SAND2016-0859J

Construction of Tolerance Bounds for a Multivariate Response Associated with a Covariate: A Case Study

E. V. Thomas, C. King, J. Cap, and A. Montoya
Sandia National Laboratories, Albuquerque, NM 87185

Abstract

The purpose of this paper is to describe and illustrate practical methods that can be used to construct tolerance bounds for a multivariate measurement associated with a covariate. These methods rely on principal components analysis and the parametric bootstrap. The methods are illustrated with an example in which the vibration environment experienced by a test object being carried by an aircraft (known as captive carry) is characterized. In this example, the dominant dimension extracted by principal components analysis is related to a measureable covariate (dynamic pressure). Utilizing this relationship, tolerance bounds are constructed that are specific to a target value of the covariate. The multivariate measurement (vibration amplitude) relates to an ordered index (frequency). However, the ordering of the multivariate response is not useful for constructing tolerance bounds due to the discontinuous nature of the vibration amplitude across the range of frequencies. Thus, the methods that are described here are simple and do not assume an underlying functional relationship of the response across the ordered index.

Keywords: Bootstrap, Captive Carry, Principal Components Analysis

Introduction

A tolerance bound is a data-dependent confidence limit on a specific quantile of a characteristic of interest (Wald and Wolfowitz (1946) and Krishnamoorthy and Mathew (2009)). Construction of a tolerance bound often involves the need for a statistical model. Tolerance bounds are used across a wide variety of situations including making a reliable assessment of exposure to hazards in the workplace (Tuggle (1982)), characterizing the distribution of groundwater constituents (Gibbons (1991)), and assessing the sensitivity of explosives to an ignition stimulus (Neyer (1994)). In manufacturing applications, tolerances bounds are frequently utilized in acceptance sampling contexts as evidence for accepting or rejecting a production lot (Montgomery (2012)). Consider the case where one is interested in the weld strength of a particular component, where the weld strength varies from unit to unit within the lot. In such a situation, one might want to develop a lower tolerance bound for a specified percentile of weld strengths for the production lot. Alternatively, one might be interested in an upper tolerance bound or in a tolerance interval (involving both upper and lower bounds). In this case, weld strength measurements conducted on a representative sample of units from the population of interest can serve as the basis for

constructing a lower tolerance bound. Suppose a sample of size n is selected from a population for testing resulting in estimates of the population mean and standard deviation of the characteristic of interest, given by $\bar{X}$ and S. Assuming that the population characteristic is normally distributed, a lower tolerance limit which bounds $(1-\alpha)\cdot 100\%$ of the population with $\gamma \cdot 100\%$ confidence is given by $LTL = \bar{X} - k_{1-\alpha,\gamma;n} \cdot S$, where $k_{1-\alpha,\gamma;n}$ is a factor computed to ensure accurate coverage. The lot will be accepted if the LTL exceeds a required level. Thus, by construction, a tolerance bound accounts for the variability within the population (through specification of a percentile) and the uncertainty associated with a small sample size (through specification of a confidence level).

Tolerance bounds usually relate to an individual (univariate) characteristic of interest, such as weld strength in the above example. The construction of tolerance bounds in a multivariate context is less common. Recently, Rathnayake and Choudhary (2015) described a method for constructing "tolerance bands" for multivariate functional data. Functional data often take the form of repeated measurements over time but also may involve multivariate data collected across some other ordered index such as frequency (or wavelength). In the functional data context, an observation consists of an ordered multivariate measurement acquired from an object (or subject). One characteristic of functional data is that the multivariate measurement is assumed to consist of noise superimposed on a smooth function of the ordered index (Ramsay and Silverman (1997)). Other recent examples where functional data analysis was used to characterize multivariate data include Storlie et al. (2013) and Tucker et al. (2013). Rathnayake and Choudhary (2015) and Storlie et al. (2013) represent the underlying function as a Gaussian process. In Storlie et al. (2013) the underlying function is also modelled by B-spline basis functions. Functional principal components analysis (Ramsay and Silverman (1997)) can also be useful. However, in order for such approaches to be effective and provide additional value, it is important that the underlying relationship be relatively smooth / continuous.

As described in Rathnayake and Choudhary (2015), the ordered multivariate data can also be classified as either "sparse" or "dense". The distinction between the two cases lies in the regularity of the ordered index across the set of observations. In addition, the multivariate data can be described in terms of "balance", which reflects the consistency to which the observations are represented across all values of the ordered index. In this paper we consider an example in which the objective is to develop tolerance bounds for an ordered multivariate measurement that is both dense and balanced, but not particularly smooth and so does not fit the functional data paradigm. Nevertheless, the correlation is strong across the various dimensions of the response. A simple version of principal components analysis is able to provide a useful basis for constructing both pointwise and simultaneous tolerance bounds. Furthermore, the dominant dimension extracted by principal components analysis is interpretable and strongly related to a measureable covariate. The statistical uncertainty of the relationship between scores of the

dominant dimension and the covariate is incorporated into the constructed tolerance bounds specific to a target value of the covariate.

The remainder of the paper discusses details of the example as well as the methods used to construct tolerance bounds.

Illustrative Example

Over the course of a lifetime, munition systems developed for the military are exposed to various sources of mechanical vibrations that can affect system performance. A significant exposure to vibration environments occurs when the munition is being carried by an aircraft (known as "captive carry"). During captive carry, the vibration level can be affected by flight conditions such as airspeed and altitude. Often, as in the case of the example presented here, it is of interest to characterize the distribution of vibration levels that can be encountered during "straight and level" flight conditions, which is the predominate source of vibration exposure. In the illustrative example presented here, data were acquired from sensors located on various positions on a test system mounted on an aircraft as it was flown over a variety of flight segments. The flight segments are considered to be a representative, random sample of "straight and level" captive carry conditions. For each sensor, the vibration data for each flight segment were summarized by the Acceleration Spectral Density (ASD). The multivariate ASD's or "spectra" are known to depend on dynamic pressure (Q) and Mach Number (M). The dynamic pressure is a particularly useful summary measure of the aerodynamic forces acting on an object as it moves through the atmosphere (Piersol (1971)).

Figure 1 displays a set of ASD's (log scale) that were acquired during 18 "straight and level" flight segments across a broad range of Q values in the transonic Mach regime from a sensor of interest. The transonic Mach regime, defined by $0.9 < Mach < 1.0$, is of particular interest since it is known to produce the worst-case levels of vibration. It will be shown that within the ensemble of spectra displayed in Figure 1, there is a large predictable effect of Q such that the ensemble of data can be sensibly normalized to a target value of Q representing a specific worst-case setting for vibration. The normalized ensemble can then be used to construct an upper tolerance bound specific to a target dynamic pressure of interest, denoted by Q_{target}. The remainder of this document describes the assumptions and details used to develop this upper bound.

First, it is interesting to examine the nature of the spectra displayed in Figure 1. One observation is that the spectra do not appear to be smooth (note the numerous sharp peaks and valleys in the spectra). A second observation is that the features are generally well aligned with respect to frequency. Hence, there appears to be, at most, negligible registration error.

