SAND2014-19097C

The Role of Model V&V in the Defining of Specifications

Todd Simmermacher, Greg Tipton, Jerry Cap, Randy Mayes
Sandia National Laboratories
PO Box 5800
Albuquerque New Mexico 87123

Abstract

Model validation plays an important role in estimating the confidence in a simulation. Many times, the simulation is the only source of response that is available. When an expensive simulation is developed, the predictions from that simulation will usually be used in an absolute sense, with more conservatism added if the model validation analysis indicates low confidence. The results of the simulation could be scaled based upon the confidence of the model. This paper will describe how the outcome of a model validation exercise determines the conservatism applied to the environmental specifications and the estimation of margins derived from those simulations.

Keywords

Validation, Specifications, Qualification, Uncertainty, Model

Introduction

Model validation and uncertainty quantification has, from necessity, become increasingly embedded into the design and qualification process. Model results are not being used in a vacuum for high consequence decisions, but are compared and formally evaluated against relevant experimental data. The degree and rigor to which the evaluation is performed varies widely based upon programmatic constraints such as budget, schedule, and risk tolerance of the end user of the results.

Model validation and uncertainty quantification has a large body of literature that has advanced the field. The iconic reference is [1]:

> *Verification*: The process of determining that a model implementation accurately represents the developer's conceptual description of the model and the solution to the model.
>
> *Validation*: The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.

The objective of verification is to confirm that the computational representation of the mathematics of the model is, in fact, representing the intended mathematics to some quantified degree. There will be errors due to discretization and convergence of solvers, but these errors should be quantified and included in the uncertainty estimates. The verification of a simulation involves both the analysis code used to develop the simulation, as well as the specific model that is implemented in the code. Verification involves comparison to known analytical solutions and other known solutions to the mathematics that is implemented in the analysis code. The specific model that is developed in the analysis code also requires verification to insure that the assumptions made by the analyst are being correctly represented such that the analysis code provides a valid simulation.

The objective of validation, in contrast, is to evaluate the appropriateness of the assumed computational model of the system of interest. Whereas verification is purely an evaluation of the mathematics of the model, validation is the tie to the physical system and experimental data. Validation involves identifying quantities of interest such as peak velocity, maximum temperature, gRMS (standard deviation of a time signal in the units of acceleration of gravity), or other scalar measures that represent the end use of the model. The results of a model validation effort should also include an uncertainty analysis of the experimental and analysis processes as well as the confidence in the comparison. This information is important to assess the adequacy of the model and to inform the risk affiliated with using the simulation results.

This paper will demonstrate a few examples where the adequacy of the model influences how the model results are utilized. Typically, regardless of the results of the validation, the model is still useful. A very poor model may only be good for relative comparisons between modeled configurations; however, it can still provide valuable insight. Many times, the model is assessed to be fair to good, and can still be used for its intended application. How the model results are used can be affected by the results of a validation. Results from a fair model may be modified by a safety or uncertainty factor whereas results from a good model may be used with no modification.

The first two examples where the adequacy of the model influences how the model results are utilized is for component test specifications. Specifically, this paper will focus on component mechanical shock test specification and component random vibration test specification. These two examples show how simulation results and the adequacy of the results are used for development of component test specifications. The final example in this paper demonstrates where the quality of the model necessary for decision making is determined by the outcome of the analysis. The example is the calculation of failure margins. It is shown that if the predicted margin is large, then a large effort into improving the model and doing model validation may not be worth the expense.

Component Specification

An application that requires a high fidelity, validated model is the derivation of test environments to be used to qualify components for use in a full system. A full system model is developed and simulations of the various environments that have been defined as requirements for the system are used to define component level laboratory tests. These laboratory level tests are designed to represent the in service excitation such that a component can be adequately tested without requiring the hardware to represent the full system.

The environments at the component level can be derived from either a full system physical test or from a computational simulation. Both the test and the computational simulations have limitations in the level of fidelity in replicating the operational environment, however, typically the test results are assumed to be sufficiently representative to develop component test specifications. Computational simulations, however, require model validation for the results to be considered valid for further use. The reality is for many applications, the model will be used unless it is excessively wrong.

