SAND2012-2632C

Assessment of the Accuracy of a Model for use in Prediction of Component Environments

Angel Urbina[1] and Mike Ross[2]
Sandia National Laboratories, Albuquerque, NM, 87185

Thomas L. Paez[3]
Thomas Paez Consulting, Durango, CO 80301

Modeling and simulation to aid in decision making, such as qualification of high consequence systems, is becoming more prevalent. With this increased use of models, there arises a need to formally assess the accuracy of the model underlying these decisions. This process is referred to as model validation. Furthermore an understanding of the uncertainty in the model responses is necessary if the model is used to predict the response of internal system components for establishing their performance. These responses, presumably, cannot be directly measured and/or multiple pieces of hardware are not available to quantify their inherent variability and thus the value of the model is heightened. Additionally, requirements on component assessment that require quantification of margins, in the presence of uncertainty, provide yet another motivation to use modeling and simulation. In this paper, we describe a model validation process to assess the accuracy of a system level model and perform an assessment of the model relative to existing data.

Nomenclature

$f_X(x)$ = Probability density function of random variable, X
$F_X(x)$ = Cumulative density function of random variable, X
$\hat{f}_X(x)$ = Kernel density estimator of random variable, X
μ_X, σ_X = Mean and standard deviation of random variable X
H(f) = Frequency response function (FRF) as a function of frequency
x_{lim} = Least favorable response or windowed frequency response function as a function of frequency
α = Statistical significance level

I. Introduction

IN recent times there has been an increase in the use of modeling and simulation to support high consequence decision making such as the qualification of complex aerospace systems for certain use environments. Qualification involves the assessment of whether or not a system and its components will operate in its intended manner given a set of inputs. Traditionally, this qualification assessment has been based on the results of experiments performed on a very limited number of hardware (usually one) which tests the components of interest to some predefined threshold of failure. Given today's economic and political constraints, it is anticipated that the availability of physical experiments for use in the qualification of systems and their components will be greatly diminished if not completely gone. Under this scenario is where the value of modeling and simulation is realized and thus necessitates the careful specification of circumstances under which a prediction from a mathematical model can be used in lieu of an experimental result. The use of modeling and simulation to aid in the qualification process is

[1] Member of the Technical Staff, Sandia National Laboratories, Uncertainty Quantification and Model Validation Dept., P.O. Box 5800, MS-0828, aurbina@sandia.gov.
[2] Member of the Technical Staff, Sandia National Laboratories, Analytical Structural Dynamics Dept., P.O. Box 5800, MS-0346, mross@sandia.gov.
[3] Consultant, Thomas Paez Consulting, Durango, CO 80301, tlpaez@bresnan.net

conceptually shown in Figure 1. Notice in the figure the use of a "validated" model. This refers to the use of model predictions when the mathematical model is "proven" to be a satisfactory substitute for the experiment. The process of "proving" that a prediction from a mathematical model can be used in place of an experiment is referred to as model validation.

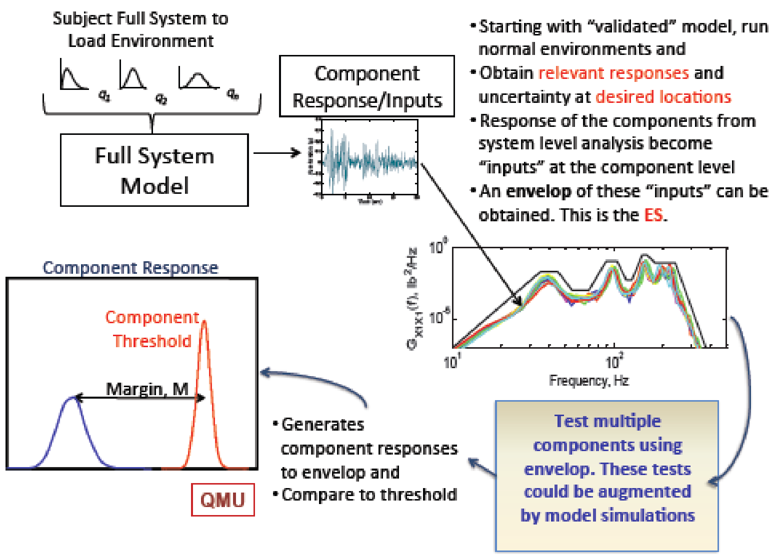

Figure 1. Conceptual use of modeling and simulation to aid in component qualification.

One formal definition of validation is the "process of determining the degree to which a computer model is an accurate representation of the real world from the perspective of the intended model applications."[1,2] The definition implies that validation is accomplished through comparison of the responses predicted with a mathematical model to the responses realized by a physical system during an experiment. Therefore, such comparisons are an integral part of the validation process. But, of course, some knowledge of the predictive accuracy of the mathematical model away from the points in the parameter/environment space where the comparisons are made is usually sought. This knowledge relies on an understanding of the interpolative and extrapolative fidelity of the mathematical model. Much remains to be learned about how these types of inferences can be made and are not treated in this paper.

An important physical reality that influences the comparisons to be performed in a validation analysis is that real systems are stochastic. In a structural dynamic system, this is so for many reasons. Among others, nominally identical structures differ randomly from one to the next because of dimension tolerance uncertainties, material property differences, and differences in fabrication details. Further, different assemblies of one collection of structural parts behave differently because of variations in details of connections, continuous changes in structural component behavior during tests, and variations in test conditions. These facts indicate that a validation process must accommodate the comparison of deterministic-to-random and random-to-random quantities. The experimental/numerical example presented in this paper exercises this requirement. To simulate the stochastic variations of structural systems, the mathematical models of these systems have stochastic forms. An application problem which revolves around the pictorial representations shown in Figure 2 will be used to demonstrate the validation process described in this paper.

A. Physical System

The schematic shown in Figure 2 refers to a physical system which is made up of components and it is subjected to external loads which in turn excite some responses. The details of the physical system shown in the figure are:

- External loads excite the entire structural system.
- Entire structure and all the components it contains respond.
- System responses near input points to system components are component environments.
- Component environments excite system components.
- Components respond to inputs.

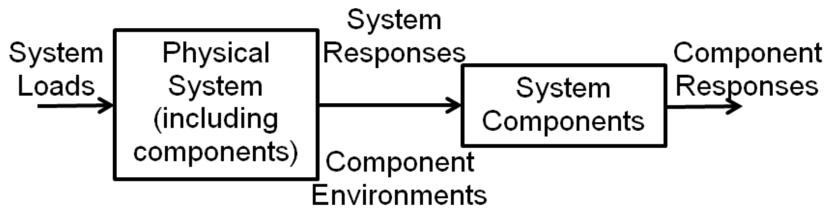

Figure 2. External load excites physical system. System responds, and motions near components are input environments for components. Components respond to environments.

B. System Model

A system model simulates the behavior of a system in pre-established use environments. When inputs are specified, system responses can be predicted. The model can be:

- Deterministic: Represents, in some sense, the average behavior or extreme behavior, of the physical system. System identification (calibration) techniques should be used to assure accuracy of model.
- Stochastic: Represents the spectrum of behaviors that physical system might realize.
  - Normally, collection of components modeled as random.
  - Probability laws of behaviors of components must be identified (calibrated).
  - Realizations of component parameters might be generated via Monte Carlo analysis.
- External loads obtained from analysis of physical experiments excite the model of the entire structural system. (Components in this model constructed to have the resolution required to obtain valid results.)
- Response computed for entire structure and all the components it contains.
- Computed system responses near input points to system components are component environments.
- Computed component environments excite system components in laboratory experiments.
- Physical components respond to experimental inputs.

