Advanced Statistical Concepts of Fracture in Brittle Materials

C.A. Johnson and W.T. Tucker, General Electric Corporate Research and Development

CERAMICS are increasingly being considered as load-bearing structural materials that are suitable for use in both ambient and elevated temperature applications. These applications are driven by the need for specific combinations of critical properties (mechanical, thermal, chemical, optical, electrical, magnetic, and others) that are often only available in ceramic materials. The function of a load-bearing component, of course, is threatened by fracture from stresses generated during service. Therefore, an understanding of the unique characteristics of fracture in ceramics is important for successful implementation.

Many of the unique mechanical properties of ceramics can be traced to the difficulty of dislocation motion in their typical ionic and covalent bonded crystal structures (Ref 1). The absence of dislocation motion and the associated absence of inelastic deformation cause most monolithic ceramics to fail in a brittle manner, with relatively low energy absorption. With only elastic deformation available to accommodate crack tip stresses, conditions of crack stability are defined by the Griffith/Orowan criteria, where the fracture stress is inversely proportional to the square root of the flaw size. This inverse square-root relationship, combined with the nonuniform flaw sizes that are inevitable in ceramics, results in two common observations. First, multiple observations of fracture strength in otherwise similar specimens generally include a great deal of scatter or variability about the average strength. Second, the average strength is a function of specimen size such that larger specimens are lower in strength than smaller specimens.

Quantitative descriptions of the variability and size dependence of strength are necessary to design structural ceramic components, to predict component reliability, and to compare the merits of different ceramics for a given application. The purpose of this article is to review some recent advances in this area of fracture statistics.

Distributional Models

Quantitative descriptions of the statistical nature of fracture in brittle materials rely on a distribution function to define the cumulative probability of failure, P, as a function of all dependent variables that influence failure, such as stress level and component size and geometry. The majority of distribution functions considered for this purpose are based on the weakest link concept, which simply states that the entire body will fail when the stress at any defect is sufficient for unstable crack propagation of that defect. Examples of applicable weakest link distribution functions include the Weibull distribution (Ref 2, 3) and extreme value distributions (Ref 4).

The Weibull distribution has become very popular in descriptions of ceramic fracture for a number of reasons. The function is mathematically simple and easy to manipulate, and it is a weakest link distribution with the proper "limiting conditions." But most importantly, the distribution has been successful in describing a great deal of fracture data representing many different materials from numerous investigations. In many respects, the Weibull distribution is to weakest link problems what the Gaussian distribution is to situations where the central limit theorem applies. For these reasons, the remainder of this article concentrates on the Weibull distribution.

In his 1939 paper (Ref 2), Weibull introduced both a uniaxial model of probabilistic failure and a related multiaxial model that treats some aspects of failure produced by multiaxial stresses. His uniaxial model is given by

$$P = 1 - \exp\left[-\int^{V}\left(\frac{\sigma}{\sigma_0}\right)^m dV\right] \qquad \text{(Eq 1)}$$

where P is the cumulative probability of failure, σ is the stress of a unit volume dV at the time of failure, m is the Weibull modulus, and σ_0 is a normalizing parameter. (Note that the symbol F is more commonly employed for cumulatives in the statistical literature.) If failure is caused by defects that are distributed randomly throughout the volume of a specimen or component, then the integration is carried out over all elements of volume, dV. If the defects are restricted to surfaces (such as machining defects), then the integration is carried out over surface elements, dA.

For the general case where stress is a function of position in the body, the integration can be carried out to yield

$$P = 1 - \exp\left[-kV\left(\frac{\sigma_{max}}{\sigma_0}\right)^m\right] \qquad \text{(Eq 2)}$$

where k is a dimensionless "load, or structure factor" and σ_{max} is the maximum stress in the structure at the time of failure. For the case of uniform uniaxial tension, k is unity and Eq 2 reduces to the form of the two-parameter Weibull distribution commonly encountered in the statistical literature. For all other loading geometries, k is a function of m and, when evaluated, is always less than unity. The product of k times V is often termed the "effective volume" and, as the term implies, is the volume of material that is effectively under uniform uniaxial tension.

If the uniaxial Weibull model described above is valid for a given material, then the two adjustable parameters, m and σ_0, are material constants. The value of the Weibull parameters are estimated from experimental fracture strengths using an "estimator," such as linear regression or maximum likelihood (Ref 5). In the simplest practical situation, Eq 2 can then be used to estimate the probability of failure of a component of interest where the maximum stress, the component size, and the load factor are known for that structure. Conversely, Eq 2 can be solved for σ_{max} to estimate the maximum stress that can be survived in the component at a specified probability of failure (specified failure rate).

Unfortunately, the procedure outlined above is oversimplified. In most real-life situations, one or more complications are present and a more complex analysis is required. A partial list of such complications, with key references where available, includes:

- Multiple active flaw populations (Ref 6)
- Multiaxial stress states (Ref 7, 8)
- Censored and/or proof-tested data (Ref 9)
- In-service modification of flaws by oxidation, slow crack growth, and other factors (Ref 10)
- Directional anisotropy of strength

If not accounted for, the presence of any of these complications will reduce the accuracy of strength/probability estimates in unpredictable ways. Although many recent efforts in fracture statistics have addressed complications such as these, and progress has been made, many difficult problems remain. In particular, progress is needed in approaches that allow more than one complication to be present simultaneously.