How the results from a simulation are used can be affected by the outcome of the model validation [2]. A model that is shown to be highly predictive with little uncertainty will be used with confidence and with little modification. If the model is shown to be a poor representation of reality or with a large degree of uncertainty, a large safety factor or scaling may be applied to ensure conservatism in the final test specification. Declaring a priori how the results of a simulation are going to be used, processed, and modified based upon the fidelity demonstrated by the validation is as important as declaring the model validation metric and acceptance criteria. In this case, identifying the modifications to the data that will be performed on the data for various levels of success is important so that the processing can be agreed upon prior to adjudicating the model. This assures that the process can be developed in an objective manner.

Component Mechanical Shock Test Specification

High level mechanical shock test specifications are simulated on either a drop table or a resonant plate or beam test setup. The shock specifications for either test are shown in Figure 1 and 2. For a drop table, the test specification is in the form of a haversine. The duration and amplitude of the haversine are the parameters to be specified. The duration is determined by the bandwidth of the shock. Normally the specification is developed by enveloping the test or model data in the shock response spectra (SRS) domain. The amplitude and duration of the haversine is adjusted until the SRS of the haversine is higher in amplitude than the data being enveloped. An important feature of the SRS is that the high frequency asymptote has amplitude equal to the peak acceleration of the corresponding time history. It is this characteristic that will be used in the development that follows.

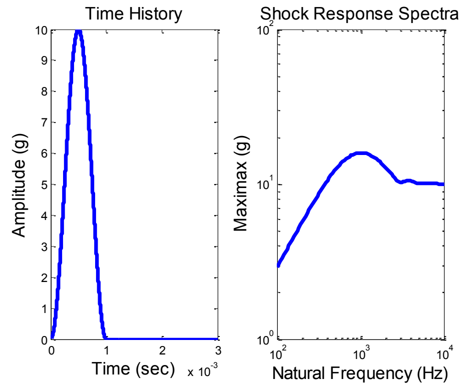

Figure 1 Representative Haversine

A shock can also be simulated on a resonant beam or plate fixture (Figure 2). This shock is appropriate for a pyro shock type application where there is ringing affiliated with the operational environment. The resonant test is performed by fixing the test article to a flat plate or beam, then exciting the plate or beam by firing a projectile at the fixture. As with the haversine, there are two parameters that are needed to define the resonant plate test, the amplitude and the "knee" frequency. The amplitude is determined by the magnitude of the impact. The knee frequency is defined as the break point where the slope changes in the SRS which is 200 Hz in Figure 2. The knee frequency is controlled by the first natural frequency of the beam or plate used as a fixture for the test. By adjusting the size or thickness of the beam or plate, the knee frequency can be tuned.

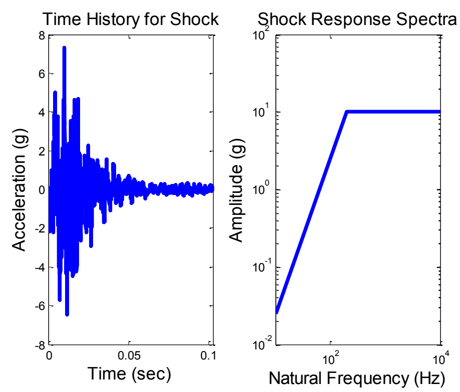

Figure 2 Representative Resonant Plate Specification

For the application described in [2], the duration of the haversine is assumed to have been determined through a combination of test data and model results. By reviewing data from both validation tests and simulation results, the frequency content of the SRS was estimated and was assumed known. The model results were to be used to determine the amplitude of the haversine for use in a component level test specification. The validation of the model was performed at the response level. Each measurement location was judged individually against test data to determine the degree of conservatism present in the model. The model was then used to simulate similar, untested environments and the results were used to derive a composite shock specification for the various components in the system.

The model validation conditions and the resulting acceptance of the data are characterized by the quantities identified in Figure 3. The mean of the quantity of interest (QOI) from the model ($\bar{M}$) and the mean of the QOI from the experiment ($\bar{E}$) are defined relative to the probability distribution of the model and experimental QOI. The uncertainty is given as a standard deviation of the respective distributions. The appropriate standard deviation for which to define U_M and U_E based upon the PDF of the model and experimental uncertainty is negotiated with a team consisting of the customer, test engineers, analysts, and environment engineers.