II. Validation Procedure

Model validation is a comparison of some measure(s) of behavior of a mathematical model to the corresponding measure(s) of behavior estimated from an experimental system. A widely accepted tenet of model validation is that any comparison between a mathematical model **calibrated** to a physical experiment and a measure of data from the same physical experiment cannot form the basis for a meaningful validation. The mathematical model used in the comparison must be predictive – built from fundamental principles – or at least calibrated to experimental results different from those used in the validation analysis. The comparison would be an easy one if physical systems were deterministic and our measurements of their behaviors were noise-free. We would simply compare the outputs of a deterministic mathematical model to the deterministic measure of physical system behavior. If they agreed "satisfactorily" then the model would be judged valid. However, measurements of physical system behavior are always noisy. The boundary conditions and excitations of physical systems are practically always random. Unmeasured factors in the environment affect physical system response. And distinct physical systems drawn from ensembles of nominally identical structures differ randomly - sometimes substantially[3,4]. Modern mathematical modeling techniques acknowledge this latter source of structural randomness to form stochastic models[5,6]. These factors complicate the comparison of the behaviors of mathematical models to the behaviors of the structures they are meant to represent by requiring the use of probability and statistics to perform the comparison. Figure 3 shows the conceptual idea behind model validation in the presence of uncertainty. Note that the model needs to agree not only in its mean behavior but ideally it should represent the level of uncertainty present in the actual physical system is trying to represent (this is represented by the σ's shown in the figure). Although the figure shows a response that follows a Gaussian distribution, this is not a requirement for the methodology described herein to work.

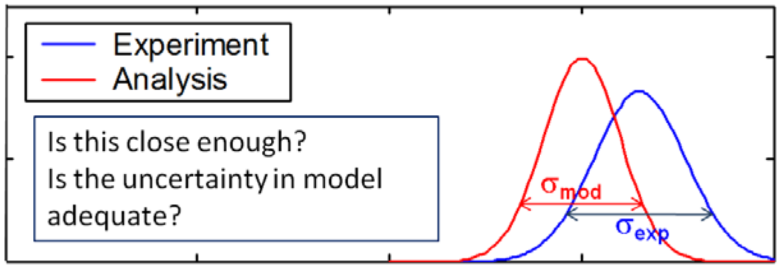

Figure 3. The model validation conceptual picture and basic questions.

From a procedural point of view and before a well-structured validation comparison can be performed, there are several decisions that must be made and criteria to be defined. A listing of items to be considered and decisions to be made before commencement of calibration and validation experiments, and model predictions is given below[7]:

- *Specify the model use/purpose, and the response measures of interest*. Model use/purpose should indicate whether the model is intended for use in preliminary/intermediate/advanced analysis and prediction. Measures of interest are the quantities the mathematical model was constructed to predict.
- *Specify validation metrics and domain of comparison*. The validation metric defines the way in which the model/experiment comparison will be made, for example, differences of response measures. Domain of comparison indicates ranges of excitation, boundary conditions, etc., within which validation comparisons are to be carried out.
- *Specify calibration experiments*. Physical experiments to be used to calibrate the mathematical model parameters (i.e., identify the model parameters) must be defined.
- *Specify validation experiments*. These identify physical system environments, boundary conditions, etc., whose responses must be satisfactorily predicted by the mathematical model in order for the model to be deemed adequate. They may involve interpolation among and/or extrapolation outside the points in environment/parameter space where calibration was performed.
- *Specify adequacy criteria*. These define the degree of accuracy the model-predicted measures of interest relative to experimentally inferred measures of interest required of the mathematical model. In an ideal scenario, these criteria will be specified by the customer and will be relevant to the intended application. In practice, specification of this is not a straightforward proposition and most of the times it is an iterative process.
- *Perform computational model predictions of validation experiment responses*. To maintain the integrity of the validation process, results are normally withheld from experimentalists.
- *Perform validation comparisons***.** Compute measures of interest of experimental system and computational model responses, then validation metrics are computed. Judge the adequacy of the computational model.

In the following section, we describe a probabilistic/statistical framework for the comparison of stochastic mathematical models and the systems they are meant to represent. The construct demonstrated here permits the characterization of the physical systems using the tools of probability and statistics. This part of the framework can be used when experimental data are available. In addition, the framework permits the incorporation of results from stochastic system analysis into the comparison. We assume that the physical system under consideration and its mathematical model will be compared using one of the metrics defined in Table 1. The advantages and disadvantages related to each of the metrics are also listed on the table.

Table 1. Metrics for Comparison of Transient Responses of Structural Dynamic Systems

Metric	**Advantages**	**Disadvantages**
Time history	❑ Directly measured ❑ Easy to interpret	❑ Difficult to match, directly
Fourier transform spectrum	❑ Easy to compute ❑ Does not rely on phase ❑ Contains "important" part of information from time history	❑ May be difficult to match, unless smoothed ❑ Non-unique
Temporal moments	❑ Easy to compute ❑ Characterizes signal shape ❑ Potential for match is reasonable	❑ Non-unique ❑ Use of few moments may yield apparent match that is unsatisfactory for some purposes
Shock response spectrum	❑ De facto standard in most shock labs	❑ Insensitive to shock details ❑ May yield misleading conclusions ❑ Difficult to understand ❑ Non-unique
Choi-Williams expansion	❑ Estimate of time-varying spectral density ❑ Mathematically well-founded	❑ Theoretical details may be difficult ❑ Non-unique ❑ Difficult to obtain with few measured realizations
Wavelet decomposition	❑ Easy to compute ❑ Compresses signal information into a few important components	❑ Phase mismatches may cause apparent differences where match is good ❑ Does not yield moments directly (but problem may be circumvented)
Time-varying spectral density	❑ Fundamental measure of non-stationary random process ❑ Estimates other than Choi-Williams available	❑ Non-unique ❑ Difficult to obtain with few measured realizations
Principal components/KL expansion	❑ Easy to compute ❑ Compresses signal information into a few important components	❑ Requires metric to compare principal shapes

The measure to be used in this paper is a measure of the frequency response function (FRF) of a structure. To define the FRF let $q(t), t \geq 0$ denote the single force excitation applied to a structure at a point, and let $\ddot{x}(t), t \geq 0$ denote the response that the force excites at a point. (The FRF can also be defined for motion inputs.) The Fourier transforms of the excitation and response are defined as:

$$Q(f) = \int_0^\infty q(t) e^{-i2\pi f t} dt \qquad -\infty < f < \infty \tag{1}$$
$$X_2(f) = \int_0^\infty \ddot{x}(t) e^{-i2\pi f t} dt \qquad -\infty < f < \infty$$

where the subscript "2" in the second expression indicates that this is the Fourier transform of a second derivative of a function. The FRF relating response to excitation is defined

$$H(f) = \frac{X_2(f)}{Q(f)} \qquad -\infty < f < \infty \tag{2}$$

for all $X_2(f) \neq 0$.

For the purposes of a validation comparison it is sometimes not convenient to compare one function to another. The reason is that in some situations functions may compare quite favorably over most of their domain of comparison, but not as well over a small interval. In such a situation it may be difficult to assess the adequacy of the comparison. Therefore, we define a measure of the FRF known as the least favorable response (LFR). It is

$$x_{Lim} = \int_0^\infty W(f)|H(f)|df \tag{3}$$

Where $W(f), f \geq 0$ is a nonnegative, absolutely integrable function; it is a windowing function. The LFR, x_{Lim}, is a bound on the response of a linear system with FRF *H(f)* to excitations whose Fourier transform moduli are bounded by *W(f)*. The LFR will be used later in the validation examples where *W(f)* is a symmetric function, and will be associated with the center frequency of *W(f)*. Conceptually, this process is shown in Figure 4. The LFR is the integration shown in Equation 3.