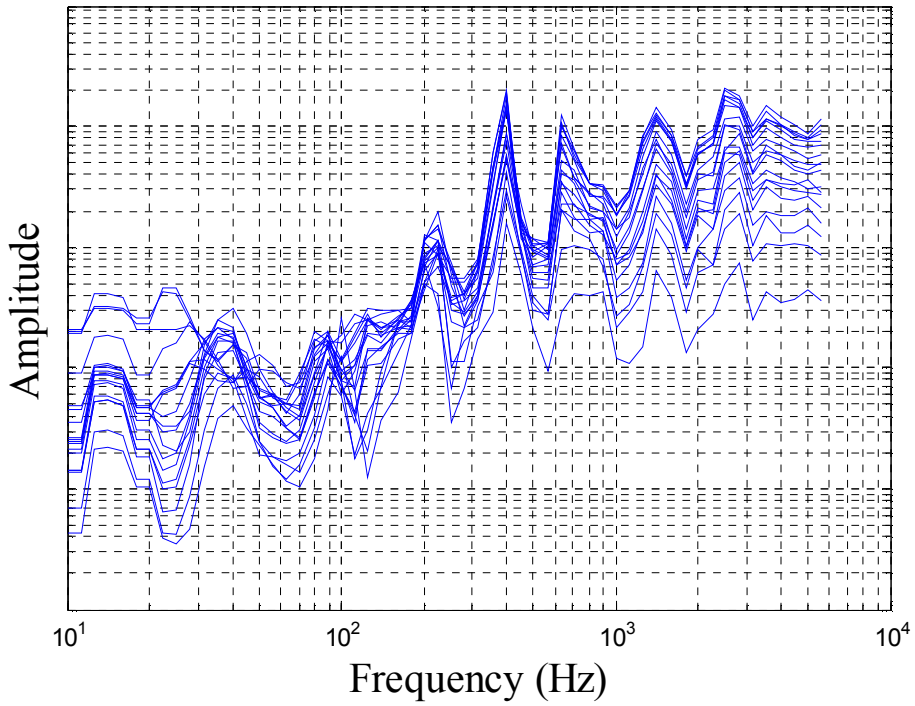

Figure 1 – Acceleration Spectral Densities (Spectra)

Based on these observations, we have developed methods for constructing tolerance bounds that are based on a version of principal components decomposition of an assumed representative set of spectra similar to what is described by Paez, et al. (2004). In the random shock literature (see e.g., Bellizzi and Sampaio (2009)), this technique is known as the Karhunen-Loeve transform (KLT). Note that if the underlying spectra appeared to be smooth and/or were affected by registration error, it would have been prudent to characterize the spectra using the functional data analysis methods discussed earlier. The KLT projects the log-transformed high-dimensional spectral data onto a smaller dimensional orthogonal space defined by latent variables, which facilitates more feasible analysis. The results from the analysis in this lower-dimensional space can easily be translated to the original dimensions as long as the lower-dimensional space adequately represents the original spectra.

The following nomenclature is used to help describe the process.

Y = matrix of spectra (number of spectra *by* channel dimension)

Y_c = matrix of mean-centered spectra (number of spectra *by* channel dimension)

$Y_c = Y - \bar{Y}$, where $\bar{Y}$ is the average spectrum

E = matrix of eigenvectors (channel dimension *by* number of latent variables)

e_k = k^{th} column of E

S = matrix of scores (number of spectra *by* number of latent variables)

$$S = Y_c \cdot E$$

s_k = k^{th} column of S

λ = vector of singular $values$ ($\lambda_1 \geq \lambda_2 \geq \cdots > 0$)

In summary, the KLT provides an approximation to the spectra via an orthonormal set of eigenvectors (shape-vectors) and an orthogonal set of scores (i.e., $\hat{Y} = Y_c + \bar{Y}$, where $Y_c = S \cdot E^T$). Typically, the row dimension of E is much smaller than the number of spectra; hence, the dimension reduction. Singular-value decomposition (SVD) is a method commonly used to obtain E from Y_c (see Golub and Van Loan (1996)).

The SVD can produce an effective basis set E if it is applied to a set of spectra that are representative and are consistent in terms of modes of variation, such as is the case in Figure 1. Figure 2 displays the mean-centered spectra associated with the 18 observations from Figure 1. The SVD was applied to the set of mean-centered spectra in order to produce an orthogonal basis of eigenvectors and scores. Each latent-variable dimension is also associated with a singular value λ_k. The magnitude of each singular value reflects the contribution of its associated dimension in modeling the total variation across all channels of the spectra. Hence, the singular values can be used to determine the importance of each eigenvector. While the SVD is capable of producing a decomposition that has as many latent variables as there are channel dimensions, the intent is to develop a basis set with a smaller dimension that also provides a good approximation to the mean-centered spectra. A graphical means to identify an appropriate number of latent variables is via a scree plot (Cattell (1966)). Figure 3 illustrates the scree plot produced from the SVD of the 18 observations. The scree plot displays the relationship between the logarithm of each singular value versus its order. Theory underlying the scree plot suggests that important, structural dimensions lie above a hypothetical line constructed to fit the least important dimensions. Thus, in this case there appears to be a single structural dimension. Figure 4 displays the eigenvector associated with the largest singular value. The score values associated with this structural dimension clearly depend on Q (see Figure 5). In fact, a simple linear model is able to accurately relate the score-value to Q. From Figures 4 and 5, it is clear that vibration levels increase as Q increases, with the rate of increase depending on frequency. The other dimensions may be related to random measurement error or some other non-structural process.

