For Publication in the Proceedings of IMAC-XIX, the 19th International Modal Analysis Conference.
February 5-8, 2001, Kissimmee, Florida.

MODEL VALIDATION AND UNCERTAINTY QUANTIFICATION

A Special Session of the SD-2000 (Structural Dynamics 2000) Forum

François M. Hemez[1] and Scott W. Doebling[2]
Los Alamos National Laboratory
Engineering Sciences & Applications Division
Engineering Analysis Group (ESA-EA)
P.O. Box 1663, M/S P946, Los Alamos, New Mexico 87545

ABSTRACT

This session offers an open forum to discuss issues and directions of research in the areas of model updating, predictive quality of computer simulations, model validation and uncertainty quantification. Technical presentations review the state-of-the-art in nonlinear dynamics and model validation for structural dynamics. A panel discussion introduces the discussion on technology needs, future trends and challenges ahead with an emphasis placed on soliciting participation of the audience.

One of the goals is to show, through invited contributions, how other scientific communities are approaching and solving difficulties similar to those encountered in structural dynamics. The session also serves the purpose of presenting the on-going organization of technical meetings sponsored by the U.S. Department of Energy and dedicated to health monitoring, damage prognosis, model validation and uncertainty quantification in engineering applications. The session is part of the SD-2000 Forum, a forum to identify research trends, funding opportunities and to discuss the future of structural dynamics.

NOMENCLATURE

The recommended "Standard Notation for Modal Testing & Analysis" is used throughout this paper [1].

1. INTRODUCTION

In many fields of computational sciences, numerical models are developed for predicting the response of a system when the phenomenon is not accessible by direct measurement or when numerical simulations are cheaper than testing. Predictions are also sought after for studying phenomena that can not be tested, for example, failure, catastrophic events or any other occuring in the tails of the probability density functions.

Nevertheless, developing sophisticated physics-based models does not necessarily guarantee accuracy and predictability. It must somehow be verified that the many assumptions involved in the successive steps of idealization, discretization and modeling yield satisfactory predictions. This is known as model validation and it is usually carried out by comparing the predictions of a model or family of models to test data. If the agreement between the two sets is not satisfactory, design parameters can be optimized to improve the predictive quality of the models. In structural dynamics, the conventional approach to inverse problem solving is the finite element model updating technology that has been extensively studied for many decades [2]. Most techniques documented in the literature are formulated in the frequency domain and apply to linear systems and stationary responses. It is only in the past decade (1990-2000) that serious attempts have been made at extending this technology to nonlinear systems and transient signals [3].

As structural dynamics becomes increasingly non-modal, stochastic and nonlinear, the finite element model updating technology evolves into the broader notions of model validation and uncertainty quantification. For example, particular re-sampling procedures must be implemented to propagate variability information through a forward calculation; non-modal features must be defined to analyze nonlinear data sets [3]. It is attempted to show how other scientific communities (physics, statistics, computational biology, climate science, etc.) are approaching and solving difficulties similar to those encountered in structural dynamics. The session consists of a combination of technical presentations, panel discussions and open discussions where the active participation of the audience is encouraged.

[1] Technical Staff Member, hemez@lanl.gov, 505-665-7955 (Phone), 505-665-2137 (Fax), Member AIAA, SEM.
[2] Technical Staff Member, doebling@lanl.gov, 505-667-6950 (Phone), 505-665-2137 (Fax), Member SEM.

Approved for public release on October 17, 2000. ***LA-UR-00-4873 — Unclassified.***

Presentations introduce the subject, review the technical issues and motivate the discussion. Part of the open forum is focused on discussing predictability, model validation, uncertainty quantification and related issues. The session also serves the purpose of presenting the on-going organization of technical meetings scheduled throughout 2001-2002, sponsored by the U.S. Department of Energy (DOE) and dedicated to health monitoring, damage prognosis, model validation and uncertainty in engineering applications.

In the remainder, current trends in modeling and predictability are discussed. The intent is to show that what can be observed in structural dynamics is not specific to our community. Also, we would like to promote the idea that experience may be gained from learning what is being achieved in other scientific communities. The discussion is then focused on the evolution of structural dynamics and consequences in terms of model validation needs. To conclude, several research trends are briefly introduced.

2. WHAT MODEL FOR WHAT PURPOSE?

The dynamics of systems commonly analyzed in most computational sciences is strongly influenced by the nature of particular partial differential equations. When complex phenomena are studied, the evolution from the system's initial conditions typically exhibits a separation of scales behavior. An example is the modeling of wild fires where small-scale phenomena must account for the turbulent nature of fire while large-scale phenomena exhibit coherent structures that mathematical operators such as the Laplacian may represent with satisfactory accuracy. In structural dynamics, this certainly applies to the phenomena by which energy is dissipated in a structure. For example, the Coulomb damping model provides a deterministic, large-scale description while the phenomenological behavior is highly stochastic and represented by the so-called "stick and slip" at the microscopic level. Statistical models explain these behaviors by coupling mean-field theories to large deviation principles that characterize the system's most probable states.