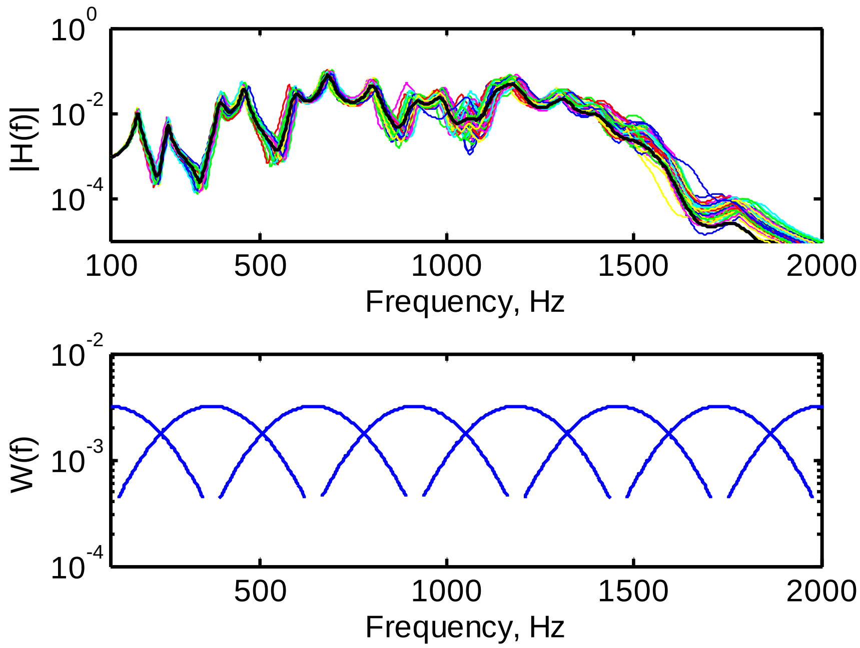

Figure 4. Top Figure: Some FRFs of a system response; Bottom Figure: An example of windowing functions centered at some selected frequencies

One or more FRFs of a structure can be computed for experimentally measured and model-predicted responses. When the FRFs of some model-predicted responses are satisfactorily near the FRFs of experimentally measured responses, then the mathematical model is a satisfactory representation of the physical system with respect to force in/acceleration out FRF measures. The latter statement is what the model validation process seeks to address. The next section shows a specific approach to perform this assessment in a probabilistic framework.

III. Validation assessment based on Test of Hypothesis

During the performance of every model validation analysis, there is a need to establish a framework and procedure for comparing measures of system response obtained from the model to corresponding measures of system response obtained from validation experiments on the structure that is the object of the model. There are less-formal approaches for performing the comparison including, for example, simply observing whether or not the model-predicted measures of response are within some arbitrary, customer-selected, distance of the experimental measures of response. However, more formal approaches may be desired. In the case where the model and/or experiment are stochastic and the measure of response is discrete, methods of fundamental statistics might form the basis for validation comparisons. This section summarizes a formal approach to validation comparisons with roots in statistical test of hypothesis[8] (TOH).

The idea underlying TOH is to establish whether or not a particular statistic of computed data might plausibly have arisen from a particular random source. In a validation analysis we would establish a hypothesis: A measure of structural dynamic response obtained from a model is a realization of the random source of the corresponding measure of structural dynamic response obtained from experiment. We would then proceed to establish whether or not the hypothesis can be rejected. If not, and if the comparison has been fairly specified, then we would pronounce the model validated, at least, for the measure of response considered, at the location where the comparison was performed, in the frequency range considered, etc. (Of course, we acknowledge that the model-predicted measure of response cannot actually come from the experimental random source – the model produces, at best, accurate simulations of the experimental source. Still, the comparison can work, in practice.)

In order to develop the method mathematically, denote a random measure of experimental structural dynamic response $M^{(e)}$; the quantity is a random variable. Assume that we know (or can obtain, approximately) the probability distribution of $M^{(e)}$; its probability density function (PDF) is $f_{M^{(e)}}(\beta), -\infty < \beta < \infty$. For now, assume that the model is deterministic. Denote the model-generated prediction of the measure of response of the experiment $m^{(a)}$. (The superscript "a" indicates that the value comes from analysis.) Our hypothesis is: The model-predicted measure of response, $m^{(a)}$, comes from the random source $M^{(e)}$. In order to test the hypothesis at the α-level of significance, we establish the $(1-\alpha) \times 100\%$ probability interval surrounding the mean of $M^{(e)}$, then observe whether or not the quantity $m^{(a)}$ lies within the interval. If it does, then we cannot reject the hypothesis; otherwise, we can. We establish the probability interval $\left[E\left[M^{(e)}\right] - \Delta_{(1-\alpha/2)}, E\left[M^{(e)}\right] + \Delta_{(1-\alpha/2)}\right]$ by solving for $\Delta_{(1-\alpha/2)}$ in the relation

$$\begin{aligned} &P\left(E\left[M^{(e)}\right] - \Delta_{(1-\alpha/2)} < m^{(a)} \leq E\left[M^{(e)}\right] + \Delta_{(1-\alpha/2)}\right) \\ &= F_{M^{(e)}}\left(E\left[M^{(e)}\right] + \Delta_{(1-\alpha/2)}\right) - F_{M^{(e)}}\left(E\left[M^{(e)}\right] - \Delta_{(1-\alpha/2)}\right) = 1 - \alpha \end{aligned} \tag{4}$$

where $F_{M^{(e)}}(\beta), -\infty < \beta < \infty$ is the cumulative distribution function (CDF) of the random variable, $M^{(e)}$. Depending on the form of the CDF, the interval may be established via closed-form calculation, or numerically. This is the simple basis for TOH-based validation analysis. When a model "passes" the validation comparison, it is judged valid for the measure of response considered, at the time or frequency considered, at the location considered, etc.

It is typical, however, that multiple measures of structural dynamic response are generated by the model, for multiple locations, multiple times or frequencies, and/or multiple physical measures of response. In such a situation we are confronted with the need to simultaneously validate the model with respect to a collection of measures, $M_k^{(e)}, k = 1, \ldots, n_e$, where n_e denotes the number of measures. In this case we simply determine a collection of n_e intervals as in (1) for each $M_k^{(e)}$. Clearly, with increasing n_e the chances of passing all the validation comparisons diminishes, no matter how accurate the model.

To accommodate that tendency, we note that level of significance, α, in the TOH corresponds to the chance that an accurate model will satisfy the requirement that $m_k^{(a)}$ will lie within the interval $\left[E\left[M_k^{(e)}\right] - \Delta_{(1-\alpha/2),k}, E\left[M_k^{(e)}\right] + \Delta_{(1-\alpha/2),k}\right]$. If the model is an accurate representation of the experiment, then in approximately $(1-\alpha) \times 100\%$ of the comparisons performed, $m_k^{(a)}$ will lie within the interval. Whether or not $m_k^{(a)}$ falls within the interval is a binary event, therefore, we note that successful satisfaction of the requirement, "$m_k^{(a)}$ lies within the interval $\left[E\left[M_k^{(e)}\right] - \Delta_{(1-\alpha/2),k}, E\left[M_k^{(e)}\right] + \Delta_{(1-\alpha/2,k)}\right]$," is the outcome of a Bernoulli trial. Denote by N_s the random variable that defines the number of times in n_e comparisons that $m_k^{(a)}$ lies within the interval $\left[E\left[M_k^{(e)}\right] - \Delta_{(1-\alpha/2),k}, E\left[M_k^{(e)}\right] + \Delta_{(1-\alpha/2,k)}\right]$. The random variable N_s has a binomial probability distribution with probability of success $(1-\alpha)$, and range of realizations $(0,1,\ldots,n_e)$.