In addition, there are three further requirements for a comprehensive probabilistic fracture methodology:

- The analysis should allow data to be combined or pooled from multiple specimen sizes and geometries, so that all available fracture data can be integrated into strength/probability estimates and size-scaling aspects of the model can be tested
- The analysis must be capable of providing a measure of the statistical uncertainty in estimates of strength/probability (confidence and tolerance bounds)
- The analysis must be compatible with goodness-of-fit tests to determine if the experimental data are consistent with all aspects of the assumed model of probabilistic behavior

A research effort funded by the U.S. Department of Energy, Office of Transportation Systems, which was subcontracted through Oak Ridge National Laboratory and carried out by the authors (Ref 11) has studied many aspects of the complications and requirements listed above. The following section reviews some results of that effort in the areas of combining data and confidence/tolerance bounds on estimates. The approaches developed for combined data and confidence/tolerance are described and demonstrated for the uniaxial model of Eq 2. Importantly, it has been shown that similar techniques are applicable to more comprehensive models of multiaxial failure, such as those of Ref 7 and 8. Thus, the form of Eq 2 is quite general.

Weibull Estimators for Combined Data

A Weibull estimator is a method, or algorithm, to analyze fracture data and estimate useful quantities. These quantities may include both point estimates, such as distribution parameters and predicted strengths, as well as estimates of intervals, such as confidence and tolerance bounds. As mentioned above, there are advantages in efficiency and model validation that result from combining or pooling fracture data from multiple specimen sizes and geometries.

In the studies of combined data, it has been useful to categorize problems according to the type of loading and the presence or absence of size scaling. The following four "classes" of problem have been considered:

- Uniform tensile stress and single specimen size (Class I)
- Uniform tensile stress and multiple specimen sizes (Class II)
- Common load factor (k) and multiple specimen sizes (Class III)
- Differing load factors and multiple specimen sizes (Class IV)

In each, it is assumed that the Weibull distribution with a size-scaling term (volume, area, or edge length) adequately describes the strength to failure. Class I problems of strength/probability estimation are the easiest to analyze, but the least useful for practical applications. On the other hand, Class IV problems are the most useful of the four but, accordingly, are the most difficult to properly analyze. The progression of complexity from Class I to IV has helped in the development of flexible and rigorous Class IV estimators.

Example of Class IV Strength Data. To demonstrate the derivation and application of Weibull estimators for Class IV problems, strength tests were performed on specimens of boron-doped sintered silicon carbide (SiC) (Ref 12). Testing was done in six bending configurations. The A, B, and C specimen geometries of the MIL-STD-1942MR (virtually identical to ASTM Standard C1161, adopted in September 1990) were used in both 3-point and 4-point bending. Specimen A is defined to have a cross section of 1.5 by 2.0 mm (0.059 by 0.079 in.) and is tested on a 20 mm (0.787 in.) outer span (10 mm, or 0.394 in., inner span when 4-point testing is done). Specimen B doubles every dimension of specimen A, and specimen C doubles those dimensions yet again. The specimen volumes therefore vary by a factor of 64, and the areas vary by a factor of 16. The difference between 3-point versus 4-point bending contributes another factor of approximately 10 to the ra of effective volumes and areas. Therefore, data span a range of 500 to 1000 in effec volume and a range of 100 to 200 in effec areas.

Specimens were prepared by isopress and sintering billets from which the spe mens were cut and ground to shape. T specimen matrix was planned to allow det tion of complications, such as billet-to-bil differences and strength dependence on lo tion within a billet. No such complicatic could be resolved and, therefore, all spe mens were considered to have a comm family of defects. A total of 137 bend spe imens were tested. The results of analyzi the specimens, one group at a time, are pr vided in Table 1.

The Weibull parameters listed in Table are the output of conventional maximum lik lihood analysis. Group-to-group variations i the two Weibull parameters are typical of st tistical sampling error on data sets of this siz Therefore, the variation in Weibull modulu for instance, is not believed to indicate th different groups contained different funda mental flaw distributions.

Low-power (50×) stereo microscopy wa used to identify the location of each fracture initiating defect in the SiC and to determin whether it was a surface, subsurface, or edge related defect. Because of the predominanc of surface-related fracture origins observed i this data set, all Class IV analyses and all ref erences to effective size discussed belov assume that strength was controlled by sur face defects.

In addition to the six estimates of m in Table 1, the strength as a function of specimen size can be used to derive a seventh, independent estimate of m. The seventh estimate adds to the confusion of determining the "best" estimate of the true Weibull modulus. Should the seven estimates be averaged? Should they be weighted, somehow, by the number of specimens contributing to each estimate? The Class IV maximum likelihood estimator described below is a more rigorous and statistically efficient approach to extracting all useful information in data sets such as the above example.

Maximum Likelihood Estimator. The two adjustable parameters of the Weibull distribution are derived from fracture data using a

Table 1 Analysis of strength tests performed on boron-doped sintered silicon carbide

Geometry	Number	Average strength, MPa	Average strength, ksi	Standard deviation, MPa	Standard deviation, ksi	Weibull modulus, m	σ_0, *$MPa \cdot mm^{2/m}$	σ_0, *$ksi \cdot in.^{2/m}$
3-pt A	18	388.11	56.29	33.36	4.838	14.57	430.99	40.093
4-pt A	17	312.85	45.375	34.83	5.052	9.43	459.24	33.546
3-pt B	18	350.96	50.902	31.12	4.514	12.20	449.93	38.405
4-pt B	48	303.31	43.991	24.17	3.506	14.28	430.28	39.666
3-pt C	18	325.65	47.232	21.69	3.146	16.39	418.83	40.929
4-pt C	18	283.78	41.159	22.35	3.242	14.48	441.08	40.921

Weibull estimator. For the first three classes of problems described above, methods of estimating the parameters from fracture data are available in the literature. However, only a few methods have been developed for estimating parameters in Class IV problems (Ref 13, 14).

Two new Weibull estimators have recently been developed to analyze Class IV problems of combined data (Ref 12). One is based on linear regression and the other on maximum likelihood. Both require successive approximation methods and are therefore best suited for computer analysis. The estimator based on maximum likelihood has been found to be more efficient than linear regression in maximizing the use of information from a data set. Therefore, only the maximum likelihood method is described herein.