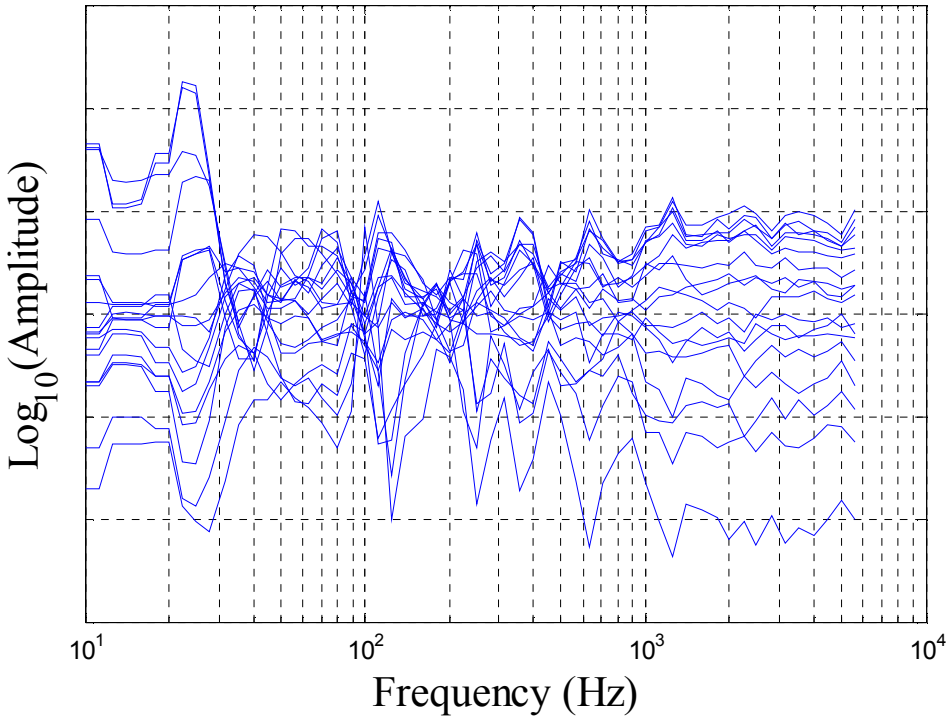

Figure 2 – Mean-Centered Spectra

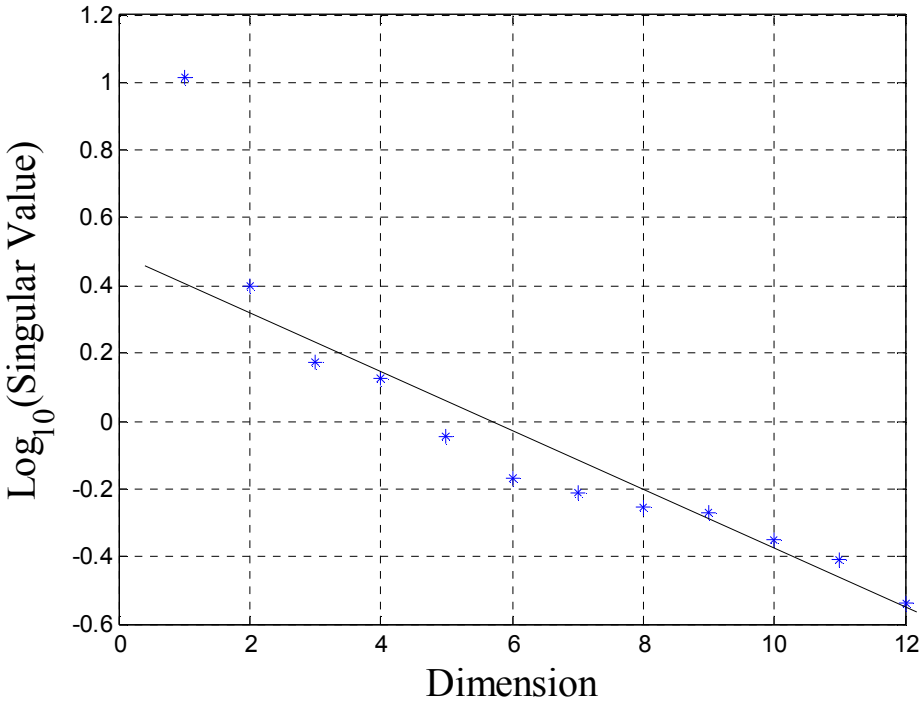

Figure 3 – Scree Plot

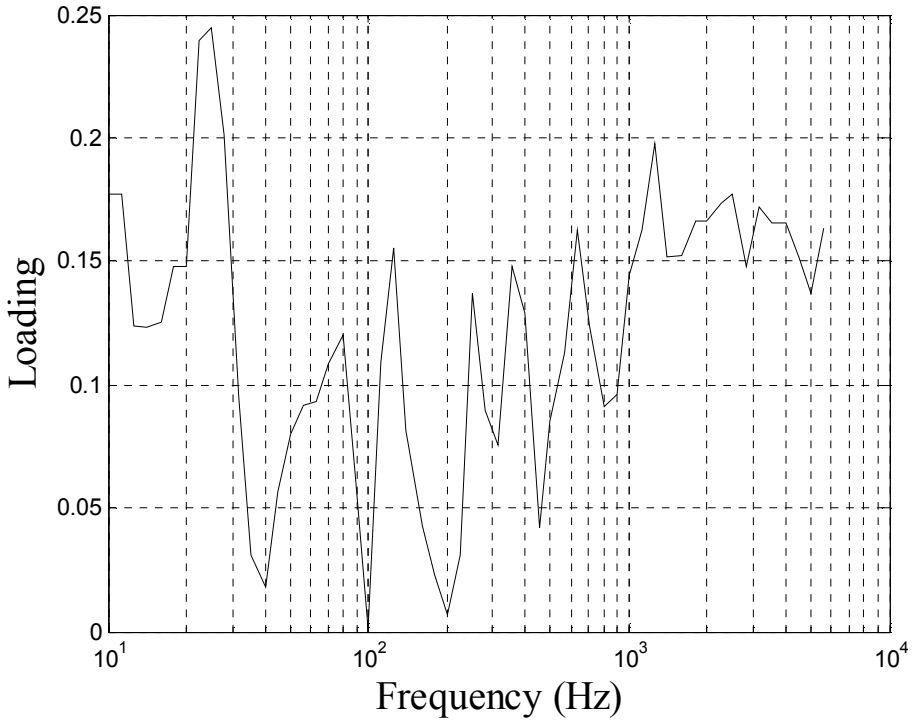

Figure 4 – Eigenvector Associated with Largest Singular Value

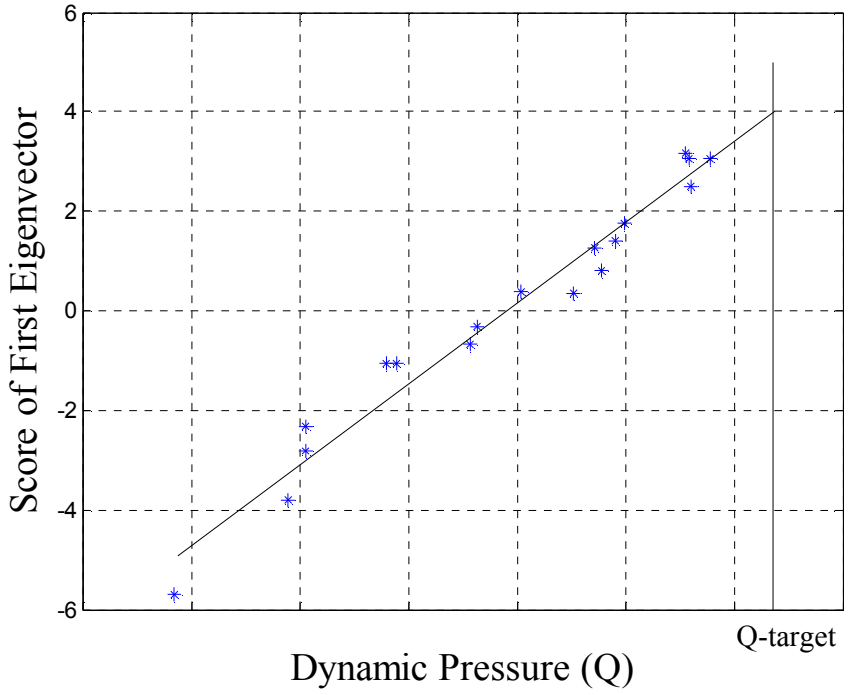

Figure 5 – Scores Associated with Eigenvector #1 versus *Q*

For the analysis that follows, additional dimensions are included to allow for a closer approximation to the spectra (i.e., $\widehat{Y} = \bar{Y} + S \cdot E^T$). Figure 6 compares the levels of variability across the original (Y) and reconstructed spectra ($\widehat{Y}$) based on using SVD with 5 latent variables. It is clear that the levels of variability of the original and reconstructed spectra are similar. For computation of tolerance bounds, we use 10 latent variables, which allows for an even closer approximation. In other situations, one might specify the number of latent variables to include as the smallest set that accounts for some target percentage of the total spectral variation (say 95%).