During a validation analysis we would, normally, not simply wish to accept a model that is "good on average," but rather, we might wish to accept a model that cannot be rejected in a second-level TOH. To develop such a TOH, we first establish the second-level hypothesis: Each comparison performed in the first TOH is a Bernoulli trial with the probability of success $(1-\alpha)$; the collection of n_e comparisons is mutually, statistically independent. To test the second-level hypothesis at the γ level of significance, we form the CDF of an n_e-trial, binomial random variable with probability of success, $(1-\alpha)$. That CDF is

$$P(N_s \leq n) = F_{N_s}(n) = \sum_{i=1}^{n} \binom{n_e}{i} (1-\alpha)^i \alpha^{(n_e - i)} \qquad 0 \leq n \leq n_e \tag{5}$$

where $\binom{n_e}{i}$ is the binomial coefficient. We find the $(1-\gamma) \times 100\%$ percentage point, n_γ, of the binomial distribution through numerical inversion of (2). (The percentage point will, normally, not be an integer, so we round the result to the nearest integer.) If the number of positive comparisons (i.e., comparisons in which $m_k^{(a)}$ lies within the interval $\left[E\left[M_k^{(e)}\right] - \Delta_{(1-\alpha/2),k}, E\left[M_k^{(e)}\right] + \Delta_{(1-\alpha/2,k)}\right]$) is smaller than n_γ, then the hypothesis is rejected;

otherwise, the hypothesis is not rejected. In the latter case, the model "passes" the validation comparison and it is judged valid for the measures of response considered, at the times or frequencies considered, at the locations considered, etc. We can reverse the framework when the model is stochastic and there is only one experiment or a few experiments.

IV. Demonstrative Example

A. Overview

To demonstrate the validation methodology described in the previous section, we use a structural dynamics example for a system that schematically shown in Figure 5. The system is excited on the right end with a low level impact from a soft-tipped, instrumented hammer and acceleration responses are measured at 3 locations: 1) at mid-span on an interior point (location not shown in the Figure), 2) at mid-section on an exterior location and on the left most end. The measurements are made using triaxial accelerometers. To perform an analysis, a 3D model using solid elements was constructed and analyzed using Sandia National Laboratories' (Sandia) developed structural dynamics code, Salinas[9]. This model has on the order of 1 million degrees of freedom and to perform linear transient analysis, the run times are in excess of 12 hours each. In order to enable uncertainty quantification, multiple runs are necessary and thus a more efficient model needed to be created.

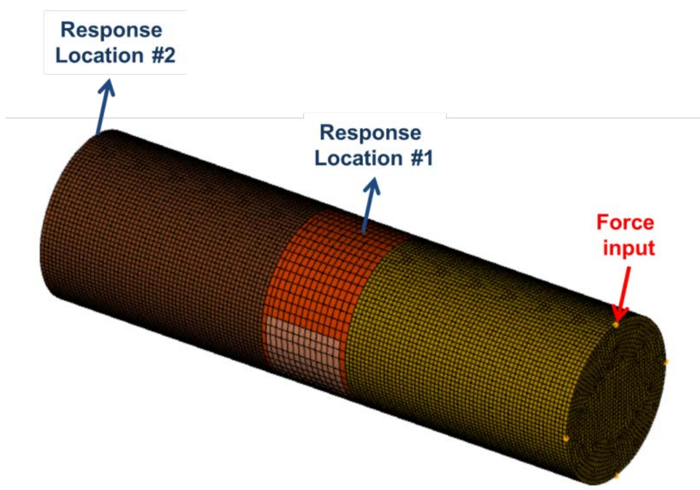

Figure 5. Schematic of system used to demonstrate validation methodology.

As described in the paper by Ross at el[10], the full 3D model shown in Figure 5 was reduced using a Craig-Bampton reduction approach and a schematic of this reduction is shown in Figure 6. This reduced order model will be referred to as the high efficient model (HEM). Details on how the reduction was done are in Reference 10. For comparison, the computational time to run this model was on the order of 6-8 hours on a computer cluster.

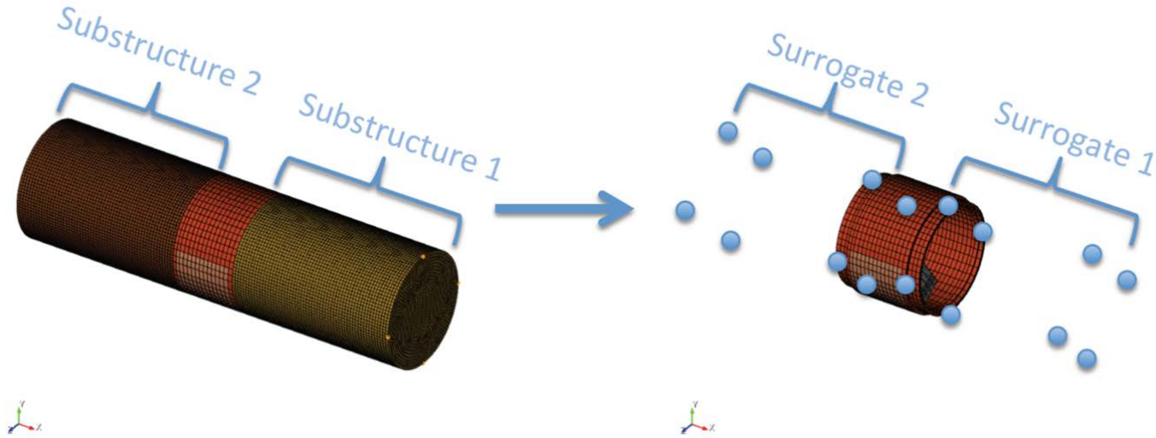

Figure 6. Craig-Bampton reduction of full model.

For completeness, an acceleration time history comparison of the full 3D model to the reduced order model is shown in Figure 7. This figure shows that the HEM model is a good surrogate of the 3D model, at least from a visual comparison. A more quantitative comparison between the full 3D model and the HEM model is given in Reference 10

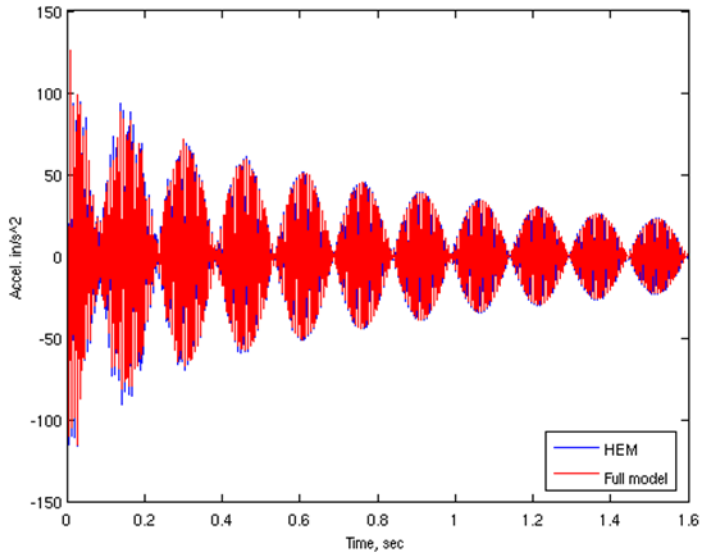

Figure 7. Acceleration time history comparison of full model to HEM model.