Derivation of the Class IV maximum likelihood estimator parallels the derivation of the Class I estimator as published by Trustrum and Jayatilaka (Ref 5), which in turn is based on maximum likelihood derivations, such as those of Lawless (Ref 15), Mann *et al.* (Ref 16), and Nelson (Ref 9). Unlike earlier derivations, the Class IV analysis must account for the loading factor, k, which is generally a function of the Weibull modulus.

Maximum likelihood finds the maximum, as a function of the unknown parameters, in a quantity known as the "log likelihood." The log likelihood, l, is defined as the sum of the logs of the probability density, f, for each observed specimen. The probability density of Eq 2 is:

$$f = \frac{dP}{d\sigma} = \frac{mkV}{\sigma_0^m}\sigma^{m-1}\exp\left[-kV\left(\frac{\sigma}{\sigma_0}\right)^m\right] \qquad \text{(Eq 3)}$$

For simplicity in upcoming equations, the subscript "max" from Eq 2 has been dropped from the stress in the numerator of the exponential, but the term still represents the maximum stress in the structure. The log likelihood, l, for a group of n strength measurements is then:

$$\begin{aligned} l &\equiv \sum_{i=1}^{n} \ln(f_i) \\ &= n\ln(m) - nm\ln(\sigma_0) \\ &\quad + \sum_{i=1}^{n}\ln(k_iV_i) + (m-1)\sum_{i=1}^{n}\ln(\sigma_i) \\ &\quad - \sum_{i=1}^{n} k_iV_i\left(\frac{\sigma_i}{\sigma_0}\right)^m \end{aligned} \qquad \text{(Eq 4)}$$

Partial derivatives of Eq 4 are then taken with respect to the two Weibull parameters, m and σ_0. The derivative with respect to m must account for the functional dependence of k_i on m. Each of these partial derivatives is then set equal to zero to find the combination of the two parameters that leads to the maximum in the log likelihood. The two partial derivatives can be combined such that σ_0 drops out, leaving the relationship:

$$0 = \frac{n}{m} + \sum_{i=1}^{n}\frac{dk_i/dm}{k_i} + \sum_{i=1}^{n}\ln(\sigma_i) - \frac{n\sum_{i=1}^{n}[k_iV_i\sigma_i^m\ln(\sigma_i) + V_i\sigma_i^m\,dk_i/dm]}{\sum_{i=1}^{n}V_ik_i\sigma_i^m} \qquad \text{(Eq 5)}$$

The maximum likelihood estimate of m is then evaluated by iteratively determining the value of m that satisfies Eq 5. Under usual conditions, only a single value of m satisfies Eq 5 for any given data set. After m is known, the second Weibull parameter can be determined without iteration by rearranging the partial of Eq 4 with respect to σ_0 that was earlier equated to zero:

$$0 = n\,\sigma_0^m - \sum_{i=1}^{n} k_iV_i\sigma_i^m \qquad \text{(Eq 6)}$$

In order to evaluate the maximum likelihood estimates of the Weibull parameters, the dependence of both k and dk/dm must be known as a function of m for every specimen geometry tested.

The true parameters of the distribution are referred to as m and σ_0. By standard convention, estimates of these parameters from analysis of data are known as $\hat{m}$ and $\hat{\sigma}_0$. An efficient algorithm for the evaluation of maximum likelihood estimates for Class IV problems has been developed that requires only five to ten iterations to determine $\hat{m}$ and $\hat{\sigma}_0$ to approximately five significant digits for most data sets that have been analyzed. The 137 strength measurements of SiC described above result in Weibull parameters of $\hat{m}$ = 14.22 and $\hat{\sigma}_0$ = 433.1. These estimates assume that failure is controlled by a distribution of surface defects. The Weibull modulus is dimensionless. In order for the exponent of Eq 2 to be dimensionless, the dimensions of σ_0 must account for the surface area term in the exponential. Therefore, the dimensions of σ_0 are a function of the Weibull modulus and are (MPa · mm$^{2/m}$).

Knowledge of the Weibull parameters allows the fracture strength to be estimated for any component size and shape that may be of interest at any probability of failure of interest using Eq 2. These estimated strengths can be displayed graphically, as in Fig 1, where log of fracture stress is plotted versus log of effective stressed area (again, fracture is controlled by surface-related defects). Figure 1 includes the complete set of 137 fracture strengths, with different symbols used to identify the various specimen geometries. The effective area used to plot each data set is the product of k times the surface area of the specimen. Because k is a function of $\hat{m}$, the plotting position is a function of the Weibull modulus.

Also included in Fig 1 are two parallel straight lines representing the strength versus

