SAND2007-7583C

Validation of Mathematical Models Using Weighted Response Measures

Thomas L. Paez
Jordan E. Massad
Terry Hinnerichs
Chris O'Gorman
Patrick Hunter
Sandia National Laboratories
MS 0557, PO Box 5800
Albuquerque, NM 87185-0557

Abstract

Advancements in our capabilities to accurately model physical systems using high resolution finite element models have led to increasing use of models for prediction of physical system responses. Yet models are typically not used without first demonstrating their accuracy or, at least, adequacy. In high consequence applications where model predictions are used to make decisions or control operations involving human life or critical systems, a movement toward accreditation of mathematical model predictions via validation is taking hold. Model validation is the activity wherein the predictions of mathematical models are demonstrated to be accurate or adequate for use within a particular regime. Though many types of predictions can be made with mathematical models, not all predictions have the same impact on the usefulness of a model. For example, predictions where the response of a system is greatest may be most critical to the adequacy of a model. Therefore, a model that makes accurate predictions in some environments and poor predictions in other environments may be perfectly adequate for certain uses. The current investigation develops a general technique for validating mathematical models where the measures of response are weighted in some logical manner. A combined experimental and numerical example that demonstrates the validation of a system using both weighted and non-weighted response measures is presented.

Nomenclature

$f_X(x)$	Probability density function of random variable X
$F_X(x)$	Cumulative distribution function of random variable X
$P(.)$	Probability of an event
L, U	Lower and upper bounds of a probability interval
S	A validation score
E	Modulus of elasticity of a foam
ρ	Weight density of a foam
x	Measure of response of model
y	Measure of response of experimental system

Introduction

Model validation is the process of testing the adequacy of predictions obtained from a mathematical model relative to the realized experimental behavior of a physical system. The model validation process is described in references [1,2]. Here, we consider the validation of mathematical models of structural dynamic systems. Numerous techniques have been developed for performing validation comparisons on structural dynamic systems. (See, for example reference [3].) We focus on a technique that performs weighted comparisons of discrete response measures of structural dynamic system behavior related to the frequency response functions of linear structures.

We assume that the model under consideration is a stochastic finite element model (FEM) that can be used to approximate the probability distribution of a measure of structural response. The model may accomplish this directly or indirectly, via a Monte Carlo approach. We also assume that there is a relatively small number of, realizations of the physical system that the model is meant to simulate. The experimental responses are treated as deterministic. However, if there are many realizations of the physical system, then statistics of responses can be computed and system behavior can be treated probabilistically. Each physical system realization can be excited, and have its response measured to compute the response measures of interest.

The validation technique developed here is demonstrated for the model of a structural dynamic system, most of whose elements are deterministic, but which includes encapsulating foam that has random characteristics. The foam, its stochastic model and its calibration are described, in detail, in [4,5]. The linear viscoelastic model used for the foam was implemented in the SALINAS finite element code and is described in reference [7]. The purpose of the model is to predict the temperature-dependent mechanical response of the foam to mechanical forces. The purpose of the overall model, including the foam, is to predict spatially dependent displacements, velocities, and acceleration responses.

This paper summarizes the validation algorithm and the results of a validation comparison involving a structure that includes the encapsulating foam mentioned above, and its FEM. The validation is conducted using a measure of acceleration frequency response functions (FRFs). The following section describes an approach to develop a probabilistic description of model-predicted structure behavior, and an approach to performing validation comparisons. Section 2 shows how to develop weighted validation comparisons for an accuracy-based criterion. Section 3 shows how to relax the accuracy-based criterion into an adequacy-based criterion. Section 4 presents an experimental and numerical example demonstrating some validation comparisons. Finally, concluding remarks are made in Section 5.

1 Probabilistic Description of Model-Predicted Response and Validation Comparisons

In order to validate a structural dynamic model it is necessary to select a measure of system response or structural behavior for use in comparison of the model to the experimental system. It is assumed that such a measure of physical system behavior has been chosen, and that the measure is available at *n* values of the system independent parameter of interest. Denote the model-predicted response measures $x_{ij}, i=1,...,m_{mod}, j=1,...,n$. The subscript *i* indexes the realization of the stochastic model response, and the subscript *j* indexes the independent variable. Denote the measures of experimental response $y_{ij}, i=1,...,m_{exp}, j=1,...,n$. The indices represent the same quantities as defined, above.