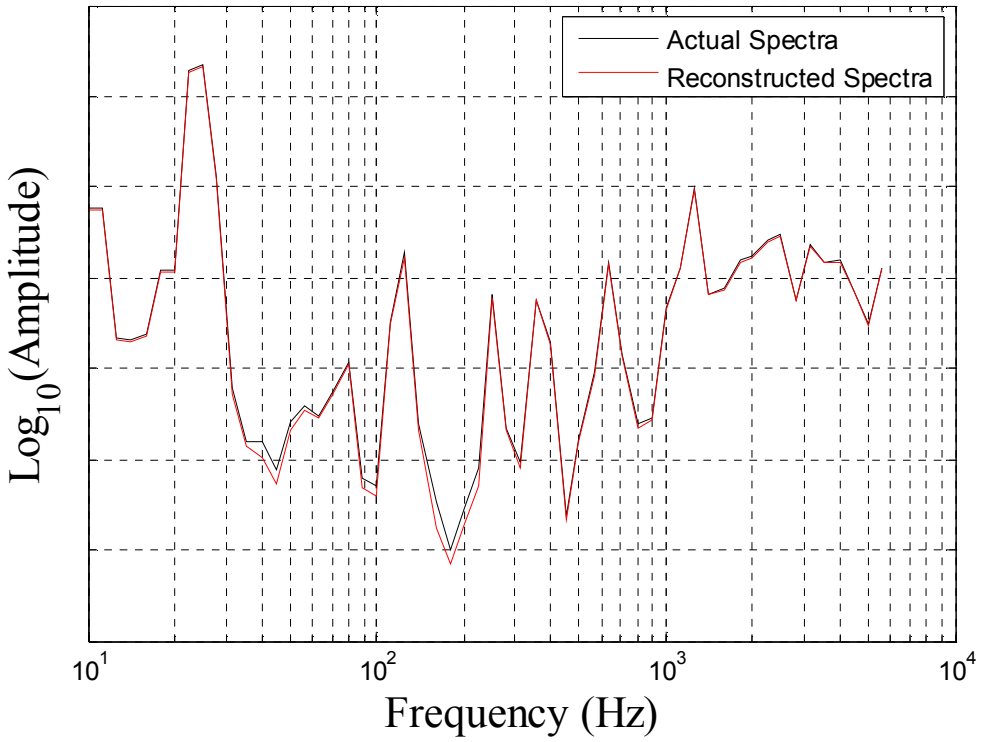

Figure 6 – Standard Deviations of Spectra (Actual and Reconstructed)

Since the first dimension (eigenvector #1) accounts for a substantial portion of the variation across the spectra and the associated scores relate well with Q, it follows that Q has a dominant influence on the spectra. Thus, it is not surprising that the shape of the standard deviation of the spectra (Figure 6) closely approximates eigenvector # 1.

The fitted relationship between the scores of the first dimension and Q (Figure 5) can be used to normalize the spectra to a particular target value of Q.

To effect this normalization, the approximated spectra $\widehat{Y} = \bar{Y} + S \cdot E^T$ are adjusted by correcting the values in the first column of S to account for the difference in the expected value between the observed values of Q and the targeted value of Q, given by Q_{target}. That is,

$\hat{Y}_{adj} = \bar{Y} + S \cdot E^T - \hat{\beta}_1 \cdot \left(Q - Q_{target}\right) \cdot e_1^T$, where Q represents the column vector of values of dynamic pressure and $\hat{\beta}_1$ is from the fitted model displayed in Figure 5 ($\hat{S} = \hat{\beta}_0 + \hat{\beta}_1 \cdot Q$). Thus, the adjusted spectra given by $\hat{Y}_{adj}$ and displayed in Figure 7 have been normalized to resemble an ensemble of spectra associated with Q_{target}.