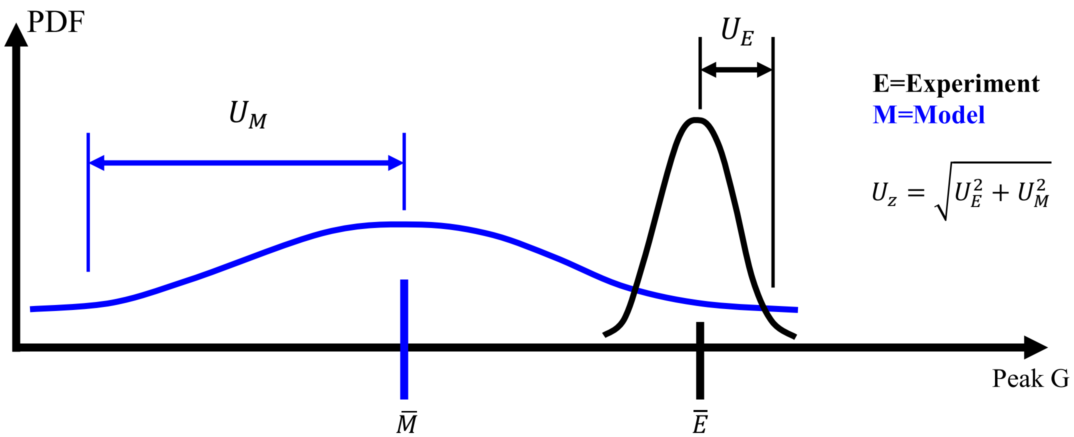

Figure 3 Model and Experimental Probability Density Functions Definitions

The criteria developed in [2] are shown in Figure 4 and 5. The underlying model to be validated is linear, however localized non-linearities exist in the physical system. At the locations where the response is predominately linear, the criterion uses a combination of model and experimental uncertainty. At locations where the response is influenced by a local nonlinearity, the linear model should be conservative. This definition favors a conservative model which is preferable for component qualification. A component that passes an over-test is guaranteed to function in the operational environment while a significantly under-tested component may have an undetected failure prior to reaching the full environment. An overly conservative specification, however, can add unnecessary cost to the design in an effort to overly ruggedize the component. A factor of two in conservatism was negotiated to be acceptable for the application and is based on the rule of thumb for enveloping SRS responses that the peak acceleration of the resulting specification should not be over a factor of two from the underlying data.

For a location where the response is predominately linear, the model is considered valid if the model is within the root mean square of the model and experimental uncertainty below the mean of the experiment or within sum of the mean of the experiment plus the experimental uncertainty above the mean of the experiment (Figure 4). The example in Figure 4 shows a valid model that is slightly non-conservative in predicting peak acceleration. This allowed for the possibility that the model results could under-predict the test, but still be considered valid.

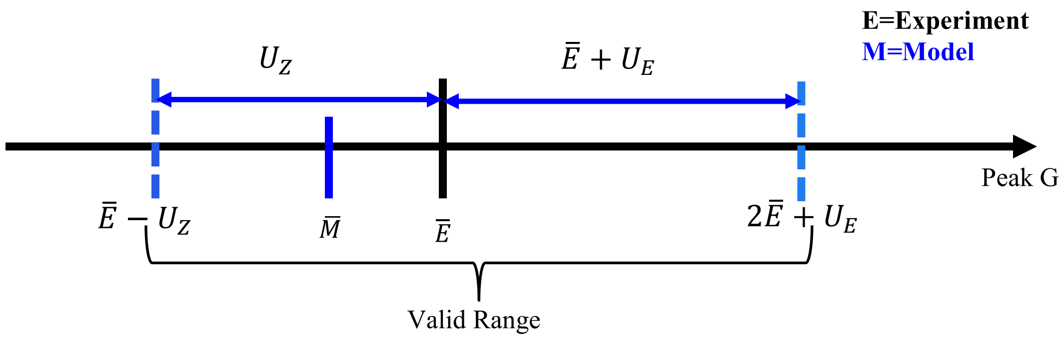

Figure 4 Graphical Representation of Validation Metric for a Linear Model in a Linear Regime

For a location where the response is primarily non-linear (Figure 5), the expectation is that a linear model should predict that location conservatively, and therefore the valid region for the model only spans between the experimental data and the same upper bound as Figure 4. This restriction only allows for the experimental uncertainty at the lower bound.