The number of model predicted response measures, m_{mod}, is assumed to be sufficient to establish a probabilistic description of the model. At a particular index, *j*, of the independent variable, all the measures of experimental response $y_{ij}, i=1,...,m_{exp}$, will be compared to the probabilistic description of the model-predicted response. Therefore, there will be a total of nm_{exp} comparisons. If in a sufficient number of weighted comparisons the model compares favorably with the experimental results, then the model will be judged valid.

Comparisons are performed using a probabilistic representation for the model-predicted measures of response. Numerical predictions of the response measure are used to form the kernel density estimator (KDE) of response measures at each value of index *j*. The KDE [6] is an estimator of the probability density function (PDF) of data at index *j*.

$$\hat{f}_{X_j}(x)=\frac{1}{m_{mod}}\sum_{i=1}^{m_{mod}}\frac{1}{\sqrt{2\pi}\varepsilon}exp\left(-\frac{(x-x_{ij})^2}{2\varepsilon^2}\right) \qquad j=1,...,n,-\infty<x<\infty \tag{1}$$

where $X_j, j = 1,\ldots,n$ is the random variable from which the realizations of the experimental response measure are sampled, and ε is a smoothing parameter of the KDE. The KDE can be integrated to obtain the estimator of the cumulative distribution function (CDF) of the random source of the model-predicted measures of response.

$$\begin{aligned}\hat{F}_{X_j}(x) &= \int_{-\infty}^{x} \hat{f}_{X_j}(\alpha)d\alpha \\ &= \frac{1}{m_{mod}} \sum_{i=1}^{m_{mod}} \Phi\left(\frac{x - x_{ij}}{\varepsilon}\right) \qquad j = 1,\ldots,n, -\infty < x < \infty\end{aligned} \tag{2}$$

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

The *p*-valued probability interval of X_j is the interval within which the probability is *p* that when a single random experiment is performed, the outcome will fall within the interval. The quantity *p* is a number in the interval $[0,1]$. We can use the estimated CDF to approximate the *p*-valued, symmetric probability interval. The *p*-valued probability interval is defined as the interval $\lfloor L_j, U_j \rfloor$ for which

$$P(L_j < X_j \leq U_j) = p \qquad j = 1,\ldots,n \tag{3}$$

The probability interval is symmetric when $P(X_j \leq L_j) = P(X_j \geq U_j) = \alpha / 2$. The limits of the interval are established from

$$L_j = F_{X_j}^{-1}\left(\frac{1-p}{2}\right) \qquad U_j = F_{X_j}^{-1}\left(\frac{1+p}{2}\right) \qquad j = 1,\ldots,n \tag{4}$$

where both expressions can be evaluated numerically using Eq. (2). A validation comparison consists in establishing whether or not a realization of the experimental response measure at index *j* of the independent variable lies within the *p*-valued probability interval $\lfloor L_j, U_j \rfloor$. If it does, the comparison is a success for the model, and if not, the comparison is a failure for the model. If a sufficient number of comparisons result in successes for the model, then the model will be validated; if not, then the model will not be validated.

In order to keep track of the model-to-experiment comparisons, an indicator variable is defined.

$$r_{ij} = \begin{cases} 1 & L_j \leq y_{ij} \leq U_j \qquad i = 1,\ldots,m_{exp}, j = 1,\ldots,n \\ 0 & otherwise \end{cases} \tag{5}$$

