APPLICATION OF NONPARAMETRIC METHODS TO RAINFLOW STRESS DENSITY ESTIMATION OF GAS TURBINE ENGINE USAGE

M.P. Enright  
R.C. McClung  
S.J. Hudak  
Southwest Research Institute  
San Antonio, TX 78238  
H.R. Millwater  
University of Texas at San Antonio  
San Antonio, TX 78238

ABSTRACT

The risk of fracture associated with high energy rotating components in aircraft gas turbine engines can be sensitive to small changes in applied stress values which are often difficult to measure and predict. Although a parametric approach is often used to characterize random variables, it is difficult to apply to multimodal densities. Nonparametric methods provide a direct fit to the data, and can be used to estimate the multimodal densities often associated with rainflow stress data. In this paper, a comparison of parametric and nonparametric methods is presented for density estimation of rainflow stress profiles associated with military aircraft gas turbine engine usages. A nonparametric adaptive kernel density estimator algorithm is illustrated for standard parametric probability density functions and for rainflow stress pairs associated with F-16/F100 engine usages. The kernel estimates are compared to parametric estimates, including a hybrid approach based on separate treatment of maximum stress pairs. The results provide some insight regarding the strengths and weaknesses of parametric and nonparametric density estimation methods for gas turbine engines, and can be used to develop improved stress estimates for probabilistic life predictions.

INTRODUCTION

The high energy rotating components in aircraft gas turbine engines may contain rare material anomalies that can lead to uncontained failure of the engine [1,2]. This event is extremely rare (the design target for commercial aircraft is on the order of $1 \times 10^{-9}$ [3]), so a probabilistic damage tolerance approach is used to predict the risk of fracture associated with this event [4,5]. Due to the power law relationship among stress and crack growth life, the fracture risk can be sensitive to small changes in applied stress values [6].

The usage history associated with a commercial flight is typically well established with known stress pair magnitudes. Since there may be some variability in the stress values, a stress scatter factor can be applied to the established values to account for this uncertainty. The influence of stress scatter on the probability of fracture is commonly estimated by applying it repeatedly to the deterministic stress values over the entire life of a disk.

However, for military missions, the number and magnitude of stress pair values may not be well defined. In addition, the variability associated with the stress values may be greater than for commercial missions. Since the stress pair values could vary significantly from flight to flight, a more accurate model is needed to estimate the influence of stress variability on the probability of fracture.

For gas turbine engines, the applied stress is a stochastic variable that is often difficult to measure and predict. A number of researchers have characterized cyclic stress values using rainflow techniques based on the power spectral density (e.g., [7], among many others) or standardized algorithms (e.g., ASTM rainflow computer algorithm [8]). Many of these efforts have focused on simulation of rainflow data from an assumed distribution through application of Gaussian (e.g., [9,10]) and Markov processes (e.g., [11-13]). However, the distribution type must also be identified for use in probabilistic fracture computations associated with damage tolerance-based risk assessment.

Two primary approaches have been proposed for density estimation of rainflow stress values: parametric and nonparametric. The parametric approach consists of the
identificati on of a standard probability density function (PDF) that most closely matches the sample data. The PDF is identified from the statistical descriptors of the data (e.g., moments, Weibull parameters [14]), or from a regression fit of the data plotted on a cumulative probability scale [15-17]. The primary drawback to this approach is that it is often difficult to apply to multimodal densities. On the other hand, nonparametric methods provide a direct fit to the data and are routinely used to estimate multimodal densities. The kernel density estimator, the most commonly used nonparametric method, has been recently applied to rainfall stress data [18,19].

In this paper, a comparison of parametric and nonparametric methods is presented for density estimation of rainfall stress profiles associated with aircraft gas turbine engine usages. A brief summary of the nonparametric adaptive kernel density estimator is presented including details regarding a specific algorithm. Algorithm performance is illustrated using data generated from parametric uni- and multi-variate PDFs and is also compared to parametric estimates based on the same data. The adaptive kernel density estimator is applied to rainfall stress pairs associated with F-16/F100 engine usages and compared to parametric estimates, including a hybrid approach based on separate treatment of maximum stress pairs associated with engine startup and aircraft takeoff.