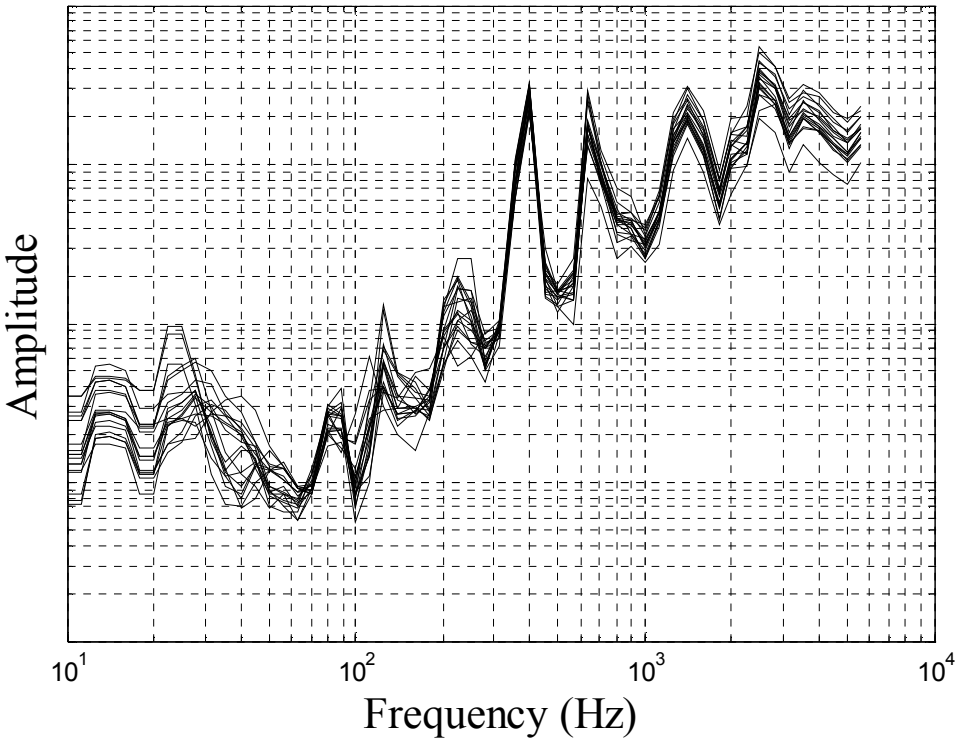

Figure 7 –Spectra Adjusted to Q_{target}, $\hat{Y}_{adj}$

Use of the Parametric Bootstrap Method for Generating Pointwise Tolerance Bounds

The parametric bootstrap can be used to develop pointwise tolerance bounds that are specific to a target level of dynamic pressure. The "parametric bootstrap" is a variant of the general bootstrap procedure (see Efron and Tibshirani (1993) and Hinkley (1988)) and assumes the form of the probability distribution associated with the observed variation in the adjusted spectra. The method assumes that the sampled flight segments are a random sample of "straight and level" captive carry conditions. It is also assumed that an appropriate basis set has been determined such that *p* latent variables are sufficient to adequately represent the spectral variation. For this analysis, it is assumed that the score values are distributed as Gaussian random variables for each of the *k*=1:*p* dimensions. The parametric bootstrap procedure is an iterative procedure such that, at each of *B* iterations, plausible values for the true (unknown) parameters of the assumed normal distributions (σ_k and μ_k) are simulated for *k*=1:*p* based on the estimate $\hat{\sigma}_k$. That is, $\tilde{\sigma}_k \sim \hat{\sigma}_k \cdot \sqrt{\frac{df}{\chi^2_{df}}}$, where χ^2_{df} is a chi-squared distributed random variable with *df* degrees of freedom, and $\tilde{\mu}_k \sim Normal\left(0, \tilde{\sigma}_\mu^2\right)$. First, note that $\hat{\sigma}_1^2$ represents 1/16th of the sum of the squared differences between the observed data and solid curve in Figure 5 ($df = 16$). In the case of the other dimensions ($k > 1$), $\hat{\sigma}_k$ is simply the standard deviation of the scores associated with the k^{th} dimension ($df = n - 1 = 17$). In the case of μ_k, the variance $\tilde{\sigma}_\mu^2 = \hat{\sigma}_k^2 \cdot x_0^T \cdot (X^T \cdot X)^{-1} \cdot x_0$ for $k = 1$ and $\tilde{\sigma}_\mu^2 = \frac{\tilde{\sigma}_k^2}{n}$ for $k > 1$. The first dimension ($k = 1$) is treated differently here since it is demonstrably related to *Q*. *X* is the so-called "design matrix" with 18 rows and 2 columns. The first column of *X* consists of ones. The second column of *X* consists of the values of *Q*. The column vector, x_0 (reflecting a dynamic pressure of Q_{target}), is defined by $\left[1\ Q_{target}\right]^T$.

For each bootstrap iteration, a set of separate plausible parameter pairs are generated independently: $\{\tilde{\mu}_k, \tilde{\sigma}_k\}_{k=1:p}$. This set of parameter pairs could be used as the basis for simulating a set of scores which in turn could be used to simulate individual spectra. Instead, the (1-α) level percentile for the spectrum (point wise) is constructed for that particular set of parameter pairs as follows.

For the i^{th} bootstrap iteration (i^{th} set of parameter pairs), the mean bootstrap spectrum is constructed as $\bar{Y}^i_{boot} = \bar{Y}_{adj} + [\tilde{\mu}_1 \quad \cdots \quad \tilde{\mu}_p] \cdot E^T$. For the same iteration, the standard deviation of the spectrum is reconstructed point wise. Due to the assumed independence of the score values (based on their orthogonal construction), the standard deviation of the j^{th} channel is $S^i_{boot}(j) = \sqrt{\sum_{k=1}^{p} E_{jk}^2 \cdot \tilde{\sigma}_k^2}$. The (1-$\alpha$) percentile for the j^{th} spectral channel for the i^{th}

bootstrap iteration is $P^i_{boot}(j) = \bar{Y}^i_{boot}(j) + z_{1-\alpha} \cdot S^i_{boot}(j)$, where $z_{1-\alpha}$ is the (1-α) percentile of the standard normal distribution. The γ -percentile of the values within $\{P^i_{boot}(j)\}_{i=1:B}$ is the γ-level upper confidence bound for the (1-α) percentile of the j^{th} spectral channel. Figure 8 presents the 90% upper confidence bounds for the 99th percentile (pointwise, by frequency) specific to Q_{target}. The appendix provides the MATLAB code that was used to produce the tolerance bounds illustrated in Figure 8.

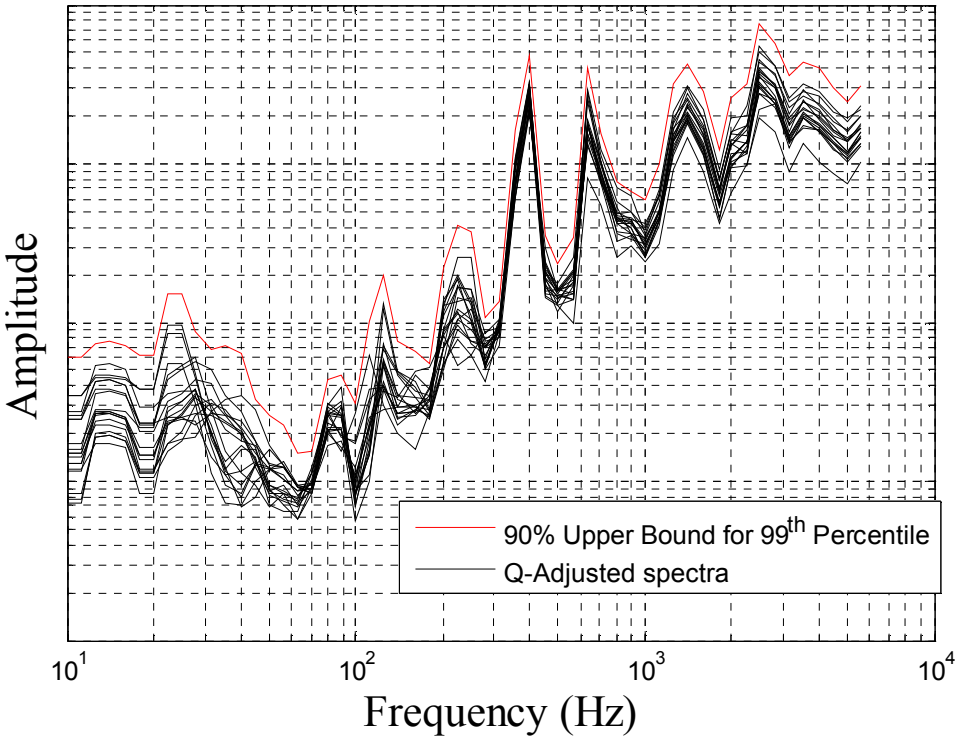

Figure 8 – 90% Upper Confidence Bounds for the 99th Percentile at Q_{target}

It is important to evaluate the accuracy of the computed tolerance bounds obtained by using this method. A straightforward way to do this is as follows. First, assume that the model forms used are accurate. Furthermore, assuming that the estimates of the model parameters reflect the truth, one can compute the pointwise 99th percentiles for the ASD frequencies conditioned at Q_{target}. Note that with these assumptions, the (1-α) percentile for the j^{th} spectral channel is

$$P_{1-\alpha}(j) = \bar{Y}_{adj}(j) + z_{1-\alpha} \cdot \hat{\sigma}(j), \text{ where } \hat{\sigma}(j) = \sqrt{\Sigma_{k=1}^{p} E_{jk}^2 \cdot \hat{\sigma}_k^2}.$$

Then, simulate the set of n spectra by adding random amounts of each the $p = 10$ eigenvectors (according to the assumed distributions of scores) to the estimated models evaluated at the

observed values of Q (i.e., $\bar{Y}_{adj}(j)$). Next, using the simulated set of data, redo the complete analysis (principal components analysis, model fitting, and construction of pointwise tolerance bounds). Then, repeat the simulation and analysis processes a large number of times and compare the constructed 99/90 upper tolerance bounds with the computed 99th percentiles for the ASD frequencies at Q_{target}. Figure 9 displays the set of tolerance bounds that were acquired through 10,000 simulation trials involving the parametric bootstrap. If the method provides accurate coverage, the 99th percentiles of the ASD frequencies will be exceeded 90% of the time by the tolerance bounds. Figure 10 compares the 10th percentile of the 99/90 upper bounds with the 99th percentiles of the assumed distribution of ASD's. It is indeed the case that the 10th percentiles of the 99/90 upper tolerance bounds are nearly equivalent to the 99th percentiles. Figure 11 displays the coverage levels for the various frequencies that were estimated by the simulation. The levels of coverage range from 79% to 93%. Lower than expected coverage was observed at some frequencies from 30 to 200 hz. It is interesting that frequencies with lower than expected coverage generally are associated with relatively small loadings of the first principal component (compare Figures 4 and 11).