When $r_{ij} = 1$, this is an indication of model success. The probability interval, $\lfloor L_j, U_j \rfloor$, is the *p*-valued probability interval, so if the model were a perfect representation of the experimental system, then a fraction of approximately *p* of the comparisons would yield the result $r_{ij} = 1$. In fact, in the perfect-model scenario, each time a comparison is made, it is a Bernoulli trial, i.e., an experiment in which there are two possible outcomes and the probability of a positive outcome is *p*. Recall that there are a total of nm_{exp} comparisons. The probability of $N = n_S$ successes in nm_{exp} Bernoulli experiments is the probability mass function (PMF) of the random variable *N*. It is governed by the binomial distribution, specifically,

$$P(N = n_S) = \binom{nm_{exp}}{n_S} p^{n_S} (1-p)^{nm_{exp} - n_S} \qquad n_S = 0,1,\ldots,nm_{exp} \tag{6}$$

where

$$\binom{j}{k} = \frac{j!}{k!\,(j-k)!} \tag{6a}$$

is the binomial coefficient. Equation (6) quantifies the chance that a perfect model will yield n_S positive outcomes. The CDF of the random variable *N* accumulates the probabilities in the PMF of Eq. (6), and is defined

$$F_N(n_S) = \sum_{i=0}^{n_S} \binom{nm_{exp}}{i} p^i (1-p)^{nm_{exp}-i} \qquad n_S = 0,1,\ldots,nm_{exp} \tag{7}$$

If we seek to assure that the probability of rejecting a perfect model is p_{rej} then we will accept the model as *valid* when the number of successful comparisons, $n_{val} = n_S$, is given by

$$n_{val} = F_N^{-1}(p_{rej}) \qquad 0 < p_{rej} < 1 \tag{8}$$

where $F_N^{-1}(\cdot)$ is the inverse of the CDF in Eq. (7). Because the CDF is defined only on integer realizations, numerical evaluation of Eq. (8) requires rounding to the nearest integer.

2 Weighted Validation Comparisons - Accuracy Criterion

We now establish an approach for performing a validation comparison that weights the individual results of Eq. (5). Assume that a collection of weights, $w_j, j = 1,\ldots,n$, is defined to gage the importance of the validation comparisons. For example, validation comparisons may be weighted heavily where experimental response amplitudes tend to be high. Each of the outcomes, $r_{ij}, i = 1,\ldots,m_{exp}$, defined in Eq. (5) is multiplied by the weight w_j. The total raw score of a set of validation comparisons including nm_{exp} individual comparisons is

$$S_{raw} = \sum_{j=1}^{n} \sum_{i=1}^{m_{exp}} w_j r_{ij} \tag{9}$$

If every comparison outcome were a success, i.e., $r_{ij} = 1, i = 1,\ldots,m_{exp}, j = 1,\ldots,n$, then S_{raw} would be given by

$$S_{raw} = m_{exp} \sum_{j=1}^{n} w_j \tag{10}$$

Therefore, a normalized score for a set of validation comparisons is defined

$$S = \frac{\sum_{j=1}^{n} \sum_{i=1}^{m_{exp}} w_j r_{ij}}{m_{exp} \sum_{j=1}^{n} w_j} \tag{11}$$

The normalized validation score, S, must fall within the interval $[0,1]$. When none of the validation comparisons yields a success, i.e., all the measures of experimental response fall outside their appropriate model-related probability intervals $\lfloor L_j, U_j \rfloor$, then $S = 0$. When many of the validation comparisons yield successes, particularly where the weights $w_j, j = 1,...,n$, are high, then the validation score defined in Eq. (11) is close to one, and this indicates favorable agreement of the model with the experiment.

Given the analysis culminating in Eq. (8), the validation requirement is that the normalized validation score, S, of Eq. (11), must equal or surpass the fraction

$$S_{val} = n_{val} / (mn) \tag{12}$$

Finally, a validation metric can be defined as the difference between the realized validation score of Eq. (11) and the required validation score of Eq. (12).

$$S_{met} = S - S_{val} \tag{13}$$

When the validation metric is nonnegative the model is judged valid with respect to the criteria laid out in this and the previous sections.

The development presented here includes the special case in which all weights are equal, i.e., $w_j = 1, j = 1,...,n$. The approach presented here is what might be called a validation requiring accuracy. This is because a pre-established number of experimental measures of response are required to lie within the probability intervals of the model-predicted measures of response, and no accommodation is included for over- or under-prediction by the model. In the following section an approach that relaxes the requirements to yield a validation requiring only adequacy will be developed.