MULTIVARIATE KERNAL DENSITY ESTIMATION

A number of nonparametric methods can be used to estimate the probability density associated with a data set, such as the nearest neighbor method [20], convex intensity [21], and orthogonal series estimators [22], among many others. Many of the nonparametric methods are related to the histogram, which is commonly used to provide a measure of the relative occurrences of discrete data values. However, as shown in Fig. 1, histograms are highly dependent on the bin width selected for sampling. Furthermore, a histogram is not a true probability density function, and therefore cannot be used for probabilistic computations.

The kernel density estimator is one of the most commonly used nonparametric methods to estimate probability densities. As illustrated in Fig. 2a, each data point is assumed to have an associated probability density within a specified range of values. As shown in Fig. 2b, if data points are relatively close together, their probability densities can overlap. The probability density at a given location can be estimated as the sum of the contributions of the kernels associated with each data point [20]:

$$ \tilde{f}(t_1,t_2) = \frac{1}{nh^2} \sum_{i=1}^{n} K \left( \frac{t_1-X_{1i}}{h}, \frac{t_2-X_{2i}}{h} \right) $$  (1)

where $t_1, t_2 = $ coordinates of evaluation point, $n = $ number of data points, $h = $ window width (smoothing parameter), $K = $ kernel estimator, and $X_{1i}, X_{2i} = $ coordinates of data point $i$.

Figure 1. Bin-width dependency associated with histogram density estimates: (a) bin width = 5 units, and (b) bin width = 10 units.

Eqn. (1) provides an initial density estimate associated with a fixed window width $h$. The window width can be adaptively adjusted to account for the number of data points within a region using the following equation [22]:

$$ \tilde{f}(t_1,t_2) = \frac{1}{nh^d} \sum_{i=1}^{n} \frac{1}{(h\lambda_i)^d} K \left( \frac{t_1-X_{1i}}{h\lambda_i}, \frac{t_2-X_{2i}}{h\lambda_i} \right) $$  (2)

where $\lambda_i$ is a bandwidth parameter that is identified adaptively based on previous PDF estimates:

$$ \lambda_i = \sqrt[2n]{\prod_{i=1}^{n} \tilde{f}(X_i)} $$  (3)
An Epanechnikov kernel function can be used to describe the probability density associated with rainflow stress values [19,22]:

\[
K\left(\frac{t_i - X_{ij}}{h}, \frac{t_j - X_{ji}}{h}\right) = \\
\frac{2}{\pi h^2} \left[ 1 - \frac{1}{h^2} \left( t_i - X_{ij} \right)^2 \left( t_j - X_{ji} \right)^2 \right] \geq 0
\]  

(4)

An adaptive kernel numerical algorithm based on Eqns. (1)-(4) was developed and verified using a number of established probability density functions. For example, consider the univariate lognormal distribution (LN(3,2)) shown in Fig. 3. The initial kernel PDF estimate based on Eqn. 1 (and 200 Monte Carlo samples from the parent distribution) is very similar to the parent PDF near the median, but differs near the right tail of the distribution. However, when the window width is adaptively adjusted using Eqn. (2), it converges to the parent distribution over most of the right tail. The parametric density estimate (LN(2.89, 1.81)) is very similar to the parent PDF near the median, but differs near the right tail of the distribution. Standard algorithms are available to assess the accuracy of the kernel estimate [20] which should be strongly considered before applying the density estimate for risk predictions.

Figure 2. Nonparametric approach to probability density estimation: (a) Epanechnikov Kernal density at individual data points, and (b) combined density of all data points. Comparison of parametric and nonparametric methods for probability density estimation.

Figure 3. Comparison of parametric and nonparametric density estimation methods for a univariate lognormal parent distribution.