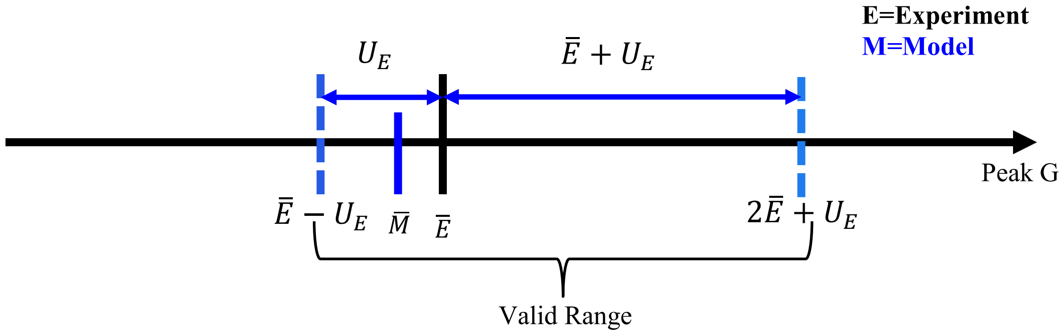

Figure 5 Graphical Representation of Validation Metric for a Linear Model in a Nonlinear Regime

Once the simulation results have been determined to be valid, the amplitude of the haversine can be set. If the valid prediction was higher than the experimental measurement, the peak acceleration was used directly from the model. However, if the valid prediction was lower than the experimental measurement, the amplitude of the haversine was set to twice the predicted level. This insured that the component specification was conservative.

A similar metric could be used to define a resonant plate test. The knee frequency, which is similar to the duration of the haversine, would be determined a priori from legacy data or through analysis of relevant test and model data. The amplitude of the resonant plate specification would be set exactly as described above for a haversine.

Component Random Vibration Test Specification

For random vibration, the test specification is defined in terms of the power spectral density (PSD). The duration of the test specification is determined by either the duration of the operational environment or a scaling based on maintaining the fatiguing nature of the insult that is determined by a variation of Miners fatigue life model. This scaling will compress the time necessary to perform the test by increasing the amplitude of the test specification. The test specification is derived from either operational, laboratory, or simulation data. The specification is defined by simplifying the underlying data by generating a straight line envelope of the data, with some conservatism added to account for variability and uncertainty.

Model validation for random vibration can be performed in a few different ways. A simple metric is a specification on the natural frequencies of the system to be within a certain percentage of the experimental frequencies as determined by a modal test. Typically 5% is assumed to be a reasonable error. A more rigorous metric is based on the least favorable response (LFR) and is developed by [3]. The LFR is attractive because it compares a weighted integral over multiple frequency ranges to determine validity. The integration minimizes the sensitivity of the metric to small errors in natural frequency that can be due to model errors or unit to unit variability.

Once the validity of the model has been established, the data is available to define specifications. Typically, due to uncertainty in the loadings, errors and uncertainty in the model, and variability in the hardware, the simulation results will differ from test data or legacy test specifications derived for a similar system. Since the model results may be the only data available to specify component specifications, processing of the simulation results is needed to normalize the results to a known basis. A method to normalize the data has been developed by [4]. This technique was developed to utilize the model results at locations that were unmeasured during a laboratory or field test. This assumes that the model is compared at measurement locations and the degree of similarity is defined at these equivalent locations. A frequency dependent scaling factor is then defined that corrects the simulation results to the measured results. This frequency dependent scaling is then applied to model results for unmeasured nearby locations. The assumption is that the bias error for nearby locations is similar to the bias error observed at the measured locations.

The correction procedure is demonstrated in Figure 6. The processing begins by either enveloping measured data or using legacy test specifications that currently exists. This is depicted as the red line in Figure 6. The raw model results are shown in blue in Figure 6. A frequency dependent function is defined as the ratio of the red curve and the blue curve evaluated at the peaks. For the example in Figure 6 this gives a value of the function at four frequencies. The ratios are then interpolated to the remaining frequencies in the bandwidth of interest. Finally the model data is scaled using this frequency dependent function with the result given as the black curve in Figure 6. Note that the peaks in the modified simulation data match the envelope of the measured data exactly as defined. The valleys are also corrected. This is critical for a properly defined test specification to insure that the test does not attempt to overdrive these regions of low response.