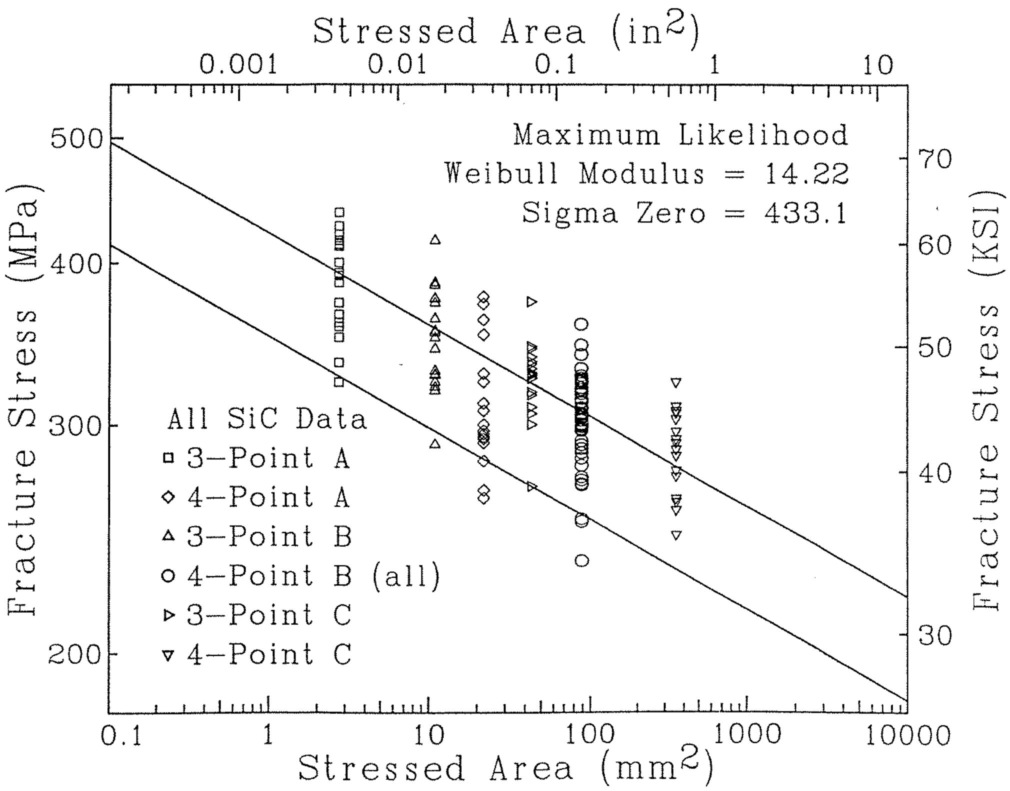

Fig 1 Dependence of strength on specimen size for boron-doped, sintered β-SiC tested in six configurations of specimen size and testing geometry. Upper line is 0.5 quantile predicted behavior; lower line is 0.05 quantile.

effective area at two different probabilities of failure. The upper straight line describes the median (or 0.5 quantile) behavior, where 50% of the specimens or components are expected to fail. The lower straight line describes the 0.05 quantile behavior where only 5% are expected to fail. Of course, an entire family of parallel straight lines exists that represents all quantile levels.

Confidence and Tolerance Bounds. Confidence bounds give a measure of statistical uncertainty associated with a quantity estimated from data sampled from a distribution. Using experimental data, one cannot exactly determine the complete distribution; this inexactness must be quantified in order to make sensible decisions. For instance, the exact location of the true 0.05 quantile line for the SiC data is unknown. Figure 1 only shows an estimate of the position of the line. Confidence bounds offer a way to quantify the inexactness and give an indirect measure of the risk associated with the use of an estimated value as the true but unknown value. In the above definition of confidence bounds, the term "statistical uncertainty" is the variation that is due to random sampling error and random measurement error; "estimated" means calculated from sampled data that were obtained from a distribution of all possible data values; and "quantity" is a parameter, quantile, or future value of a distribution.

There are three primary types of confidence bounds, which, by convention, are defined as:

- Confidence limits: bounds on a parameter of a distribution
- Tolerance limits: bounds on a quantile of a distribution
- Prediction limits: bounds on a future value sampled from a distribution

A distribution characterizes the population of strengths, for example, of all components made from a certain material using a given manufacturing process. Generally, there is only a sample of material available for use in specifying the distribution, and observations are taken in the presence of measurement error. The measurement error can either be of a random nature or occur as a bias (constant offset). In the developments herein, neither of these forms of measurement error are considered. Other sources of uncertainty, such as model "incorrectness," are not accounted for in the estimation procedures, either. The resulting bounds therefore only account for statistical sampling error.

Statistical sampling error arises from the fact that a sample quantity, for example, a sample mean, does not equal the true value associated with the population, for example, the population mean. Therefore, the underlying distribution cannot be *exactly* determined or specified with a finite number of observations. This inexactness must be quantified in order to make sensible decisions about design considerations and the ability of a component to withstand a required service stress. One way to quantify the uncertainty is to employ confidence bounds, which give an indirect measure of the risk associated with using an estimated value as the true value. When a "high" confidence coefficient is used, then it is "safe" to act as if the true value is contained within the *observed* interval. But the terms high and safe are open to subjective interpretation; what is high and/or safe for one person may not be for another. Thus, the level of acceptable confidence must be determined for each individual problem.

The qualifier "observed" is emphasized above because the location and, often, the length of a confidence interval changes from one sample to the next under identical sampling conditions. The confidence coefficient reflects this kind of uncertainty by giving the long-run probability that an interval covers the true, but unknown, value of interest. This is illustrated schematically in Fig 2, where one of the 20 intervals shown does not cover the true value, whereas the others do. Each interval has been deduced by an independent sampling of the distribution. The behavior shown is analogous to that which would occur in the long run for a 95% confidence interval on the unknown true value. In practice, of course, the validity of a specific interval in covering the true value would not be known in advance. The probability of an interval missing the true value then gives a measure of the risk one is willing to take in making a decision about the true value. The future is indeed uncertain.

In the calculation of confidence and tolerance bounds, two overall approaches have evolved: exact bounds based on sampling theory and approximate bounds based on asymptotic considerations. Exact bounds arise by inverting statistical tests and/or conditional integration methods first described by Fisher (Ref 17). These methods can be shown, analytically, to have exact coverage probabilities, as opposed to an approximate coverage. Approximate bounds include those obtained by maximum likelihood asymptotics, some likelihood ratio methods, and some bootstrap (simulation) methods. Reference 18 gives an annotated bibliography that contains further references to the large literature on confidence bounds for lifetime distributions. Because a lifetime is analogous to a strength, lifetime methodologies can apply, in some cases, to the present situation.

To some extent, existing methods for normal theory and likelihood ratio approaches offer solutions up through Class III problems. However, current bootstrap procedures are generally not applicable, without some development, to any of the four classes. These two approaches are emphasized because they appear to be the most promising for use in obtaining confidence and tolerance bounds in Class IV problems. More details on this issue and references to other approaches are given in Ref 19.

