigma_i - \tilde{\gamma}}{\sigma_i} \right) \left(\frac{\sigma_i - \tilde{\gamma}}{\tilde{\beta}} \right)^{\tilde{\alpha}} \right\} \tag{15}$$

This expression corresponds to pure bending. This is an acceptable assumption when failure of all test specimens within a sample occurs between the inner loads depicted in Figure 1. Ignoring observations that fail outside the gage section will effectively censor the sample, and the methods presented here will not be valid. In Eq. 15, V_B represents the volume of the bend-bar specimen within the inner load span. Using this expression for Z_i, the KS statistic D is once again minimized

with respect to the three Weibull parameters. Using Powell's optimization method, the results are the three Weibull parameters that minimize the statistic D for a given sample (i.e. $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\gamma}$).

If failure of the bend specimens is due to surface flaws, Z_i takes the form

$$Z_i = 1 - exp\left[- \frac{1}{2(h+b)} \left(\frac{\sigma_i - \tilde{\gamma}}{\tilde{\beta}} \right)^{\tilde{\alpha}} \left[\frac{h}{\tilde{\alpha}+1} \left(\frac{\sigma_i - \tilde{\gamma}}{\sigma_i} \right) + b \right] \right] \quad (16)$$

The dimensions h and b are the height and thickness of the bar, as identified in Figure 1. Once again failure observations must occur between the inner load span (i.e., the region of pure bending) for reasons mentioned above.

The Anderson-Darling (AD) statistic (A^2) is the second goodness-of-fit statistic considered. This statistic belongs to the Cramer-von Mises class of quadratic statistics and is defined by the expression

$$A^2 = N \int_{-\infty}^{\infty} \{F_N(x) - f(x)\}^2 [F(x)\ (1 - F(x))]^{-1}\ dF(x) \quad (17)$$

where the terms $f(x)$, $F_N(x)$, $F(x)$, and N have been previously defined. Using the notation developed for the KS statistic, the AD statistic can be expressed as

$$A^2 = -N - (1/N) \sum_{i=1}^{N} \{(2i-1) [\ln Z_i + \ln(1 - Z_{N+1-i})]\} \quad (18)$$

As before the sum of Z_i depends on the test configuration and the failure mode (assuming that the Weibull distribution characterizes the underlying failure population). For the case where the uniaxial tensile test is used, and failure is the result of volume flaws, Z_i takes the form given in Eq. 13. When failures of a uniaxial tensile specimen are due to surface flaws, Z_i takes the form given in Eq. 14. For the case where a four point bend configuration is used, and the failures are the result of volume flaws, the Z_i function is given by Eq. 15. When failures of four point bend tests are the result of surface flaws, the form for Z_i is given by Eq. 16.

POWELL'S OPTIMIZATION METHOD

As noted previously, Powell's optimization method (see Press et al., 1986) minimizes the EDF statistics for each specimen configuration presented above. This optimization method is an iterative scheme, where the search for a minimum functional value is conducted along a specified set of direction vectors. The number of direction vectors corresponds to the number of parameters(constrained or unconstrained) associated with the function. The EDF statistics (i.e., the function being optimized) will depend on specimen geometry, individual failure observations, and the estimated parameters $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\gamma}$. However, the specimen geometry will not change for a given sample, thus the EDF statistics are optimized with respect to the parameters α, β, and γ. In essence this method locates, in succession, an optimum point along each direction vector. An arbitrary set of direction vectors can be utilized to optimize a given function; however, Powell's method employs noninterfering (or conjugate) directions in order to speed convergence. This alleviates difficulties which arise when optimization along one direction vector is disturbed by a subsequent search along a new direction vector. The method formulates and updates n mutually conjugate directions, where n (for this case equals three i.e., α, β, and γ) defines the size of the parameter space. The set of direction vectors is updated by discarding the direction vector that produced the maximum change during an iteration. The average direction defined by the initial and final point of an iteration is substituted, and becomes the initial direction vector for the next iteration. Note that this method does not produce quadratic convergence, but nevertheless is very robust.