From the basic understanding of the relevant physical phenomena controlling the response of interest coupled with a sensitivity analysis based on modal responses, a set of model parameters was selected and a probability model for each one was chosen. For this example, only bounding information on the selected parameters was available; therefore, a uniform distribution was selected as the probability model for all parameters. The parameters and their bounds are shown in the table below. In all, 22 random variables were considered in this example.

Table 2. Model parameters of interest and their bounds

Parameter	Lower bound	Upper bound
Damping – mass proportional stiffness	$5.0e^{-7}$	$5.0e^{-6}$
Damping – stiffness proportional stiffness	$1.0e^{-4}$	$1.0e^{-3}$
Silicon material – modulus of elasticity	100 psi	1300 psi
Silicon material – Poisson's ratio	0.45	0.4999
Translational stiffness at interfaces (3 directions @ 3 locations)	$7.0e^{6}$ lb/in	$1.3e^{7}$ lb/in
Rotational stiffness at interfaces (3 directions @ 3 locations)	$7.0e^{6}$ lb/in	$1.3e^{7}$ lb/iin

B. Analysis and Results

To propagate parametric uncertainty through the model, 80 Latin hypercube sample sets were generated using Sandia's DAKOTA[11] software. Each sample set consisted of 22 realization of each of the parameters shown in Table 2. For each sampled parameter set, a model output consisting of accelerations at the 3 response locations of interest were obtained. From the acceleration responses, the FRFs and the LFRs were calculated using Equations 1 through 3. These results as well as the measured response from one physical system are shown in Figure 8 through Figure 10. By observing these figures, it is noted that the system model in general captures the main characteristics of the actual physical system. The beating phenomena, the principal modes of vibration and the decay in the acceleration time history, albeit at a different rate are present in the model results as compared to the experimental data. The effect of parametric uncertainty is also observed, especially in the LFRs (shown in Figure 8 through Figure 10 part c), where the cluster of dots at each frequency of interest span a large range. It is worthwhile to point out that the location and width of the windows used to calculate the LFR are determined based on the specifics of the problem. For this example, the locations (center frequencies) correspond to the modes of interest of the system. The window widths were either narrow or wide depending on how many features (i.e. modes) we felt were appropriate to include. This is a judgment call and it is problem dependent. The windows selected for this example are for illustration of the method only.

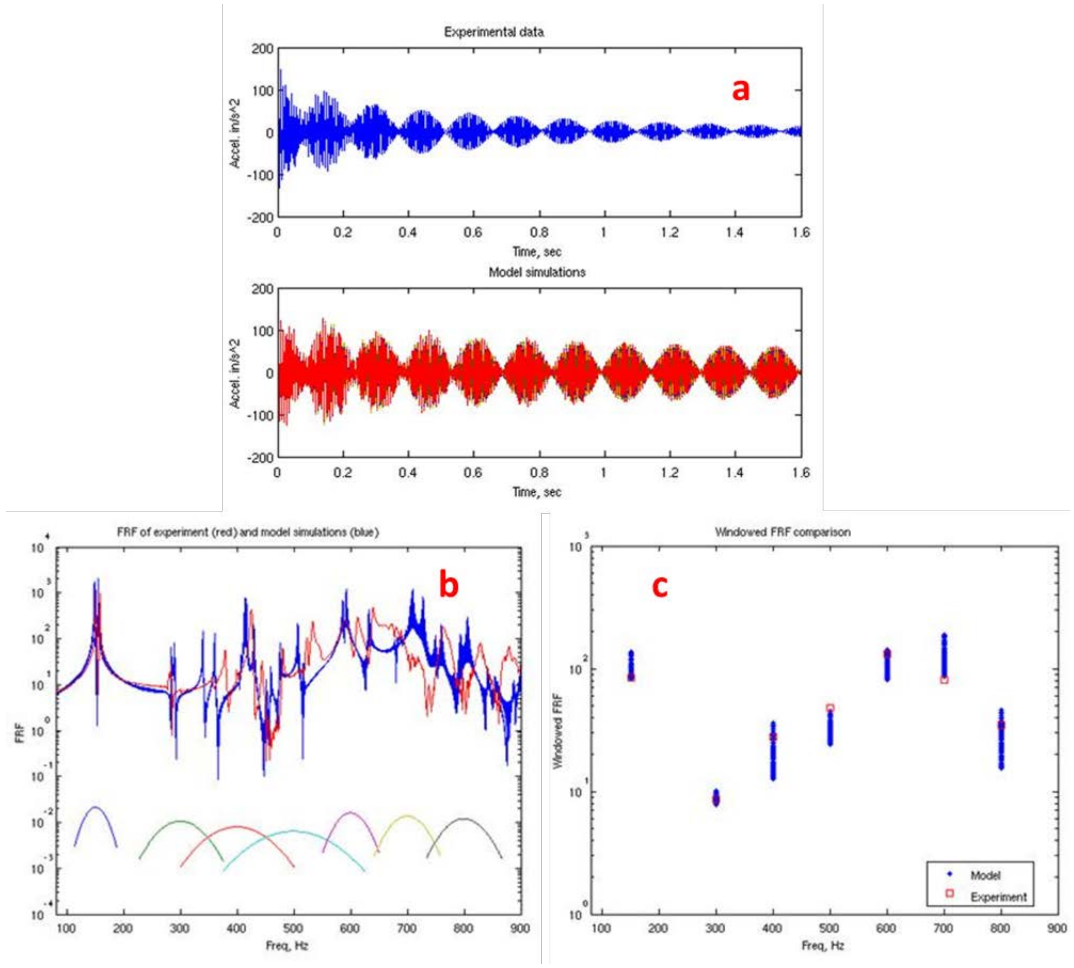

Figure 8. Results of middle section - inner point
a) Acceleration time history, b) Frequency response function (FRF) and windows, c) Least Favorable Response (LFR)