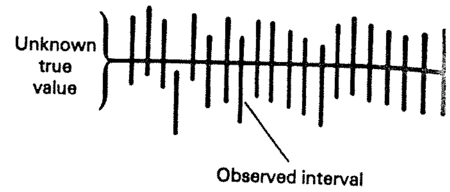

Fig 2 Projected confidence intervals from a sequ of independent estimates; for 95% bou 1 in 20 intervals on average will not cover unknown value

Bootstrap Techniques for Confide and Tolerance. As the term "bootstr implies, these methods pull themselves up their own bootstraps; that is, they gat information about variability in estimates evaluating only the data available in the c set of interest. A good review of bootst techniques and their basis can be found in 20. These techniques are readily applied t variety of distributions, but have not yet b applied to the Weibull distribution with s scaling.

Bootstrap techniques characterize a co bination of the intrinsic uncertainty of es mation that is due to sampling error and a additional uncertainty that is due to inef ciencies of the estimator. These techniqu are ultimately dependent on an estimator digest the experimental measurements fracture strength and estimate parameters the distribution and/or estimates of t strengths of components with other sizes a loading geometries. The estimator used in th discussion is the Class IV maximum like hood estimator described above.

Two types of bootstrap techniques ha been studied: parametric and nonparametri Both involve simulation and analysis of ma "artificial" data sets that are derived eith directly or indirectly from the experiment data. In the nonparametric technique, n assumption of the form of the strength di tribution is made prior to generation of th simulated data sets. The simulated data a created by randomly choosing strengths fro the original data "with replacements"; that i after a strength is chosen from the list c experimental strengths, it will still be avai able to be chosen again within that same sim ulated data set. Therefore, each simulated dat set is virtually guaranteed to have some repea observations. In the parametric technique, th estimator is used to estimate the Weibu parameters for the experimental data. Simu lated strengths are then generated by ran domly choosing simulated specimens from th infinite population of possible specimens tha are consistent with those distributio parameters.

In both the parametric and nonparametric techniques, the number of specimens gener ated for each size and geometry of test spec imen is identical to that of the original data set. The estimator is then used to analyze the simulated data set. This simulation and analysis procedure is repeated a large number of times (typically, 1000 times). The variability in estimates from the many simulated data sets reflects the intrinsic variability of the estimator in analyzing data sets similar to the original experimental data set. For instance, if the $\hat{m}$'s resulting from the simulations are ordered from smallest to largest, then the range of $\hat{m}$ values from the 25th to the 975th position in this list of 1000 $\hat{m}$'s is an estimate of the 95% confidence interval on estimates of $\hat{m}$ using that estimator. For example, in 950 of 1000 trials, the estimated $\hat{m}$ value was found to fall in this interval when sampling error was the only source of error. Similar bounds can be placed on predicted strength values at a specific component size and probability of failure by estimating the component strength for each of the bootstrap simulations (using Eq 2) and carrying out a similar ordering to find the interval in strength that contains 95% of the estimated strength values. Because there is nothing magic about 95% intervals, the confidence interval could be estimated in this fashion for any level of confidence that may be of interest.

Figure 3 is very similar to Fig 1, but includes tolerance bounds on the 0.05 quantile. The curved lines bounding the 0.05 quantile straight line are the tolerance limits as a function of effective area as determined by the parametric bootstrap technique. The corresponding 95% confidence bounds on the two Weibull parameters are 13.01 and 15.81 for $\hat{m}$ and 421.0 and 444.5 for $\hat{\sigma}_0$. The solid triangular data points are also tolerance limits on the 0.05 quantile, but were calculated by the likelihood ratio technique, discussed in the next section of this article.

The shape of the tolerance limit lines are approximately hyperbolic, as should be expected. The hyperbolic shape results in a "pinch point," where the uncertainty in strength estimates is the smallest. Regardless of the quantile being considered, this pinch point is located near the average strength of the data. Thus, as one considers component sizes with strengths that are progressively farther from this average strength (either higher or lower), the width of the tolerance interval gets progressively larger. As one considers smaller quantile behaviors, the pinch point for the new quantiles will move to the left toward progressively smaller effective specimen sizes. Typically, strength predictions are made for components that are larger than the test specimens. Therefore, the width of the tolerance interval for most components will increase as one considers progressively smaller quantile behaviors.

Figure 3 does not include the tolerance limits for the nonparametric bootstrap technique. The results are very similar for this data set. Comparisons of parametric versus nonparametric bootstrap analyses have been made for many real and simulated data sets. These comparisons have shown that the nonparametric bootstrap is more prone to bias and has a greater sensitivity to data sets that contain "apparent outliers." For these reasons, the parametric bootstrap is favored for most situations studied to date.