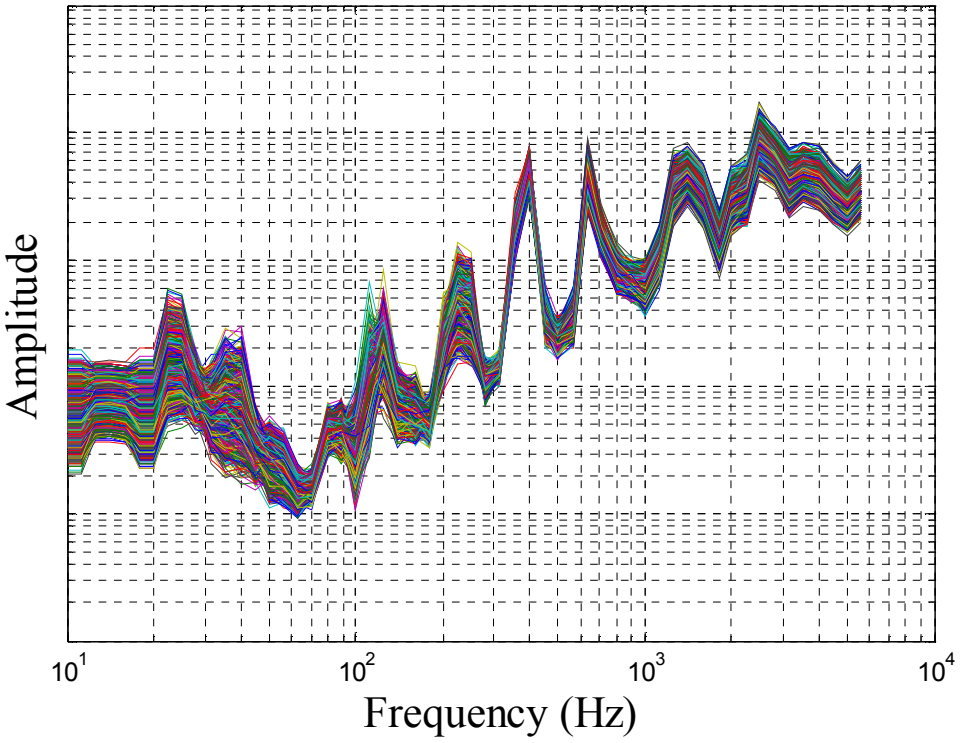

Figure 9 – Simulated 99/90 Upper Tolerance Bounds

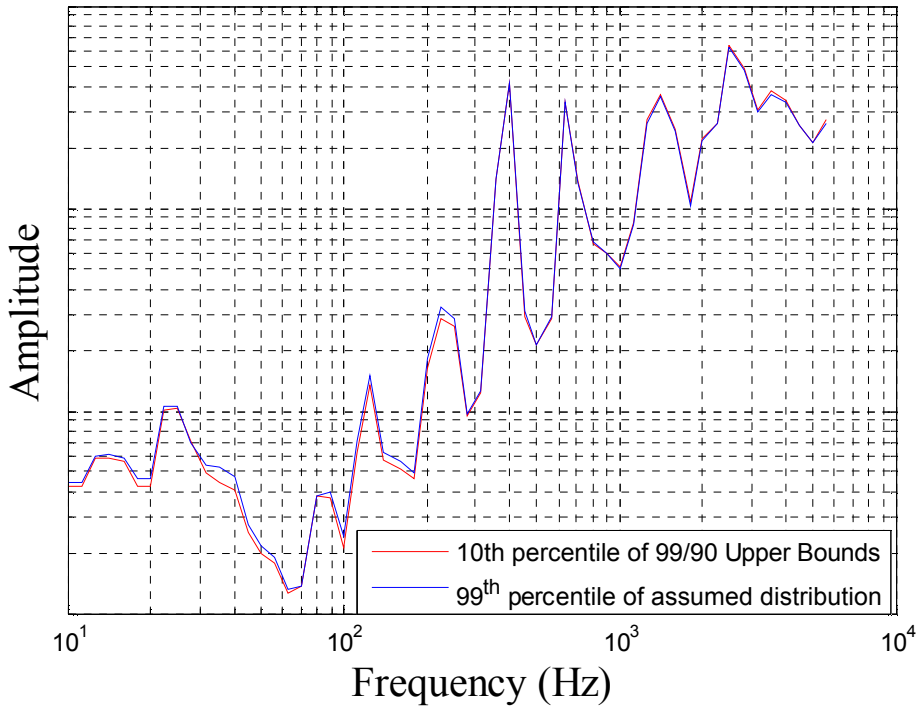

Figure 10 – Comparing the 10th percentile of the 99/90 upper bounds with the 99th percentiles

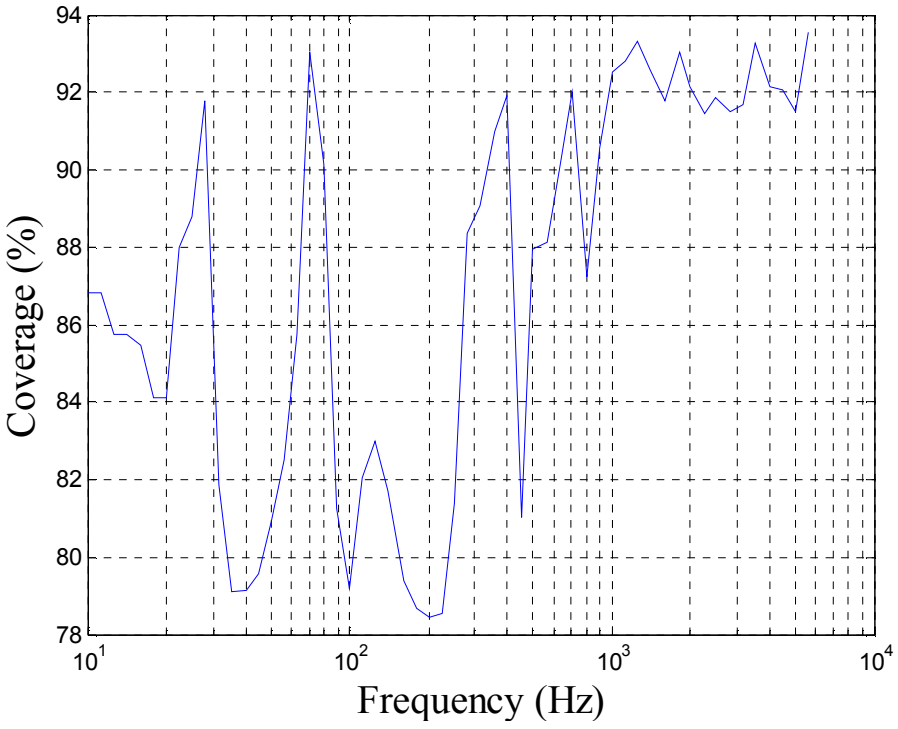

Figure 11 – Estimated levels of coverage, by frequency

To further investigate the issue of marginal lower than expected coverage an additional simulation experiment was conducted. The intent is to simulate a simpler system; one that does not include the effects of Q and by association, the first principal component. A set of n spectra was constructed by adding random amounts of all but the first of the $p = 10$ eigenvectors (according to the assumed distributions of scores) to the estimated model spectrum at Q_{target}. Next, using the simulated set of data, redo the complete analysis except for the adjustment of spectra at Q_{target} and modeling of the effect of Q. Then, repeat the simulation and analysis processes 10,000 times and compare the constructed 99/90 upper tolerance bounds with the computed 99th percentiles for the ASD frequencies at Q_{target}. Figures 12 and 13 summarize the results. In this case, 10th percentiles of the 99/90 upper tolerance bounds are almost identical to the 99th percentiles (see Figure 12). This is not surprising, since there is very little spectral variation due to dimensions 2-10. However, as Figure 13 shows, the 99/90 upper tolerance bounds uniformly provide slightly less coverage than expected.