Figure 4. The adaptive kernel method was used to estimate the probability density associated with a multivariate distribution (N(100,10)).
Figure 5. Comparison of density estimation methods for a multivariate normal distribution: (a) parametric density estimate, and (b) Kernel density estimate.

APPLICATION TO DENSITY ESTIMATION OF RAINFLOW STRESS PAIRS ASSOCIATED WITH AIRCRAFT ENGINE USAGES

The adaptive kernel algorithm was applied to the density estimation of rainflow stress pairs associated with actual aircraft gas turbine engine usages. Raw RPM values were obtained from engine flight data recorder data for a variety of F-16/F100 missions. RPM values are shown in Fig. 6 for representative flights of missions classified as “Live Fire” and “Instruments & Navigation”. The RPM values were converted to stress values using an empirical algorithm [23] including a shakedown procedure to address stress values that fall above the material yield stress [24]. The stress values were sorted into min-max pairs using an established rainflow cycle counting algorithm [8].

As shown in Fig. 6, a typical engine usage history can consist of thousands of RPM data points. However, the data can often be reduced to a few hundred values following application of the rainflow cycle counting algorithm. For example, the rainflow stress pairs associated with each of the individual flights shown in Fig. 7 are based on usages similar to the ones shown in Fig. 6. Most of the stress pairs are clustered in a region of minimum and maximum stress values ranging from 350 to 750 MPa. However, a few stress pairs associated with the aircraft major stress cycle are located in the region where the minimum stress is zero (associated with engine startup and aircraft takeoff).

If the stress pairs can be completely described by a probability density function, then the influence of usage variability on the probability of fracture can be quantified by sampling directly from the rainflow stress PDF. The desired PDF must provide a probability density for each region of the population that is consistent with the clustering of data points shown in Fig. 7. It must attempt to describe the population from which the data points were drawn to reflect the influence of data points not included in the current sample.

A parametric density estimate of the rainflow stress pairs associated with the Live Fire mission is shown in Figs 8a and 9a. The unimodal density does not appear to include the influences of the stress pairs associated with startup/takeoff. It also does not capture the correlation behavior among minimum and maximum stress values suggested in Fig. 7a.

The stress pairs associated with startup/takeoff could be considered separately from the remaining data because they must be applied only once per flight (this cannot be guaranteed if they are included in the sampling population). In addition, since the minimum stress values associated with these stress pairs are always zero (to reflect the zero RPM value prior to engine startup), they can be modeled as a univariate (and probably unimodal) distribution. A parametric probability

Copyright © #### by ASME
CONCLUSIONS
A comparison of parametric and nonparametric methods was presented for density estimation of rainflow stress profiles associated with military aircraft gas turbine engine usages. When the entire sample population is considered, the kernel method appears to provide a more realistic estimate compared to the parametric approach. It provides explicit treatment of the stress pairs associated with engine startup, and addresses the correlation behavior among the minimum and maximum stress values. However, when the parametric method is applied to a sub-population of Live Fire rainflow stress pairs that does not include values associated with engine startup, it also appears to capture the correlation structure of the minimum and maximum stress values. The results can be used to develop improved stress estimates for probabilistic life predictions.

ACKNOWLEDGEMENTS
This work was performed with funding from the U.S. Air Force Research Laboratory and Southwest Research Institute under AFRL Dual Use Agreement No. F33615-03-2-5203. The support and encouragement of Jay Jira, Jim Larsen, and Patrick Golden (technical monitor) of AFRL is gratefully acknowledged.

REFERENCES
Figure 8. Comparison of density estimation methods for rainflow stress pairs associated with Live Fire usage histories: (a) parametric density (all stress pairs), (b) parametric density (no engine startup stress pairs), and (c) kernel density estimate.

Figure 9. Contour plots comparing density estimation methods for Live Fire rainflow stress pairs: (a) parametric density (all stress pairs), (b) parametric density (no engine startup stress pairs), and (c) kernel density estimate.