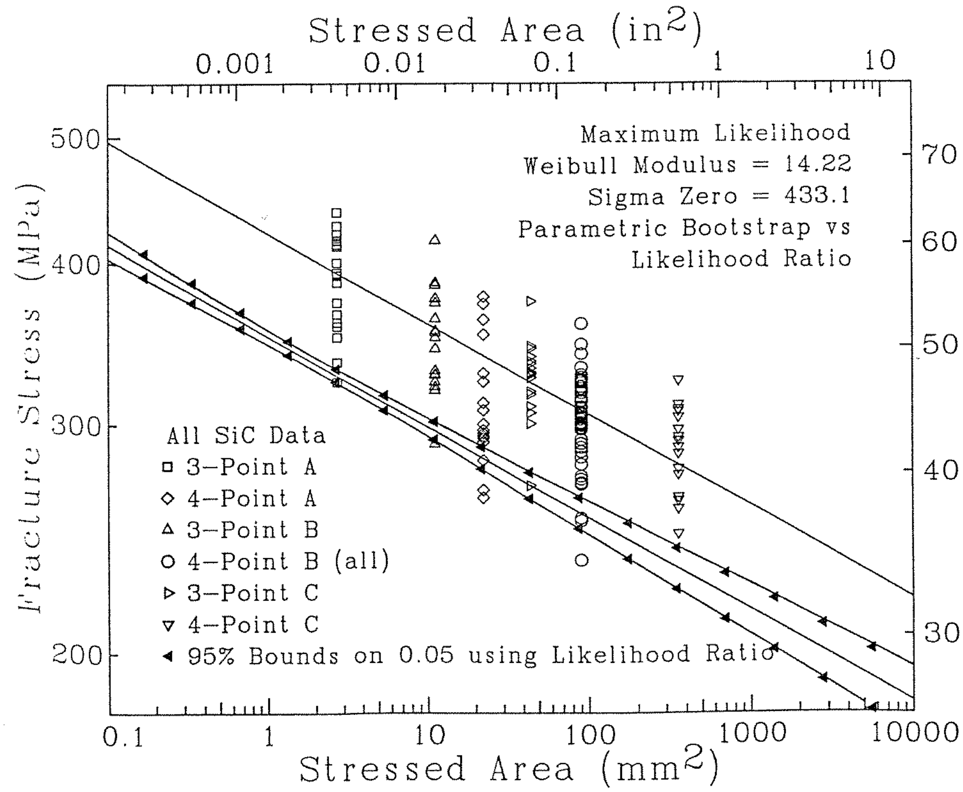

Fig 3 Plot similar to Fig 1, but includes 95% tolerance bounds on the 0.05 quantile as determined by parametric bootstrap (curved lines) and likelihood ratio (solid triangles) techniques

The bootstrap technique is computationally intensive, but it offers the potential of estimating confidence and tolerance bounds on a variety of very complex problems, even those beyond Class IV (such as those involving multiple flaw populations, multiaxial stresses, and strength degradation that is due to slow crack growth). If an estimator can be designed to estimate all the adjustable parameters of a given model, then the bootstrap technique is capable of estimating confidence bounds on all the parameters and tolerance bounds on estimated strengths. Of course, the better the quality of the estimator in using all the information available within the data set, the smaller the width of the confidence and tolerance bounds resulting from the analysis.

As with most estimators, the Weibull maximum likelihood estimator has the property of yielding offset, or biased, estimates of parameters, strengths, and confidence intervals. The magnitude of bias error often decreases as the number of samples increases. Unfortunately, the degree of bias error is difficult to predict in advance for complex estimators, such as this Class IV estimator. Bootstrap simulations for estimation of confidence and tolerance bounds contain information about the magnitude of bias introduced by the estimator. Reference 21 describes how such information can be used to correct bias errors. Although maximum likelihood estimates do generate bias, of all the Class IV estimator studies to date, maximum likelihood has generally been found to generate the least bias.

Likelihood Techniques for Confidence and Tolerance. Likelihood ratio methods offer another way to obtain confidence bounds. The development given by Cox and Oakes (Ref 22) as applied to the two-parameter (m, σ_0), Weibull distribution is closely followed. This method is based on the direct use of the likelihood ratio statistic

$$W(\sigma_0) = 2[l(\hat{m},\hat{\sigma}_0) - l(\hat{m}_{\sigma_0},\sigma_0)] \qquad \text{(Eq 7)}$$

where $l(.,.)$ is the log likelihood given by Eq 4, $(\hat{m},\hat{\sigma}_0)$ is the joint maximum likelihood estimate of (m,σ_0) and $\hat{m}_{\sigma_0}$ is the maximum likelihood estimate of m conditional on σ_0 (that is, taking σ_0 fixed). The function $l(\hat{m}_{\sigma_0},\sigma_0)$ of σ_0 is sometimes called the profile log likelihood for σ_0. Under the null hypothesis that σ_0 is the true or actual value of the second Weibull parameter, $W(\sigma_0)$ has, approximately, a χ^2 distribution with $p_{\sigma_0} = dim(\sigma_0)$ (= 1) degrees of freedom. Inverting the test yields a corresponding $1 - \alpha$ confidence region as:

$$\{\sigma_0 : W(\sigma_0) \le \chi^2_{p_{\sigma_0},\alpha}\} \qquad \text{(Eq 8)}$$

where $\chi^2_{p_{\sigma_0},\alpha}$ is the upper α point of the chi-squared distribution with p_{σ_0} degrees of freedom. The procedure of inverting a statistical test to obtain a confidence region is common practice: The set of all σ_0 values that satisfy the acceptance criterion, that is, would not result in a decision of rejection if that particular value had been the null value, are associated with a test statistic value that has a probability of $1 - \alpha$ under the prescribed null hypothesis. But this just defines a $1 - \alpha$ confidence region. Thus, in order to obtain a $1 - \alpha$ confidence region on σ_0, one must find those σ_0 values that satisfy Eq 8.

The strength quantiles for a component, as shown in Fig 1, are denoted by σ_p. In practice, p is generally chosen to be "small," for example, 0.05 or 0.01. It is well known that with mild conditions, the maximum likelihood estimator (MLE) of a function of a parameter(s) is the function of the MLE(s). For example, if $\hat{\theta}$ is the MLE of θ, then the MLE of $f(\theta)$ is $f(\hat{\theta})$. This fact is commonly used to estimate σ_p; see, for example, Ref 9. However, because we also want to determine various kinds of confidence limits via the likelihood ratio method, the somewhat different course of redefining σ_0 into σ_p is followed. Thus, let σ_p be such that

$$p = 1 - \exp\left\{-kV\left(\frac{\sigma_p}{\sigma_0}\right)^m\right\} \qquad \text{(Eq 9)}$$

where k and V are associated with the component of interest. Then, solving for σ_0 yields

$$\sigma_0 = \sigma_p\left[\frac{1}{kV}\ln\left(\frac{1}{1-p}\right)\right]^{-1/m} \qquad \text{(Eq 10)}$$