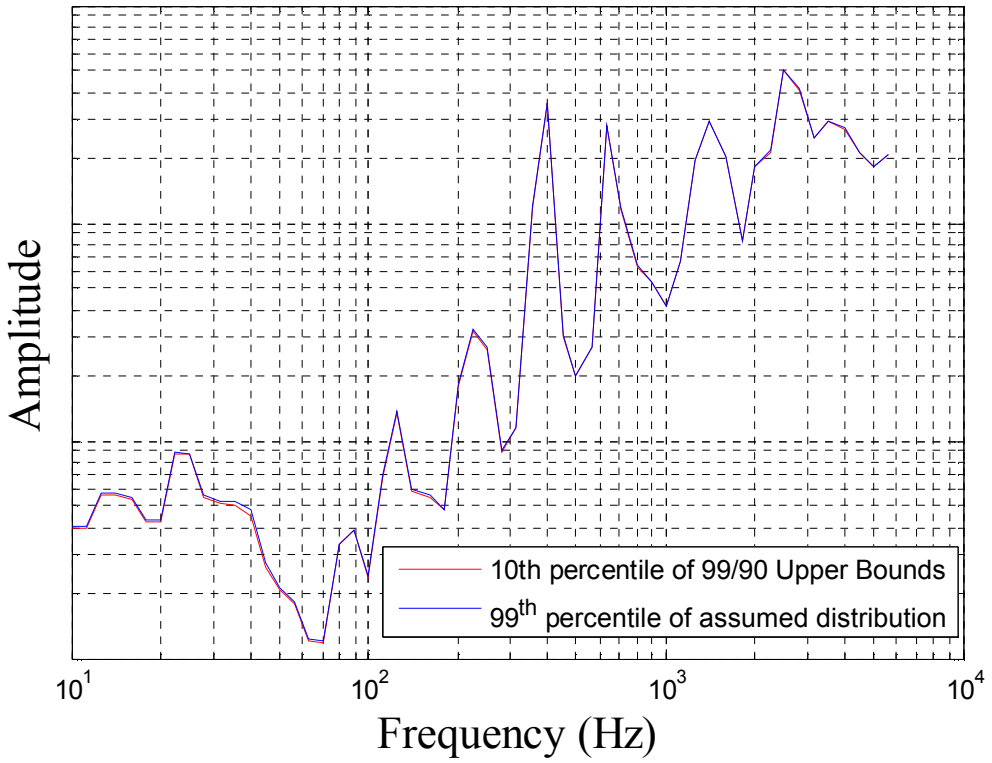

Figure 12 – Comparing the 10th percentile of the 99/90 upper bounds with the 99th percentiles: without first principal component

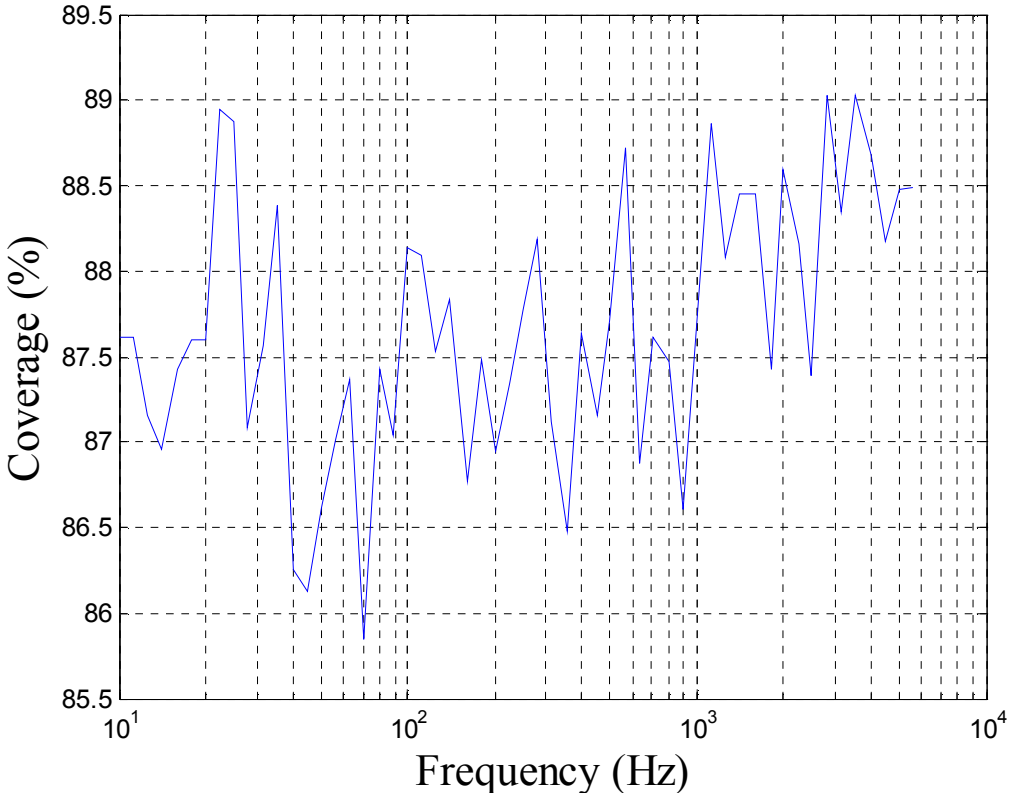

Figure 13 – Estimated levels of coverage, by frequency: without first principal component

These idiosyncrasies and risks associated with using the parametric bootstrap in conjunction with principal components analysis are not unknown. For example, Milan and Whittaker (1995) warn about extra nuisance variability when using the parametric bootstrap in association with principal components analysis. The extra nuisance variability can reveal itself in: (1) a rotation of the principal components, (2) an inversion of the ordering of the singular values, and (3) an inversion of the signs of the eigenvectors (and associated scores). In the example presented here, the effects of the extra nuisance variability on the accuracy of the constructed tolerance bounds are mostly inconsequential as shown in Figures 10 and 11. This is largely due to the manner in which the analysis is performed and the fact that the covariate relates to the "first" principal component which is dominant and hence accounts for most of the spectral variability. The nature of the first eigenvector (particularly where the loadings are large) will not be greatly distorted by spurious correlation of the first dimension with the other dimensions as created by the bootstrap iterations. Also note that any inversion of the sign of the scores associated with the first eigenvector will be properly accounted for in the regression of those scores on Q. In the illustrative example, the dominance of the first principal component is believed to be responsible for the accurate coverage that is observed in the case of frequencies where the loadings of the first principal component are large.