This remark brings us to our first point. The total predictability that may be expected from a particular model depends on the purpose intended for that model. Generally, the traditional approach for model validation is stated as follows:

> **"My model is valid because it reproduces the test data with adequate accuracy…"**

Nevertheless, reproducing test data does not guarantee predictability away from the region in the design space that relates to the test data. Also, this approach may be irrelevant when it comes to phenomenological models or statistically accurate models that do not provide deterministic outputs. Similar trends can be identified in other scientific communities such as physics, statistical sciences, biology and climate science. It can be further observed that the tools being applied to model validation problems are quite different from the tools generally used by structural dynamicists. In Figure 1, an overall description is provided of the techniques involved to validate computer simulations.

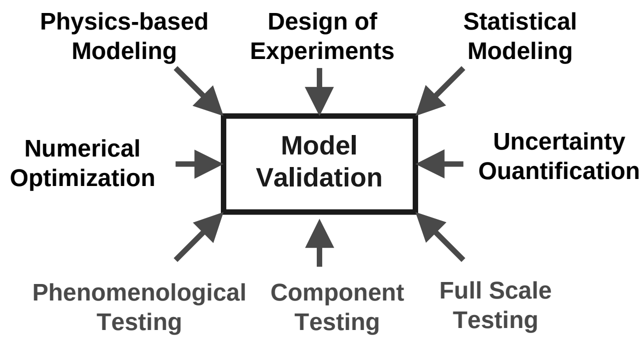

Figure 1. Conceptual view of model validation.

It is our opinion that the focus of the research in model validation should be shifted from validating deterministic models to validating statistically accurate models. This need arises increasingly because environmental variability and other sources of uncertainty in manufacturing tolerances and assembly procedures must be accounted for to fully capture the spectrum of behavior of the systems analyzed (from nominal response to catastrophic failure). Therefore, the concept of model validation should be strongly coupled to uncertainty quantification, a relationship that has generally been overlooked by the conventional finite element model updating technology.

3. WHAT DOES IT TAKE TO BE PREDICTIVE?

Even if very complex numerical simulations can be performed on today's most powerful computing platforms, the central question remains: What does it take to be predictive? The five elements generally mentioned as being critical when it comes to assessing the predictive accuracy of a model are:

1) **The geometry;**

2) **The physics;**

3) **The sources of uncertainty;**

4) **The model sensitivities;**

5) **The outcome of the model.**

Approximating the geometry and the physics remains an issue of importance in many scientific fields such as wild fire modeling, traffic modeling or global climate prediction. In structural dynamics however, the capabilities are generally available to represent any geometry at any precision level. Similarly, the physics of the systems dealt with is well described, at least at the continuum level, by the equations of solid mechanics and fluid dynamics. Therefore, our discussion will focus on other aspects even if it is acknowledged that significant research efforts are currently being spent in areas such as multi-scale, high-fidelity material modeling.

Modeling uncertainty and calculating the model's sensitivities (or estimating the statistical correlation of an output y_j to an input p_i) may offer significant computational challenges when nonlinear, stochastic models are involved. Similarly, defining the outcome of a model assumes that relevant features and metrics can adequately assess its purpose. Analysts dealing with complex numerical simulations that generate several Giga-bytes of output may be overwhelmed by the amount of data produced. Data compression and pattern recognition tools then become key components of the analysis.

4. SPECULATIVE OUTLOOK

When analyzing the dynamic response of a complex system using the finite element method, it is not acceptable to neglect the contribution of an important component, joint or interface. In the past, neglected dynamics were accounted for by tuning parameters in the model to agree with the experimental data. For example, the damping (modal or other) was determined "ad hoc" using test data obtained from testing the fully assembled system. Then, the identified damping properties were added to the model to improve its predictive accuracy. At present, some of the full-scale testing capabilities which formerly existed at the U.S. national laboratories and many other facilities in the automotive, aerospace and civil engineering communities are no longer functional. Therefore, it is no longer possible to reconcile a model with experimental data for all environments. In the future, models will be constructed with limited use of these expensive, full-scale test data sets.

To the evolution of testing facilities and practices must be added the current trends in modeling and analysis. Figure 2 illustrates what, we believe, will be a typical structural dynamics application of the 21st century. It represents a tri-axis MEMS micro-sensor developed for health monitoring. Although this system should not be referred to as a "structure" (because its primary purpose is not to carry loads), structural dynamics clearly plays a central role in reliability analysis, thermal and electromagnetic modeling.