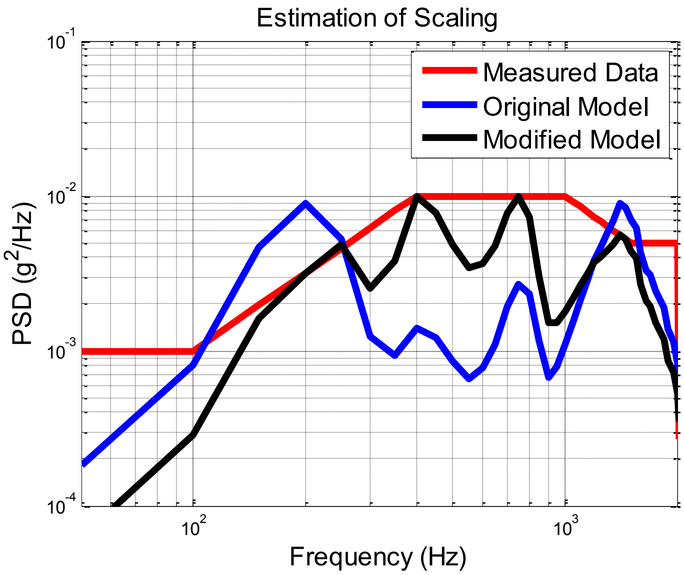

Figure 6 Localized Bias Correction using Peak Scaling

Once the frequency dependent scaling has been defined for a location, this function is assumed to be valid at nearby locations. The definition of nearby is based on engineering judgment and should be determined by a team including the environments engineer, analyst, and customer. The scaling is then used to modify predictions at locations of interest where no test data were collected. Figure 7 shows an example of nearby results scaled using the frequency dependent function determined in Figure 6. Note that the mid frequencies are amplified as a result of the scaling. The model under-predicted this frequency range at the measurement location and this amplification is carried through to the other locations as desired.

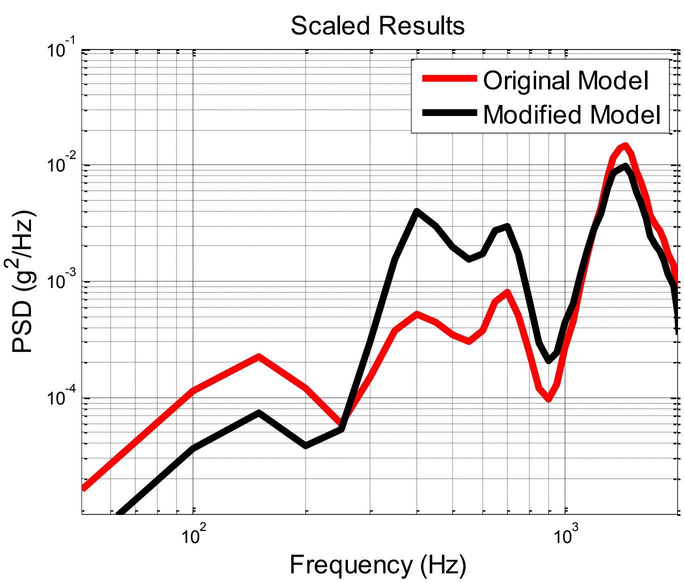

Figure 7 Localized Bias Correction

Quantifying Margins and Uncertainties

The final example is an application where the model results are used to determine the margin in a system. Quantification of margins and uncertainties has become increasing important in high consequence applications [5,6]. An example application is depicted in Figure 8. The requirement is a hard limit based off of some measured operational data. The operational environment is random and has some distribution. The requirement is derived by defining some level of conservatism with some confidence factor. This then identifies the deterministic value of the requirement, in this case peak stress at some location. The model will be used to evaluate the capability of the component by simulating environments much higher than the requirement until a failure is predicted by the model. The probability of failure is then estimated by propagating uncertainty in the parameters through the model and estimating the potential distributions of the peak stress.

The acceptance criteria of the model being deemed valid are determined in a large part by the amount of margin demonstrated in the simulation. In Figure 8, a system with a large degree of margin is shown. The model can have a large amount of uncertainty and the margin would still be large. In this case, a very refined model with minimal uncertainty may be as effective as a cruder model with large uncertainty. If the model is being used solely for margin assessment, an extensive, time consuming model development and model validation effort might not be worth the expense.