As indicated above, the optimized parameter space is defined by the estimates of the Weibull parameters α, β, and γ. Since a good choice of starting values ($\tilde{\alpha}_0$, $\tilde{\beta}_0$, and $\tilde{\gamma}_0$) is essential in quickly locating the optimum point, the results of Cooper's modified least-squares estimation method are used as the initial vector for Powell's method. Further restrictions are imposed on the optimization process. Negative values for the estimated Weibull parameters, and estimated threshold parameters ($\tilde{\gamma}$) larger than the smallest failure stress in a given sample, are not physically meaningful. Thus directions that produce these parameter values are discarded in the update of the direction vectors, and parameter values are reset to the minimum allowable values.

Example

Since failure data for CMC material systems are sparse, only failure data for a monolithic sintered silicon nitride (grade SNW-1000, GTE Wesco Division) are used to illustrate the

relative merits of the proposed estimation techniques. This data was published by Chao and Shetty (1991) and is reprinted in Table 1. These values represent the maximum stress at failure for 27 four-point bend specimens. The outer support span for the test fixture was 40.4 mm, and the inner load span was 19.6 mm. The cross sections of the test specimens were 4.0 mm wide, and 3.1 mm in height. All failures occurred within the 19.6 mm inner load span, thus it was assumed that each specimen was subjected to pure bending.

Chao and Shetty performed a fractographic analysis of each specimen using optical and scanning electron microscopy. These studies indicated that all failures were initiated at subsurface pores (i.e., volume defects). Hence, equations for bending associated with volume defects are used for parameter estimation. Five methods were used to estimate the Weibull parameters from this set of failure data. These were Cooper's three parameter least squares method, the three parameter modified least squares method outlined by Duffy et al. (1993), minimizing the KS statistic, minimizing the AD statistic, and a two parameter estimation using the maximum likelihood estimation technique outlined in the ASTM Standard Practice C-1239. The Kolmogorov-Smirnov statistic (D) and Anderson-Darling statistic (A^2) were computed for each set of parameter estimates. The values of these EDF statistics, and the estimated parameters for each method are listed in Table 2.

A comparison of estimates obtained by both least-squares methods shows small differences in the estimated Weibull threshold parameter $\tilde{\gamma}$. Larger differences are present between the two methods in the estimates of the other parameters. Specifically, the modified least squares method provided a higher estimate for $\tilde{\alpha}$ than did Cooper's method, and a lower estimate for $\tilde{\beta}$ Furthermore, both goodness-of-fit statistics (D and A^2) are smaller for Cooper's method than for the modified least- squares method. Duffy et al. (1993) demonstrated that the modified least squares method is theoretically more rigorous than Cooper's original work since the modified method attempts to minimize a true residual. However, it is apparent from this example that Cooper's original approach yields better goodness-of-fit statistics. This discrepancy in part motivated the development of estimators based on minimizing goodness- of-fit statistics.

Estimates of the Weibull parameters obtained by minimizing the KS statistic result in the smallest value of D, which is not surprising. Similarly, estimates of the parameters obtained by minimizing the AD statistic result in the smallest value of A^2 in comparison to the other estimation methods. However, the Weibull parameters obtained by optimizing the goodness-of-fit statistics differ considerably from the estimates obtained using the least-squares techniques. Specifically, the value of $\tilde{\beta}$ from minimizing the goodness-of-fit statistics is nearly twice the value obtained with the least-squares techniques. As an additional comparison, parameter estimates from using a maximum likelihood estimator assuming a two-parameter Weibull distribution are included in Table 2. These estimates produce the highest values for both goodness-of-fit statistics.

Finally, cumulative distribution functions for all of the parameter estimates are plotted on a single Weibull diagram (see Figure 2). All of the failure data fall relatively close to all four of the three-parameter curves. This type of visual assessment (along with its highly subjective interpretation) should provide the motivation for the use of quantitative measures in determining the goodness-of-fit.