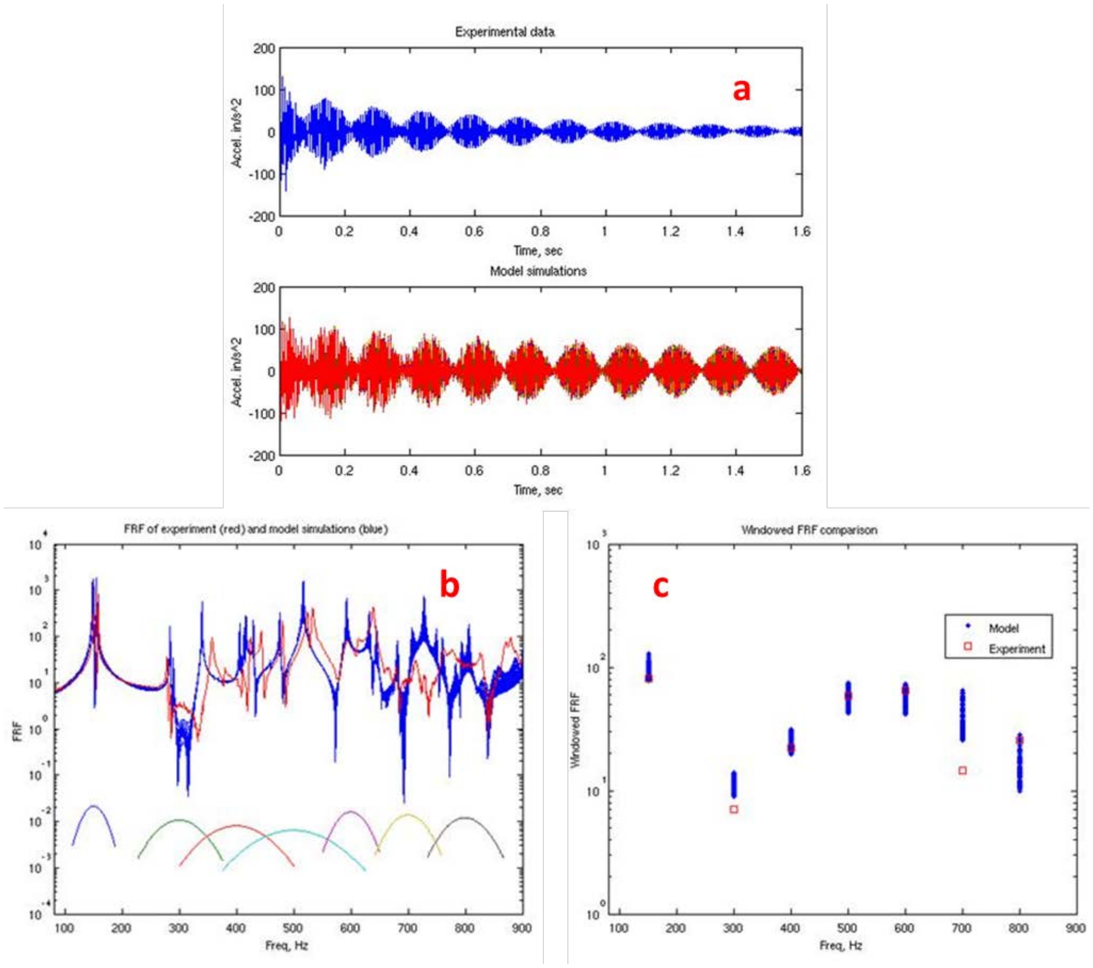

Figure 9. Results of middle section – outside
a) Acceleration time history, b) Frequency response function (FRF) and windows, c) Least Favorable Response (LFR)

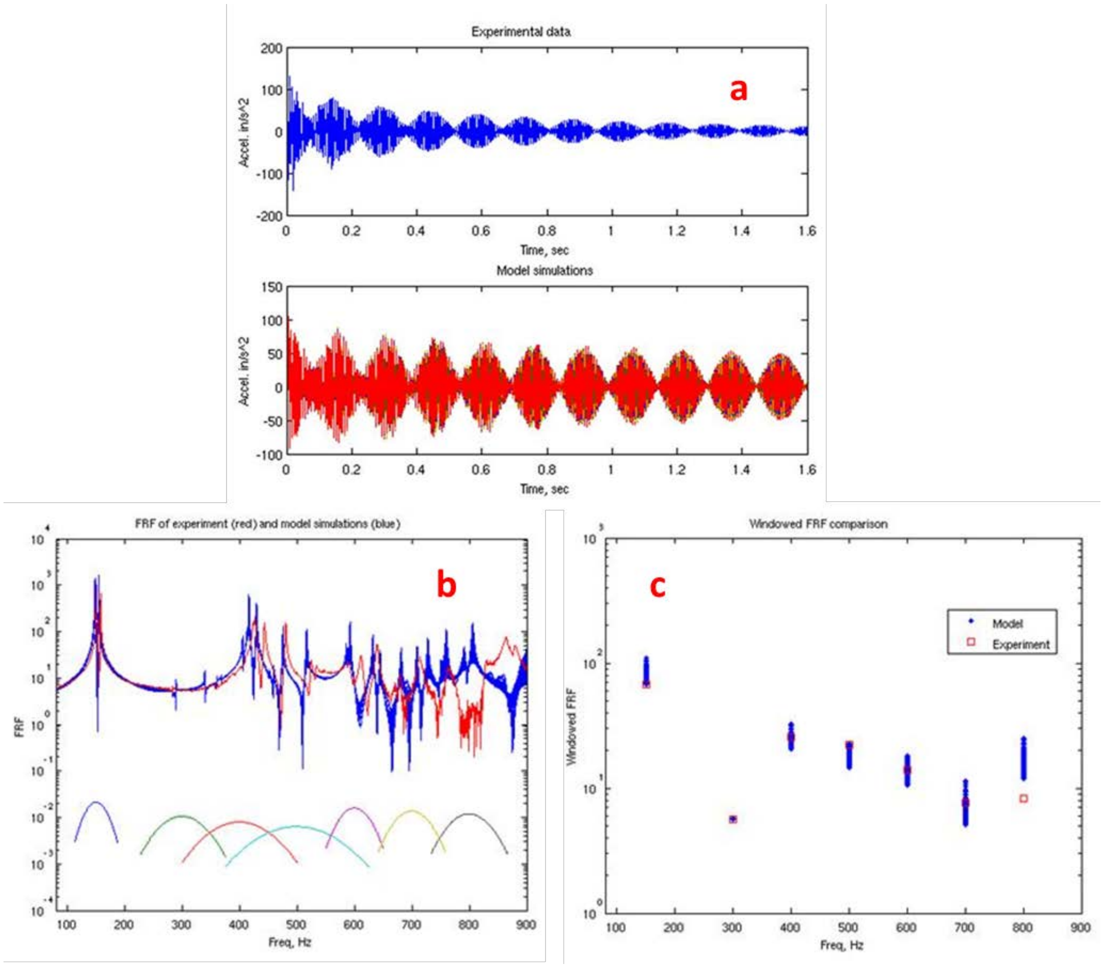

Figure 10. Results of left end
a) Acceleration time history, b) Frequency response function (FRF) and windows, c) Least Favorable Response (LFR)

One final set of results that can be obtained from the analysis is the relative contribution of different sources of uncertainty to the response quantity of interest. In this case, we examine two sources, 1) due to parametric uncertainty and 2) due to use of the HEM model. Their contribution is shown relative to the measure of interest, in this case, the LFR and these are plotted in Figure 11. The 3 plots represent the 3 locations of interest. Shown in these figures is evidence that for this particular measure of response, the parametric uncertainty encompasses the error due to using a reduced order model. It is interesting to note that on one of the response locations (shown in the upper left hand plot in Figure 11), particularly in window location #5, the reduced order model has the larger error when compared to all other window locations within the same response locations and among the other locations. This brings up the fact that the model might not have the same uncertainty at all locations and thus this fact will have to be included in the validation assessment, which will be examined in the following section.