Here, the focus has been on frequency-specific, pointwise tolerance bounds. Attempts to construct accurate joint tolerance bounds were unsuccessful. Even a Bonferonni-type approach is difficult to implement due to the frequency-dependent coverage and the idiosyncrasies associated with using the parametric bootstrap in conjunction with principal components analysis.

Conclusions

A parametric bootstrap procedure has been developed for constructing upper tolerance bounds relating to the acceleration spectral density associated with captive carry for a fixed condition defined by a target dynamic pressure. These methods rely on the ability to create an ensemble of adjusted spectra that is normalized to a specific dynamic pressure of interest via the strong empirical relationship between the acceleration spectral density and dynamic pressure. The method has been found to be applicable to similar situations where one or more covariates relate to scores associated with a dominant principal component. Furthermore, the parametric bootstrap procedure was shown to provide reasonably accurate coverage given that the distributional assumptions are met. The method is relatively simple and straightforward and do not assume continuity across the ordered index. Note that a different and more complex approach (utilizing functional data analysis methods) might be more appropriate if the multivariate data are continuous across the ordered index.

References

Bellizzi, S., & Sampaio, R. (2009), "Smooth Karhunen–Loève decomposition to analyze randomly vibrating systems," *Journal of Sound and Vibration*, *325*(3), 491-498.

Cattell, R. B. (1966), "The scree test for the number of factors," *Multivariate behavioral research*, *1*(2), 245–276.

Efron, B. and Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, New York: Chapman Hall, New York.

Gibbons, R. D. (1991), "Statistical Tolerance Limits for Ground-Water Monitoring," *Groundwater*, *29*(4), 563-570.

Golub, G. H. and Van Loan, C. F. (1996), *Matrix Computations, Third Edition*, The Johns Hopkins University Press.

Hinkley, D. V. (1988) "Bootstrap methods," *J. R. Statist. Soc. B*, 50, 321-337, 1988.

Krishnamoorthy, K. and Mathew, T. (2009), *Statistical tolerance regions: theory, applications, and computation*, New York: John Wiley.

Milan, L. and Whittaker, J. (1995), "Application of the parametric bootstrap to models that incorporate a singular value decomposition," *Applied Statistics*, 31-49.

Montgomery, D. C. (2012), *Introduction to Statistical Quality Control*, 7th Edition, New York: John Wiley.

Neyer, B. T. (1994), "AD-Optimality-Based Sensitivity Test," *Technometrics*, *36*, 61-70.

Paez, T. L., Morrison, D. and Cap, J. S. (2004), "Generation of Nonstationary Random Excitations with a Specified Tolerance Limit," *9th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability*, SAND2004-4301C, Sandia National Labs, Albuquerque New Mexico.

Piersol, A. G. (1971), "Vibration and acoustic test criteria for captive flight of externally carried aircraft stores," *DIGITEK Corporation, AFFDL-TR-71-158*.

Ramsay, J. O. and Silverman, B. W. (1997), *Functional Data Analysis*, New York: Springer.

Rathnayake, L. N. and Choudhary, P. K. (2015) "Tolerance bands for functional data," *Biometrics*, in press.

Storlie, C. B., Fugate, M. L., Higdon, D. M., Huzurbazar, A. V., Francois, E. G., and McHugh, D. C. (2013), "Methods for characterizing and comparing populations of shock wave curves," *Technometrics*, *55*, 436-449.

Tucker, J. D., Wu, W., and Srivastava, A. (2013) "Generative models for functional data using phase and amplitude separation," *Computational Statistics & Data Analysis*, *61*, 50-66.

Tuggle, R. M. (1982), "Assessment of occupational exposure using one-sided tolerance limits," *The American Industrial Hygiene Association Journal*, *43*(5), 338-346.

Wald, A. and Wolfowitz, J. (1946), "Tolerance limits for a normal distribution," *The Annals of Mathematical Statistics*, 208-215.

Appendix – MATLAB code, Parametric Bootstrap

```
function [ub, y_recons, score1, evector1]=bootstrap_par_tolerance_qadj(y, q,
p, n_boot, p_tile, gamma, target_q)
%
%   OUTPUT VARIABLES:
%   -(1) ub--> "qq x 1" vector of upper bounds for spectra
%   -(2) y_recons--> "n x qq" matrix of reconstructed spectra (at target_q)
%   -(3) score1--> n dimensional set of scores for first e-vector
%   -(4) evector1--> loadings of first e-vector: "qq x 1"
%
%   INPUT VARIABLES:
%   -(1) y--> "n x qq" matrix of all spectra
%   -(2) q--> dynamic pressure associated with each spectrum (n dimensional)
%   -(3) p--> number of latent variables (scalar)
%   -(4) n_boot--> number of bootstrap simulations (scalar)
%   -(5) p-tile--> conforming fraction of population (scalar)
%   -(6) gamma--> fractional confidence level (scalar)
%   -(7) target_q--> spectra are normalized to this level
%

[n,qq]=size(y);
mcy=meancent(y);
my=mean(y);
alpha=1-p_tile;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Create basis set from mcy_pfc
%
[~, ~, e_vectors]=svds(mcy, p);
scores = mcy*e_vectors;
std_score=std(scores);
score1=scores(:,1);
evector1=e_vectors(:,1);

des=[ones(1,n); q']';
bhat=inv(des'*des)*des'*scores(:,1);
pred=des*bhat;
pred_error=pred-scores(:,1);
rmse_sav=sqrt(pred_error'*pred_error/(n-2));

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Adjust spectra to target dynamic pressure
% Uses fitted linear regression

%
adj_scores=scores;
for i=1:n
   adj_scores(i,1)=adj_scores(i,1)-bhat(2)*(q(i)-target_q);
end;

mcy_est=adj_scores*e_vectors';
y_recons=ones(n,1)*my+mcy_est;
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Bootstrap Portion of Analysis
mean_spectrum=mean(y_recons);

ind_confid=floor(n_boot*gamma);

[n,dum]=size(scores);
m=zeros(p,1);
s=zeros(p,1);

% Replicate setup
simul_ub_spectrum=zeros(n_boot,qq);
extrap_mult=sqrt([1 target_q]*inv(des'*des)*[1 target_q]');

for k=1:n_boot

    for j=1:p;
        if j==1
            r1=chi2rnd(n-2,1,1);
            r2=normrnd(0,1,1,1);
            s(j)=rmse_sav*sqrt(n-2)./sqrt(r1);
            m(j)=r2.*s(j)*extrap_mult;
        end;
        if j> 1
            r1=chi2rnd(n-1,1,1);
            r2=normrnd(0,1,1,1);
            s(j)=std_score(j)*sqrt(n-1)./sqrt(r1);
            m(j)=r2.*s(j)/sqrt(n);
        end
    end;

    sim_mean_spectrum=mean_spectrum + m'*e_vectors';
    sim_std_spectrum=sqrt((e_vectors.^2)*(s.^2))';
    simul_ub_spectrum(k,:)=sim_mean_spectrum + norminv(1-
alpha,0,1)*sim_std_spectrum;
end;

ss=sort(simul_ub_spectrum);
ub=ss(ind_confid,:);
```