CONCLUSION

New methods of parameter estimation are proposed that are based on the minimization of goodness-of-fit statistics. These methods are used to estimate Weibull parameters from failure data whose population is assumed to be characterized by a three-parameter Weibull distribution. As an example, the proposed methods were compared with other parameter estimation methods, using failure data from a monolithic ceramic material. The proposed methods provided a better fit to the failure data in terms of the EDF statistics. However, to completely test the proposed methods, performance criteria like bias and invariance have to be evaluated through the use of Monte Carlo simulations.

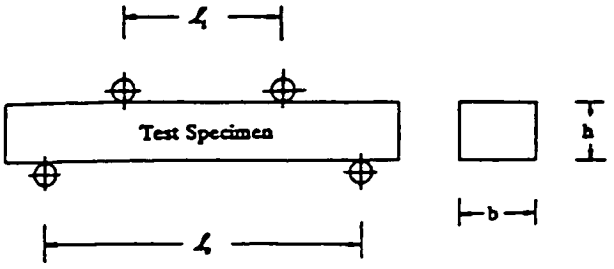

FIGURE 1 GEOMETRY AND NOTATION FOR A FOUR-POINT BEND TEST SPECIMEN.

TABLE 1 FOUR-POINT BEND FAILURE DATA FOR SILICON NITRIDE.

Specimen No.	Strength (MPa)
1	613.9
2	623.4
3	639.3
4	642.1
5	653.8
6	662.4
7	669.5
8	672.8
9	681.3
10	682.0
11	699.0
12	714.5
13	717.4
14	725.5
15	741.6
16	744.9
17	751.0
18	761.7
19	763.9
20	774.2
21	791.6
22	795.2
23	829.8
24	838.4
25	856.4
26	868.3
27	882.9

TABLE 2 PARAMETER ESTIMATES OBTAINED FROM FOUR-POINT BEND FAILURE DATA

Estimation Method	$\tilde{\alpha}$	$\tilde{\beta}$ $(MPa \cdot mm^{(3/\tilde{\alpha})})$	$\tilde{\gamma}$ (MPa)	D $(\times 10^{-2})$	A^2 $(\times 10^{-1})$
Cooper's Least Squares	1.625	892.37	560.84	9.404	1.749
Modified Least Squares	1.677	861.93	558.08	9.538	1.798
KS Estimator	1.375	1298.44	558.08	6.080	1.963
AD Estimator	1.168	1537.03	581.09	7.676	1.406
Two-Parameter MLE	10.119	974.09	0.00	11.20	5.394