3 Weighted Validation Comparisons - Adequacy Criterion

Under certain circumstances it may be desirable to perform a validation comparison where it is sufficient to assess the adequacy of model predictions in addition to, or instead of its accuracy. In these cases we may loosen the requirements for establishing positive results in Eq. (5). One way to accomplish this is to require that, in order for a comparison to be successful (or adequate), the j^{th} measure of experimental response fall within the interval $\lfloor \alpha L_j, \beta U_j \rfloor, j = 1,...,n$, where, normally, α is a quantity less than or equal to one, and β is a quantity equal to or greater than one. (The case, $\alpha = \beta = 1$ yields the accuracy validation defined above.) For example, if an over-prediction of the response measure by a factor of two by the model is considered acceptably conservative, and no under-prediction is permitted, then we may set $\alpha = 0.5, \beta = 1$.

4 Experimental/Numerical Example

The structure used in this investigation is a simple assembly of two metal, prismatic masses with equal dimensions, joined by a prismatic element of foam with cross section equal to the cross sections of the two metal elements. An experimental structure is shown in Figure 1.

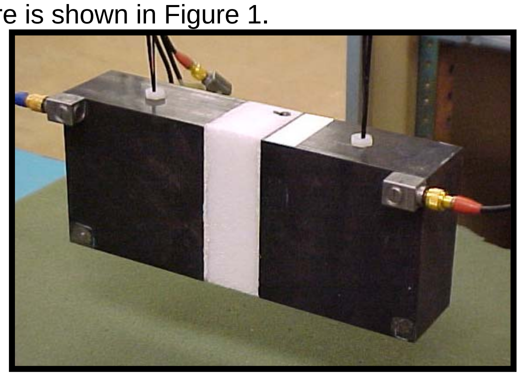

Figure 1. Photograph of the structure whose model is validated in this investigation.

The foam is cellular, with physical properties that vary as random fields and with temperature. Specifically, the foam density, modulus of elasticity, shear modulus, and energy dissipation characteristics all take random values for every finite volume specimen (average characteristic over specimen volume) and for every temperature. A sequence of experiments – known as calibration experiments - on various foam samples and on a collection of structures like the one shown in Figure 1, with both metal elements fabricated from steel, at various temperatures, were used to characterize the foam material properties. It was found that all the properties listed above vary randomly, but the randomness is dominated by the foam density, and the modulus of elasticity is a nearly (but not exactly) deterministic function of the foam density. Figure 2 shows a plot of the foam modulus of elasticity, measured and inferred form the calibration experiments at ambient temperature. The circles denote experiment results, and the line through the origin and the midst of the points is the sample mean of the data obtained via regression. The sample mean function has the formula

$$E(\rho) = C(T)\rho^2 \; ksi \qquad (14)$$

where ρ denotes the material weight density, $C(T)$ is a temperature-dependant coefficient of the square law, and $E(\rho)$ is the density-dependent material modulus of elasticity. For the data plotted in Figure 2, the standard deviation of vertical distance of data from the mean regression line is $\sigma_E = 2.30\,ksi$, the standard deviation of the weight density is $\sigma_\rho = 1.85\,lb/ft^3$, and $C(T) = 0.112$ (compatible units). The deviations of the modulus of elasticity values from the mean line are relatively small compared to the deviations of the material weight density from the mean. Therefore, the weight density is modeled as a random variable, and the modulus of elasticity is modeled as a deterministic function (Eq. (14)) of the weight density.

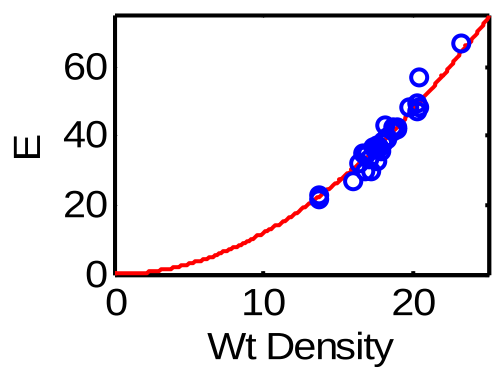

Figure 2. Foam modulus of elasticity in ksi versus weight density in *lb/ft*^{3}.
Data (blue circles), Estimated mean (red line)

Aside from the calibration experiments, a sequence of validation experiments was run on six nominally identical replicates of a system like the one shown in Figure 1. The experiments were conducted at six temperatures, in order to validate the foam model at various temperatures. For the purposes of demonstrating the validation procedures presented in this paper, we summarize a validation comparison performed on one structure, at one temperature. The validation structure is the same as the calibration structure shown in Figure 1. However, validation experiments were done above and below room temperature whereas calibration was only done at room temperature. Therefore, its dynamic characteristics differ from those of the calibration structures. The results of the validation experiments were not made available to the modelers prior to performance of comparisons of mathematical model predictions to the experimental results.