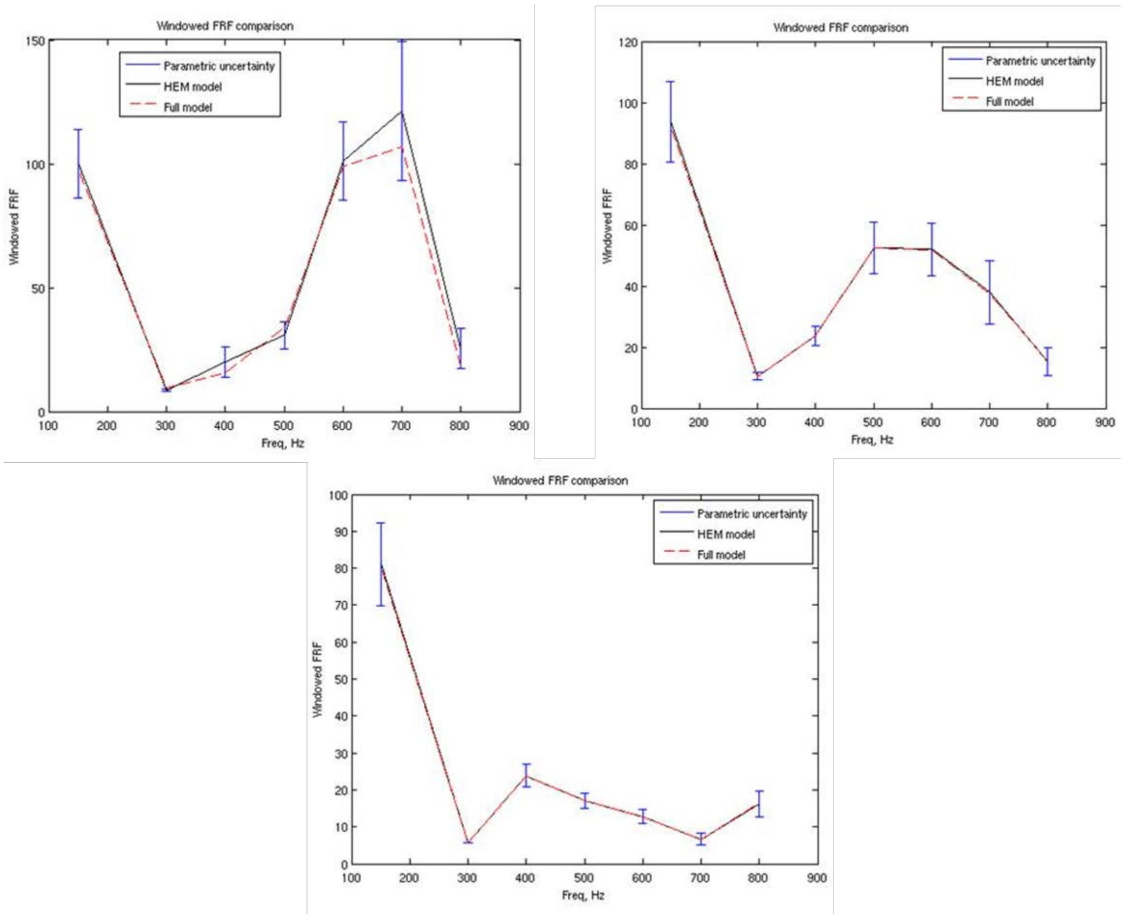

Figure 11: Sources of uncertainty and their relative contribution

C. Model validation assessment

To demonstrate the validation methodology shown in Section III, we use the result obtained in the previous section. We are provided with a collection of FRFs such as those shown in Figure 8 through Figure 10 part b. Shown there are FRFs predicted with a collection of model responses (blue lines), and one FRF modulus obtained from test data (red line). The objective in this validation assessment is to use statistical test of hypothesis (TOH) to establish whether the experimental LFR can be rejected as arising from the random source of the model LFR. From Figure 12, the blue dots show the results for the model-generated LFRs and the red square shows the result for the experimental LFR as calculated using Equation 3. To illustrate the method, the first window is shown at left, bottom plot. The blue dots have the CDF approximator shown at right, bottom. The LFR of the experimental FRF occurs at the location identified with the red line. The hypothesis is that the experimental datum arises from the random source that produced the model data. To test the hypothesis at the α level of significance $\alpha = 0.05\%$, we do the following:

- Compute the $\alpha/2$ and $1 - \alpha/2$ percentage points from the CDF. The percentage points are shown by the black lines
- When the experimental result lies within the interval indicated by the black lines, the hypothesis cannot be rejected
- Otherwise the hypothesis is rejected
- When a sufficient number of tests yield "non-rejection" results the model is validated

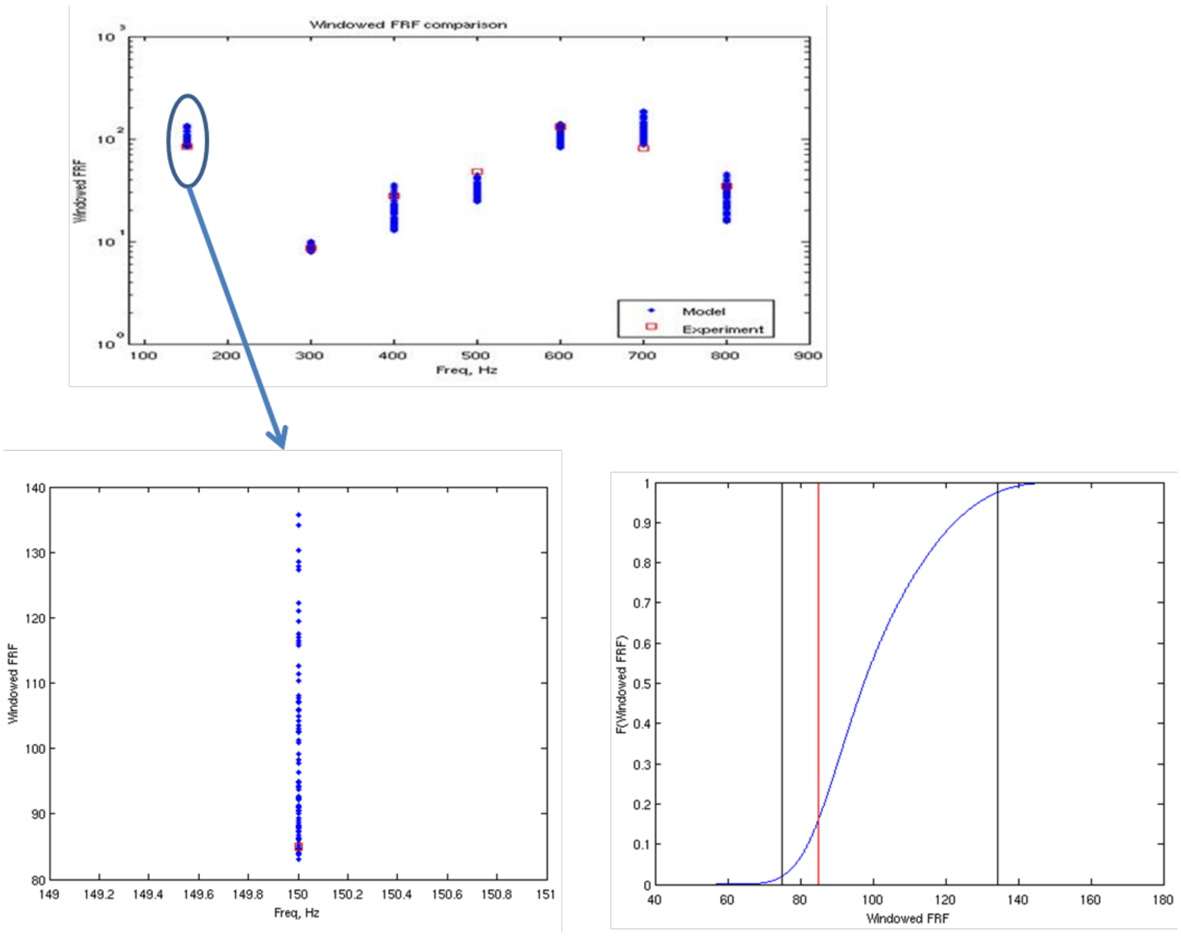

Figure 12. Validation comparison

From Figure 13, tests of hypothesis were performed for seven sets of data with $\alpha = 0.05$ (Top figure). These seven sets correspond to the 7 center frequencies used to calculate the LFR. At one of these frequency center locations, the model was rejected (REJ); in six it was not (DNR). If the experiment data arise from the random source of the model data, probability of "Do Not Reject" is:

$$p = 1 - \alpha = 0.95 \quad (7)$$

Further, if we treat the TOH as seven Bernoulli trials with two possible outcomes, the probability of j "success" is Binomial. The binomial CDF is shown in the bottom plot of Figure 13. A key question that arises is: How many outcomes must be DNR to validate the model? To determine this, we choose a probability, α_2, acceptable for rejecting the correct model. This represents the significance level of this second test. In this case and for illustration purposes, $\alpha_2 = 0.05$ which then says, j=5 positive results would be necessary to validate the model. There are six DNRs, so the model is deemed validated using this criterion. In an actual application, the level of significance will be determined by the end use of the model and the context in which this model is being use (i.e. to support design trade off studies will require a lesser level of rigor in the specification of the significance level relative to the case in which the model is directly used to support certification of a component in a large system).