and substituting Eq 10 into Eq 2 yields

$$P_i(\sigma_i) = 1 - \exp\left\{-r_i R_i \ln\left(\frac{1}{1-p}\right)\left(\frac{\sigma_i}{\sigma_p}\right)^m\right\} \qquad \text{(Eq 11)}$$

where $r_i = k_i/k$ and $R_i = V_i/V$. Equation 11 has the two unknown parameters m and σ_p and relates the component situation to the test situation. For example, the parameter σ_p is the pth quantile at size V associated with the component, appearing in the strength distribution of σ_i for the ith specimen. We note in Eq 11 that r_i is a (known) function of m and also R_i is known; m and σ_p must be estimated by taking data, σ_i. To this end, Eq 11 is employed in the remaining developments of the likelihood ratio procedure. In addition, this methodology can be employed to determine confidence bounds on other parameters. For example, it is straightforward to obtain confidence limits on either m or σ_0. Moreover, a reparameterization analogous to that given by Eq 9 to 11 can be employed to get confidence bounds on the reliability at a given strength. Thus, the likelihood ratio method is very flexible and can be employed to obtain various kinds of confidence limits. The trick is to make a suitable change of variables for the parameter(s) of interest.

The maximum likelihood solution for m is as given previously in Eq 5 with r_i replacing k_i. In the present setup, the maximum likelihood solution for σ_p becomes

$$\hat{\sigma}_p = \left[\ln\left(\frac{1}{1-p}\right)\Sigma r_i(\hat{m})R_i\sigma_i^{\hat{m}}/n\right]^{1/\hat{m}} \qquad \text{(Eq 12)}$$

In order to obtain confidence limits on σ_p, Eq 8 must be solved with σ_0 replaced by σ_p. This requires evaluating $l(\hat{m}, \hat{\sigma}_p)$ and $l(\hat{m}_{\sigma_p}, \sigma_p)$. The method to obtain $\hat{m}$ and $\hat{\sigma}_p$ has already been described. Now, take σ_p fixed so as to determine $\hat{m}_{\sigma_p}$. Inspection of Eq 4 and employing Eq 11 reveals that the solution is

$$0 = \frac{\Sigma_{\hat{r}_i'}}{\hat{r}_i} + \frac{n}{\hat{m}_{\sigma_p}} + \Sigma\ln\sigma_i - n\ln\sigma_p - \ln\left(\frac{1}{1-p}\right)\Sigma\left[\hat{r}_i R_i\left(\frac{\sigma_i}{\sigma_p}\right)^{\hat{m}_{\sigma_p}}\left\{\frac{\hat{r}_i'}{\hat{r}_i} + \ln\left(\frac{\sigma_i}{\sigma_p}\right)\right\}\right] \qquad \text{(Eq 13)}$$

where the circumflexes on the r's indicate that m_{σ_p} has been replaced by $\hat{m}_{\sigma_p}$. Equation 13 is solved in the same manner as Eq 5 in an iterative procedure to find the boundary values of σ_p such that

$$\{\sigma_p : W(\sigma_p) \le \chi^2_{p_{\sigma_p},\alpha}\} \qquad \text{(Eq 14)}$$

is met, that is, replace σ_0 by σ_p in Eq 7 and 8 via the change in density from Eq 2 to Eq 11 to carry out the likelihood ratio procedure to obtain confidence limits on σ_p.

The results of applying the likelihood ratio method to the SiC data are shown in Fig 3. Using Eq 13 and 14 with a p of 0.05 and employing various component areas yields the 95% confidence bounds on the 0.05 quantile shown on Fig 3 as solid triangles. A comparison of the two techniques indicates (in a limited way) that the likelihood ratio and parametric bootstrap methods give confidence limits that practically agree. This is consistent with the general knowledge base for bootstrap methods. These Class IV approaches of analysis combine data from all SiC specimen sizes and geometries, and make strength predictions with confidence bounds on components with arbitrary sizes and geometries. The "component" loading of Fig 3 is that of uniform tension, because no allowance was made for the loading geometry in the calculation of the confidence bounds.

Unfortunately, the likelihood ratio method has one drawback. The limits obtained are generally biased in a similar manner to those described above for the bootstrap method. Note that if the asymptotic distribution used in deriving Eq 8 were exact, then the expected value of $W(\sigma_0)$ would be p_{σ_0}. Sometimes it is possible to find an expansion such that the expected value is $p_{\sigma_0}[1 + c/n + o(1/n)]$. Then $(1 + c/n)$ is called a Bartlett correction factor and improved properties are obtained by replacing W with $W' = W/(1 + c/n)$ in Eq 7 and 8. Frequently, c must be estimated to carry out this correction procedure. As pointed out by Cox and Oakes (Ref 22), it is rarely feasible to carry out such calculations in the presence of censoring. In addition increased complexity of a problem d ishes the ability to determine c.

As mentioned above, it is also possil correct bootstrap limits for bias thı information generated during the boo simulation. However, a general study ha been carried out that compares all the r ods of bias correction with the goal of c mining a preferred way to obtain unb confidence limits. One can imagine cor ing bootstrap confidence limits via Ba corrections, on the one hand, or, on the o correcting likelihood ratio confidence li via a bootstrap procedure.

Future Needs and Directions. A s of complications that are often present in life problems of probabilistic strength a ysis was described in the section "Distr tional Models" in this article. Since the e 1980s, a great deal of progress has been n in the development of methods to acc modate such complications. Unfortunat these methodologies are not compatible w each other in many instances. Therefore several of the complications are pres simultaneously, no overall approach is av able. Additionally, in the presence of m of the complications listed earlier, meth of combining data, estimating confide bounds, and testing for goodness-of-fit either not available or are so computationa intensive that they may not be practical.