The validation structure was excited with a sequence of forces from an impact hammer, and responses were measured with tri-axial accelerometers located as in Figure 1. Experimental and model-predicted results from one set of impacts – those applied in the transverse direction near one end of the structure – will be summarized here from tests performed at 24 °C. The measured response accelerations were used to estimate the structure frequency response functions (FRFs), and one of the FRFs is shown in Figure 3. This is the FRF of a validation structure where the input excitation is in the transverse direction near one end of the structure, and the response acceleration is in the transverse direction on the mass opposite the excitation.

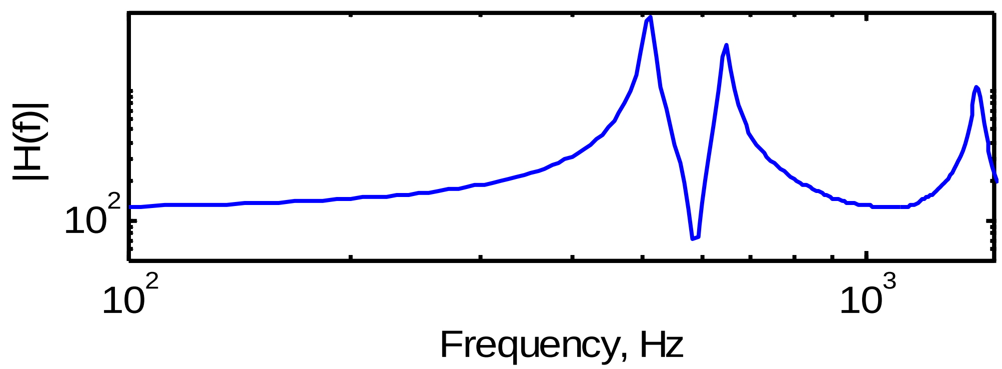

Figure 3. Typical FRF for a validation structure like the one shown in Figure 1.

The FEM of the structure was constructed in the Salinas framework [7]. It uses 3282 eight-node hexagonal elements. the elements are approximately, uniformly sized. The model contains 4120 nodes, and there are three degrees of freedom per node, therefore the model has 12,360 degrees of freedom. The model was solution verified, i.e., a convergence study was performed to show that the FEM converged satisfactorily. The metal components are assumed to have deterministic material properties. For each realization of the structure, the foam has homogeneous, uniform properties. The modulus of elasticity is assumed to be a random variable modeled as in Eq. (14), where the material weight density, ρ , is a random variable. The random variable ρ is assumed to have a normal distribution with mean equal to $\mu_\rho = 19.00\, lb/ft^3$ and standard deviation equal to $\sigma_\rho = 1.28\, lb/ft^3$.

In order to develop stochastic realizations of the model-predicted FRFs corresponding to the experimental FRFs, the values of the foam weight density random variable are sampled using the Latin hypercube approach [8]. For each sample of foam density the material modulus of elasticity is computed, and for each modulus of elasticity value we use the finite element code to evaluate the structural FRF. The FRFs of twenty modeled structures whose foam densities were chosen via Latin hypercube sampling are shown in Figure 4. As for the experiment, the input excitation is in the transverse direction, and the response acceleration is in the transverse direction on the mass opposite the excitation.