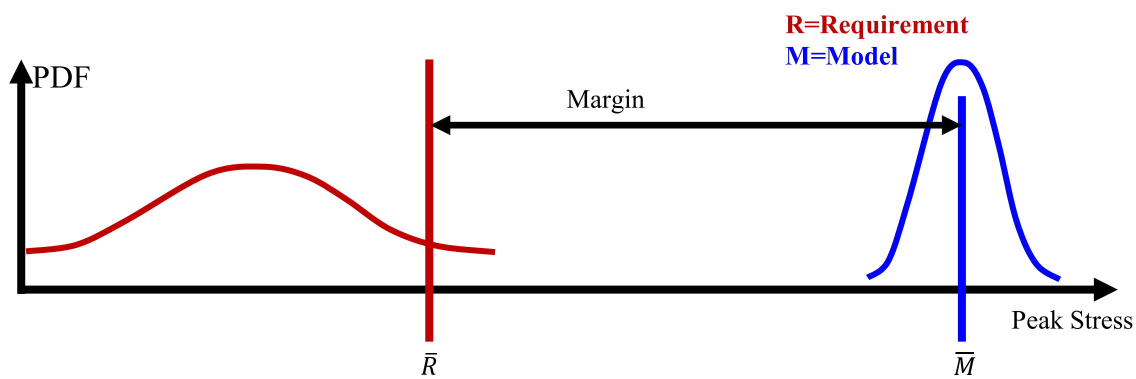

Figure 8 Quantification of Margin and Uncertainty

Figure 9 demonstrates a system with a small degree of margin with respect to the requirement. In this case, the margin is small and the required confidence in the model results is much higher than the situation in Figure 8. Here, large uncertainty in the model could indicate the possibility of negative margin which could imply that the system does not meet its requirements. A large model validation and characterization effort would be demanded for this situation prior to any assessment of the margin.

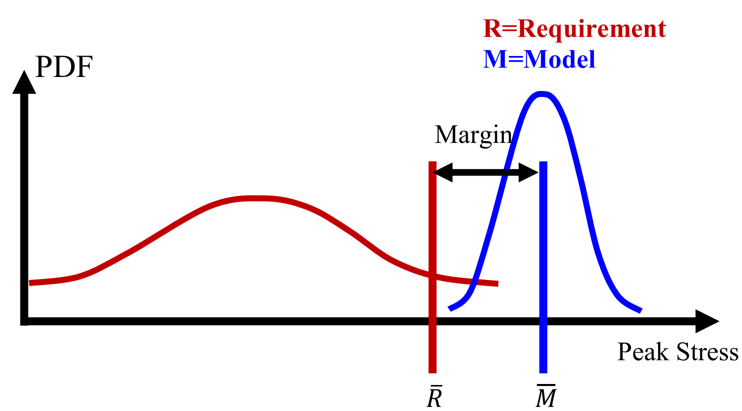

Figure 9 Quantification of Margin and Uncertainty

In the quantification of margin using modeling results, the required confidence in the model is dependent on the demonstrated margin. The results of the simulation determine the acceptance criteria for the model validation. This dependence on the simulation requires an initial estimate of the margin. A rough or initial model might be used to estimate the degree of margin present in the design. Then the path for model development can be planned and executed to the appropriate level.

Conclusions

Three examples of where the outcome of a model validation exercise determines how the data from the simulation is used have been presented. In two of these examples the end result of the model is not a yes/no decision, but the simulation results are used to provide test specifications for designers to assess their components against. The third example shows where the quality of the model necessary for decision making is driven by the outcome of the analysis. These examples again show that the application that the model was developed for drives how the model validation is performed.

Acknowledgement

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, *Guide for the Verification and Validation of Computational Fluid Dynamics Simulations*, American Institute of Aeronautics and Astronautics, AIAA-G-077-1998, 1998.

[2] Mayes, R. L., "Developing Adequacy Criterion for Model Validation Based on Requirements," *Proceedings of the International Modal Analysis Conference*, Orlando, FL, 2008.

[3] Urbina, A., Ross, M., Paez, T. L., "Assessment of the Accuracy of a Model for Use in Predication of Component Environment," *Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Honolulu, HI, 2012.

[4] Cap, J.S., "A Method for Reducing the Effect of Biased Data on the Derivation of Test Specifications," *Sandia National Laboratories Internal White Paper*, May 5, 2014.

[5] Pilch, M., Trucano, T. G., Helton, J. C., *Ideas Underlying Quantification of Margins and Uncertainties (QMU): A White Paper*, Sandia National Laboratories Internal Report, SAND2006-5001, 2006.

[6] Diegert, K., Klenke, S., Novotny, G., Paulsen, R., Pilch, M., Trucano, T., *Toward a More Rigorous Application of Margins and Uncertainties within the Nuclear Weapons Life Cycle – A Sandia Perspective*, Sandia National Laboratories Internal Report, SAND2007-6219, 2007.