The five-year period from 1991–1996 expected to be especially productive addressing such problems of compatibility probabilistic failure analysis. For example, part of the DOE-sponsored program cited Ref 11, two subcontracted efforts are inc pendently addressing problems of probabil tic life prediction. One subcontract is wi Garrett Auxiliary Power Division of Allie Signal, Inc., and the other is with the Alliso Gas Turbine Division of General Motors Cor In both cases, the objective is to develop ar test a ceramic life prediction methodology th simultaneously addresses many of the con plications and requirements described in th article.

In the case of the Garrett effort (Ref 23 a comprehensive methodology is bein developed that will address the entire list c complications. In addition, the methodolog will have the desirable capabilities of analyz ing pooled data and determining confidenc bounds on estimates. This is a very ambitiou undertaking, but there are no obvious road blocks to the planned approach. In paralle with the development of probabilistic tools i a well-designed experimental testing plan. Th resulting data will provide a challengin opportunity for the new tools to digest var ious types of test specimen data, predict the life of a simulated turbine component, and compare the predicted behavior with observed behavior. The ultimate success of comprehensive probabilistic life prediction tools will eventually require validation through numerous successful predictions of component behavior.

REFERENCES

1. W.D. Kingery, H.K. Bowen, and D.R. Uhlmann, Chapt. 14, *Introduction to Ceramics*, 2nd ed., John Wiley & Sons, 1976
2. W. Weibull, A Statistical Theory of the Strength of Materials, *R. Swedish Acad. Eng. Sci. Proc.*, Vol 151, 1939, p 1–45
3. W. Weibull, A Statistical Distribution of Wide Applicability, *J. Appl. Mech.*, Vol 18, 1951, p 293–297
4. E.J. Gumbel, *Statistics of Extremes*, Columbia University Press, 1958
5. K. Trustrum and A. De S. Jayatilaka, On Estimating the Weibull Modulus for a Brittle Material, *J. Mater. Sci.*, Vol 14, 1979, p 1080–1084
6. C.A. Johnson, Fracture Statistics of Multiple Flaw Distributions, *Fracture Mechanics of Ceramics*, Vol 5, Plenum Press, 1983, p 365–386
7. S.B. Batdorf and H.L. Heinisch, Weakest Link Theory Reformulated for Arbitrary Fracture Criterion, *J. Am. Ceram. Soc.*, Vol 61 (No. 7–8), 1978, p 355–358
8. J. Lamon and A.G. Evans, Statistical Analysis of Bending Strengths for Brittle Solids: A Multiaxial Fracture Problem, *J. Am. Ceram. Soc.*, Vol 66 (No. 3), 1983, p 177–182
9. W. Nelson, *Applied Life Data Analysis*, John Wiley & Sons, 1982
10. W.T. Tucker and C.A. Johnson, Modifications in Strength Distributions Due to Slow Crack Growth, *Proceeding of the 24th Automotive Technology Development Contractors' Coordination Meeting*, Publication P-197, Society of Automotive Engineers, Apr 1987, p 233–237
11. "Ceramic Technology for Advanced Heat Engines Project," Contract DE-AC05-840R21400 with Martin Marietta Energy Systems, Inc., Subcontract 86X00223X with General Electric Corporate Research and Development
12. C.A. Johnson and W.T. Tucker, Advanced Statistical Concepts of Fracture in Brittle Materials, *Ceramic Technology for Advanced Heat Engines Project Semiannual Progress Report for April 1987 Through September 1987*, ORNL/TM-10705, Oak Ridge National Laboratory, Mar 1988, p 200–208
13. C.A. Johnson and S. Prochazka, "Investigation of Ceramics for High Temperature Turbine Component," Final Report, Contract N62269–76–0243, June 1977
14. S.B. Batdorf and G. Sines, Combining Data for Improved Weibull Parameter Estimation, *J. Am. Ceram. Soc.*, Vol 63, 1980, p 214–218
15. J.F. Lawless, Construction of Tolerance Bounds for the Extreme-Value and Weibull Distributions, *Technometrics*, Vol 17, 1975, p 255–261
16. N.R. Mann, R.E. Shafer, and N.D. Singpurwalla, *Methods for Statistical Analysis of Reliability and Life Data*, John Wiley & Sons, 1974
17. R.A. Fisher, Two New Properties of Mathematical Likelihood, *Proc. R. Soc. A*, Vol 144, 1934, p 285–307
18. C.A. Johnson and W.T. Tucker, Advanced Statistical Concepts of Fracture in Brittle Materials, *Ceramic Technology for Advanced Heat Engines Project Semiannual Progress Report for October 1987 Through March 1988*, ORNL/TM-10079, Oak Ridge National Laboratory, Aug 1986, p 208–223
19. W.T. Tucker and C.A. Johnson, Confidence Bounds on Strength Estimates, *Proceeding of the 26th Automotive Technology Development Contractors' Coordination Meeting*, Publication P-219, Society of Automotive Engineers, Apr 1989, p 205–210
20. B. Efron and R. Tibshirani, Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy, *Stat. Sci.*, Vol 1, 1986, p 54–77
21. C.A. Johnson and W.T. Tucker, Advanced Statistical Concepts of Fracture in Brittle Materials, *Ceramic Technology for Advanced Heat Engines Project Semiannual Progress Report for October 1989 Through March 1990*, ORNL/TM-11586, Oak Ridge National Laboratory, Sept 1990, p 298–316
22. D.R. Cox and D. Oakes, Chapt. 3.3, *Analysis of Survival Data*, Chapman and Hall, 1984
23. A.M. Comfort and J.S. Cuccio, Life Prediction Methodology for Ceramic Components of Advanced Heat Engines, *Proceeding of the 24th Automotive Technology Development Contractors' Coordination Meeting*, Publication P-230, Society of Automotive Engineers, Apr 1990, p 275–282