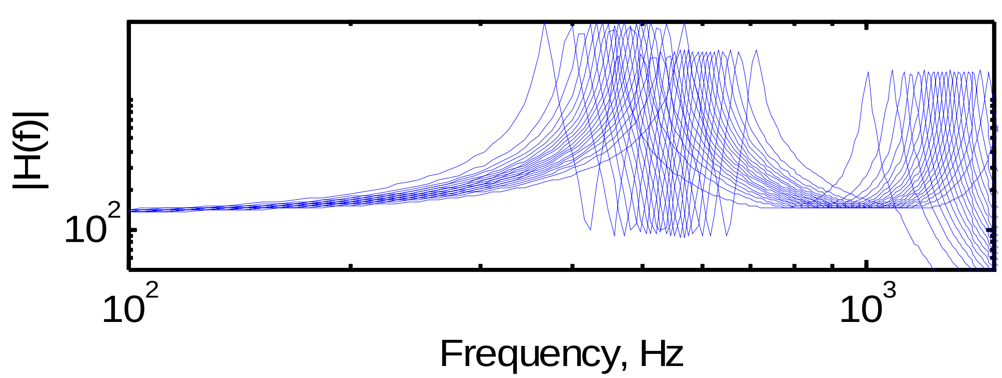

Figure 4. Typical FRFs of system behavior from FEM of structure in Figure 1.

Under some circumstances the model-predicted FRFs of Figure 4 might be compared directly to the validation structure FRF of Figure 3 to establish the validation adequacy of the model. However, to use the framework developed in previous sections we require a discrete measure of the FRF moduli. We define such a discrete measure as follows. For example, let $\left|H_i^{(mod)}(f)\right|, i=1, f \geq 0$, denote the single experimental FRF modulus in Figure 3. A frequency averaged characteristic of the function is

$$y_{ij} = \int_0^\infty w(f - f_j, \gamma) \left| H_i^{(mod)}(f) \right| df \qquad i = 1,...,m_{exp}, j = 1,...,n \tag{15}$$

where $w(f,\gamma), -\infty < f < \infty, \gamma > 0$, is a positive, symmetric, absolutely integrable function whose width depends on the positive parameter γ. The function $w(f,\gamma)$ serves as a weighting function on the modulus of the FRF, and the integral serves to average the weighted FRF over a band of frequencies established by the width of the window. The function $w(f - f_j, \gamma)$ is centered at the frequency f_j, therefore, the weighting occurs for frequencies surrounding the frequency f_j. The operation described here establishes the measures of experimental response, mentioned in Section 1, $y_{ij}, i = 1,...,m_{exp}, j = 1,...,n$. An analogous weighting can be performed on the model-predicted FRFs to obtain a frequency averaged characteristic of that data, $x_{ij}, i = 1,...,m_{mod}, j = 1,...,n$, which represent the model-predicted response measures.

When we use the approach of Eq. (15) to establish discretized response measures of the experimental and the model-predicted FRFs we can perform a validation comparisons as described in Sections 1, 2, and 3. Consider the case where the windowing function is given by

$$w(f,\gamma) = \frac{1}{\sqrt{2\pi\gamma}} e^{-f^2/2\gamma^2} \qquad -2\gamma < f < 2\gamma \tag{16}$$

The window is a truncated Gaussian density function. Application of Eq. (15) to the FRFs of Figures 3 and 4 yields the results shown in Figure 5. The center frequencies are $f_j = 100 \times (1.318)^{(j-1)}, j = 0,...,9$, and each window width equals its corresponding center frequency. The graph makes it clear that the model predictions fit the data quite well, particularly at the higher frequencies.

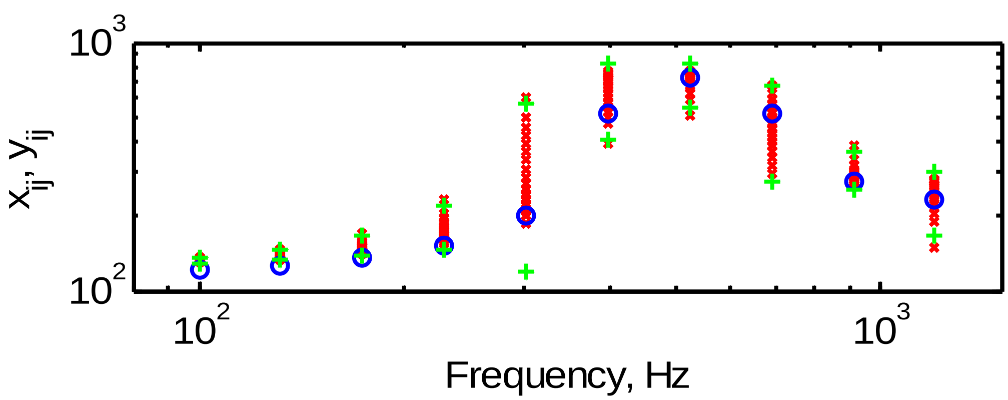

Figure 5. Windowed measures of experimental (blue circles) and model-predicted (red x's) FRFs.