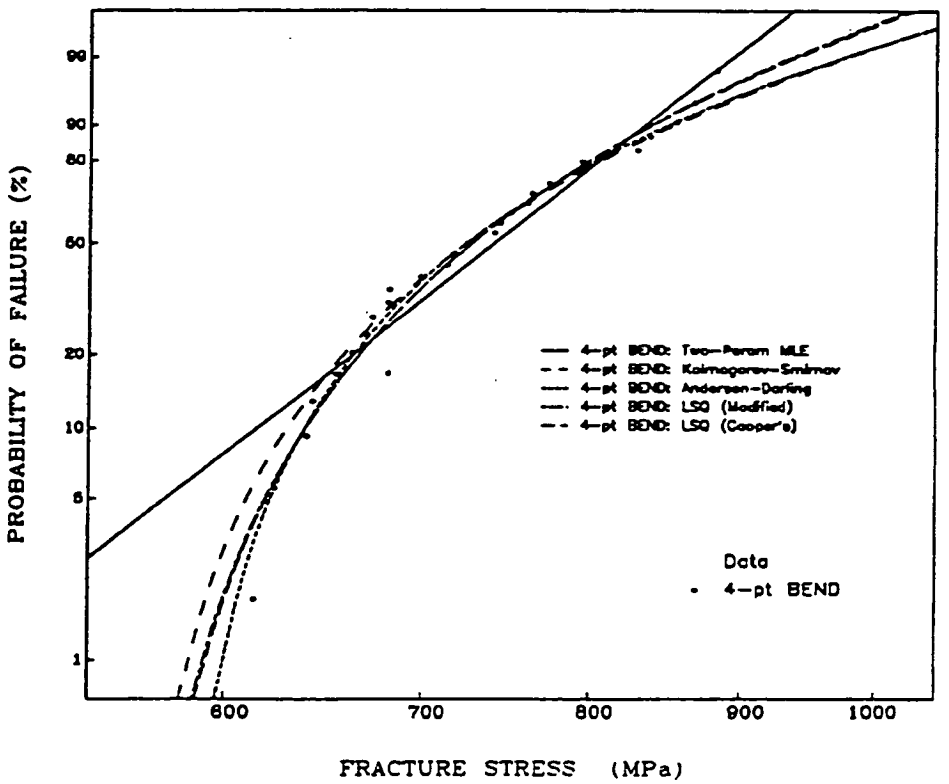

FIGURE 2 WEIBULL DIAGRAM FOR FIVE PARAMETER ESTIMATES

REFERENCES

Antle, C.E., and Klimko L.A., 1973, "Choice of Model for Reliability Analysis and Related Topics, II," ARL 73-73-0121, Aerospace Research Laboratories, Wright-Patterson AFB, Ohio, AD-772775.

Chao, L.-Y., and Shetty, D.K., 1991, "Reliability Analysis of Structural Ceramics Subjected to Biaxial Flexure," *Journal of the American Ceramic Society*, Vol.74, pp. 333-344.

Cooper, N.R., 1988, "Probabilistic Failure Prediction of Rocket Motor Components," PhD Thesis, Royal Military College of Science, Shrivenham, Swindon, Wilts.

Duffy S.F., Powers, L.M., and Starlinger A., 1993, "Reliability Analysis Using Three Parameter Weibull Distribution," *Transactions of the ASME, Journal of Engineering for Gas Turbines and Power*, Vol. 115, pp.109-116.

Duffy, S.F., Quinn, G.D., and Johnson, C.A., 1992, "Reporting Uniaxial Strength Data and Estimating Weibull Distribution Parameters for Advanced Ceramics," ASTM Standard Practice C-1239, Subcommittee C28.02.

Foley, M.R., Pujari, V.K., Sales, L.C., and Tracey, D.M., 1993, "Silicon Nitride Tensile Strength Database from CTAHE Processing for Reliability Project," Life Prediction Methodologies and Data for Ceramic Materials, ASTM STP 1201, C.R. Brinkman and S.F. Duffy, Eds., American Society for Testing and Materials, Philadelphia, PA.

Margetson, J., and Cooper N.R., 1984, "Brittle Material Design Using Three Parameter Weibull Distributions," *Probabilistic Methods in the Mechanics of Solids and Structures, Symposium*, Stockholm, Sweden, (Eggwertz, S., and Lind, N.C., ed.), Springer-Verlag.

Palko, J.L., Starlinger A., Duffy S.F., and Thomas D.J., 1993, "Ceramic Composite Analysis and Reliability Evaluation of Structures (C/CARES) - User's and Programmer's Manual," NASA TP (in review).

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T., 1986, Numerical Recipes - The Art of Scientific Computing, Cambridge University Press, New York.

Stephens, M.A., 1986, "Tests Based on EDF Statistics," Goodness-of-Fit Techniques, R.B. D'Agostino and M.A. Stephens, eds. Marcel Dekker, Inc., New York, pp. 97-194.

Thomas D.J., and Wetherhold, R.C., 1991, "Reliability Analysis of Continuous Fiber Composite Laminates", *Composite Structures*, Vol. 17, pp. 277-293.

Tracy P.G., Rich T.P., Bowser R., and Tramontozzi L.R., 1982, "On the Statistical Nature of Fracture," *International Journal of Fracture*, Vol. 18, pp. 253-277.

Weibull, W., 1939, "A Statistical Theory of the Strength of Materials," *Ingeniors Vetenskap Akadiem Handlingar*, No. 151, pp. 5-45.

Weibull, W., 1951, "A Statistical Distribution Function of Wide Applicability," *Journal of Applied Mechanics*, Vol. 73, pp. 293-297.