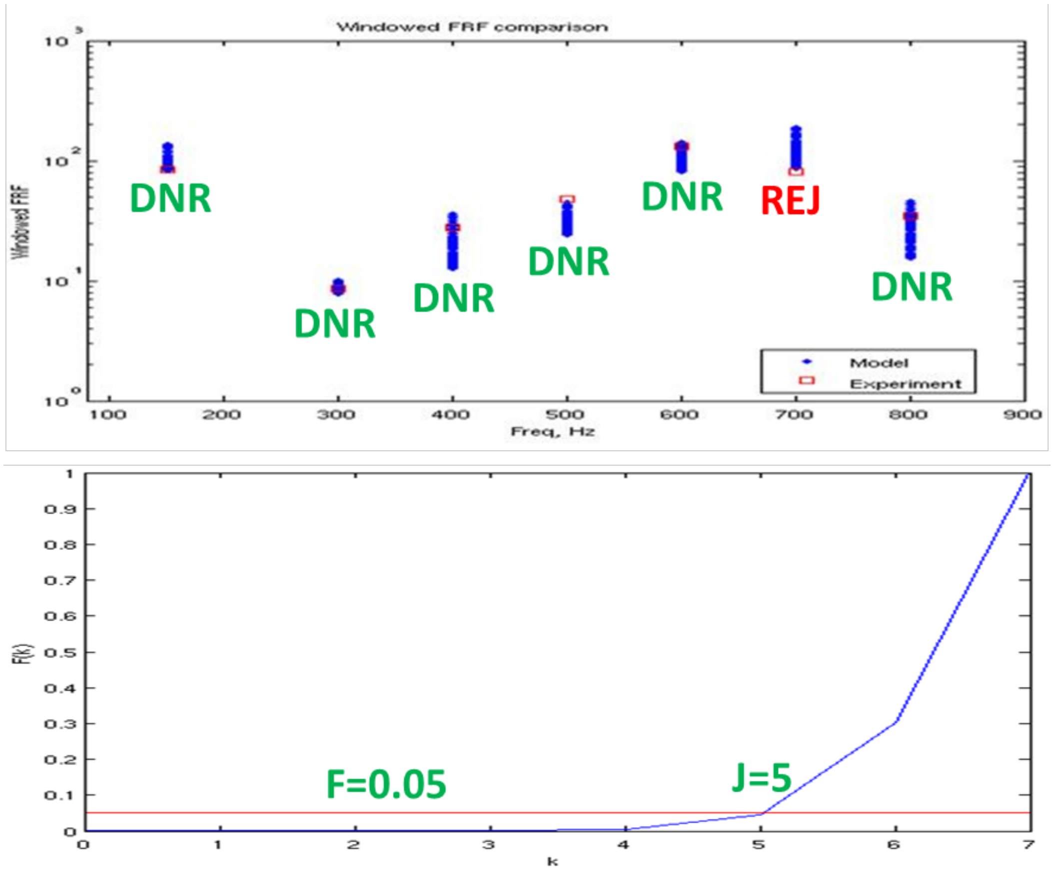

Figure 13. Validation assessment @ 7 window locations

V. Summary

In this paper, we demonstrate a model validation process to assess the accuracy of a model of a complex mechanical system relative to a specific use environment. The assessment process entails: (1) the definition of suitable response measures that are relevant in the anticipated use environment; (2) a criteria to assess whether the model is adequate or not; and (3) a metric by which to assess the difference in the model versus experiments. Of major importance in this assessment is the examination of the uncertainty captured by the modeling and simulation process. The total uncertainty predicted by the modeling and simulation effort is the “sum” (not necessarily a linear combination) of different sources of uncertainty, such as that arising from parametric uncertainty, model form, solution inadequacy, etc. These need to be accounted for if we sought to make the best informed decision relative to qualification of a system to a given environment.

Acknowledgments

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

[1]AIAA (American Institute of Aeronautics and Astronautics), (1998), *Guide for the Verification and Validation of Computational Fluid Dynamics Simulations*, AIAA-G-077-1998, Reston, VA, American Institute of Aeronautics and Astronautics.

[2]U.S. Department of Energy, (2000), “Advanced Simulation and Computing (ASCI) Program Plan,” 01-ASCI-Prog-01, Sandia National Laboratories, Albuquerque, New Mexico.

[3]Paez, T., Hunter, N., and Cafeo, J., (2002), “A Karhunen-Loeve Framework for Modeling Structural Randomness,” Proceedings of International Modal Analysis Conference, XX, Los Angeles.

[4]Paez, T., Hunter, N., and Cafeo, J., (2002a), “Dynamics of Randomly Excited Stochastic Structures,” Proceedings of the Ninth International Congress on Sound and Vibrations, The International Institute of Acoustics and Vibration, July, Orlando.

[5]Ghanem, R., Spanos, P., (1991), *Stochastic Finite elements: A Spectral Approach*, Springer-Verlag, New York.

[6]Ghanem, R., Red-Horse, J., (1999), “Propagation of Probabilistic Uncertainty in Complex Physical Systems Using a Stochastic Finite Element Approach," in Predictability: Quantifying Uncertainty in Models of Complex Phenomena, S. Chen, L. Margolin, D. Sharp, Eds., Special issue originating from the 18th Annual International Conference of the Center for Nonlinear Studies, Los Alamos, NM, USA, 11-15 May 1998.

[7]Urbina, A., Paez, T.L., Rutherford, B., O’Gorman, C.C., Hinnerichs, T.D. and Hunter, P.S., “Validation Mathematical Model: An Overview of the Process”, 2005 Annual Meeting of the Society of Experimental Mechanics, Portland, Oregon, June 2005.

[8]Miller, I., Freund, J. E., (1977), *Probability and Statistics for Engineers*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

[9]Reese, G.M., Bhardwaj, M.K., Segalman, D.J., Alvin, K.F., Driessen, B.J., Pierson, K., Dohrmann, C., Wilson, C.R. and Walsh, T., “Salinas-User’s Notes”, SAND Report 1999-2801, Sandia National Laboratories, Albuquerque, NM, updated 2012.

[10]Ross, M.R., Urbina, A., Simmermacher, T.W. and Paez, T.L., “The Verification and Uncertainty Quantification of Surrogate Models Used for Structural Analysis”, 2012 To Appear in Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Honolulu, Hawaii, April 2012.

[11]Adams, B. M., Dalbey, K. R., Eldred, M. S., and Swiler, L. P., \DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quanti_cation, and Sensitivity Analysi," Tech. Rep. SAND2010-2183, Sandia National Laboratories, August 2011.