The information provided in Figure 5 can be used to perform a validation comparison as outlined in Sections 1, 2, and 3. This was done for the response characteristics in Figure 5. To accomplish the validation, ninety percent probability intervals were used. The probability intervals are shown by the crosses Figure 5. The probability of rejecting an acceptable model was set to $p_{rej} = 0.20$. This level was chosen for purposes of demonstration of the technique; it is rather stringent. The weighting function for the first validation comparison was chosen to be uniform and equal to one at all frequencies. An accuracy-basis validation was completed; the resulting validation metric is $S_{met} = -0.1$. For the validation criteria used here – accuracy-basis validation, uniform weighting of all comparisons, stringent probability of rejection of 0.2, etc. - the model is rejected as not valid.

The validation analysis was repeated using parameters identical to those listed above, except that now the windowed measures provided at each frequency in Figure 5 were weighted by the squared mean values of the experimental measures. An accuracy basis validation was completed and the resulting validation metric is

$S_{met} = 0.16$. For this validation criterion - accuracy-basis validation, experiment-based weighting of all comparisons, stringent probability of rejection of 0.2, etc. - the model is accepted as valid.

We note that in order to preserve the integrity of the validation process, the decision as to what criteria to use to perform validation analyses should be made prior to the validation comparisons. This comparison makes it clear, though, that when model is satisfactorily capable of making accurate predictions under some important set of circumstances – for example, where amplification of response over input is greatest – then the use of weighted validation comparisons may be valuable.

5 Discussion and Conclusions

This paper develops a technique for performing model prediction to experimental system validation comparisons. It permits the weighting of system characteristics or response measures to emphasize the importance attached to specific predictive capabilities. An example is presented to demonstrate how emphasis might be placed on the predictive capability of a linear model in the high amplitude range of its frequency response function. It is expected that the range of potential applications goes far beyond the example. The weighting can be made with regard to a model's capability to make adequate predictions at various times, at various locations, and for various measures of response. All these applications can be accommodated using the present validation procedure.

Acknowledgment
Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

References

1. ASME, "Guide for Verification and Validation in Computational Solid Mechanics," V&V 10-2006, American Society of Mechanical Engineers, (2006).
2. U.S. Department of Energy, (2000), "Advanced Simulation and Computing (ASCI0 Program Plan," 01-ASCI-Prog-01, Sandia National Laboratories, Albuquerque, New Mexico.
3. Urbina, A., Paez, T., Gregory, D. L., Resor, B., Hinnerichs, T., O'Gorman, C., (2006), "Validation of a Combined Nonlinear Joint and Viscoelastic Encapsulating Foam Model," *Proceedings of the Society for Experimental Mechanics Annual Meeting*, SEM, St. Louis, June.
4. Hinnerichs, T., Urbina, A., Paez, T., O'Gorman, C., Hunter, P., (2007), "Validation of a Viscoelastic Model for Foam Encapsulated Component Response Over a Wide Temperature Range", *Proceedings of IMAC XXV*, Society for Experimental Mechanics, Orlando, FL.
5. T. Hinnerichs, A. Urbina, T. Paez and C. O'Gorman, "Development and validation of a viscoelastic foam model for encapsulated components," *Proc. 2006 SEM Annual Conference and Exposition on Experimental and Applied Mechanics 2006*, Vol. 1, pp. 422-438, 2006.
6. Silverman, B., (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.
7. Reese, Garth, Bhardwaj, Manoj, Walsh, Timothy, Salinas-Theory Manual, Sandia National Laboratories, Albuquerque, NM 87185, Apr 11, 2006.
8. Wyss, G. D., Jorgensen, K. H., (1998), "A User's Guide to LHS: Sandia's Latin Hypercube Sampling Software," SAND98-0210, Sandia National laboratories, Albuquerque, NM.