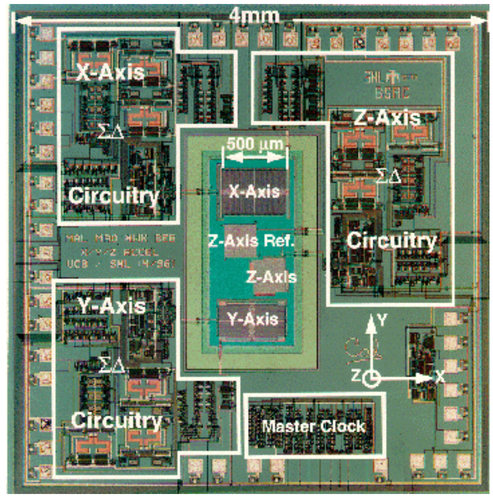

Figure 2. MEMS sensor, http://www.imi-mems.com. (Courtesy of Integrated Micro Instruments.)

This example illustrates that structural dynamics with its interaction with other scientific domains and emergent technologies will become increasingly:

1) **Nonlinear;**
2) **Non-structural;**
3) **Non-modal;**
4) **High bandwidth;**
5) **Multi-physics.**

In these conditions, can the concept of finite element model updating that has been developed for linear, modal dynamics be generalized? Is model updating the correct answer to model validation? What "features" other than the conventional frequency response functions, mode shapes and resonant frequencies can be extracted from the data to characterize the response of a nonlinear system? How to quantify the total uncertainty of an experiment? How to propagate the parametric uncertainty of a numerical simulation? These are some of the questions that we try to address in the session.

5. MODEL UPDATING VS. MODEL VALIDATION

We would like to emphasize that model validation should be thought of as a broader concept than model updating. A numerical simulation is not necessarily validated after the output has been compared to test data and the model has been updated. Instead, it is generally agreed upon that new, well-thought strategies must be established for model validation. They integrate tools such as component testing, full-scale testing, test-analysis correlation, design of computer experiments, statistical analysis and finite element model updating (see Figure 1).

For validating computer models, it is generally agreed upon that errors caused by our imperfect knowledge of "separable" physics (that is, effects that can be decoupled from each other) should be identified first. Then, the sources of variability and modeling errors that may result from the successive steps of system integration can be identified and corrected. At the separable physics or continuum levels, phenomena are generally complex but dedicated and well-controlled testing procedures can be defined. At the sub-assembly or full-scale levels, testing is difficult and variability may be a concern but few unknowns remain to be inferred from test data.

In addition to recognizing that a model must be gradually validated, great attention should also be paid to the operating conditions and the model's purpose. Clearly, two different experiments and probably two different models must be developed when the same component is subjected to random vibrations or shock response. The purpose of a model (that is, what the model needs to predict) is also of paramount importance because it dictates the features and metrics on which the validation should focus. As a result, model validation must be thought of in terms of a matrix of experiments rather than a single test-analysis correlation study. An example is provided in Figure 3 that illustrates the coupling between models and applied loads. To be complete, a third axis that would represent the model's purpose should be added.

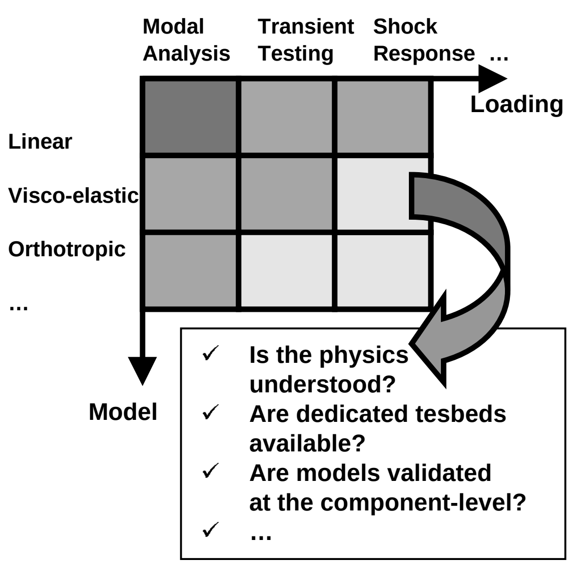

Figure 3. Matrix of model validation experiments.

An example of practical implementation of this paradigm is the validation of complex engineering simulations performed at Los Alamos National Laboratory for the Accelerated Strategic Computing Initiative (ASCI) program. The application illustrated in Figures 4-5 represents the highly transient response of a threaded joint assembly due to explosive loading. The explicit finite element model features 1.4 million elements, 480 contact pairs and more than 6 million degrees of freedom. Nonlinearity arises in the form of pre-load, contact mechanics, material modeling and thermal coupling. When running on an ASCI platform with 504 dedicated processors, one hour of CPU time is required to simulate 10^{-3} second of response.

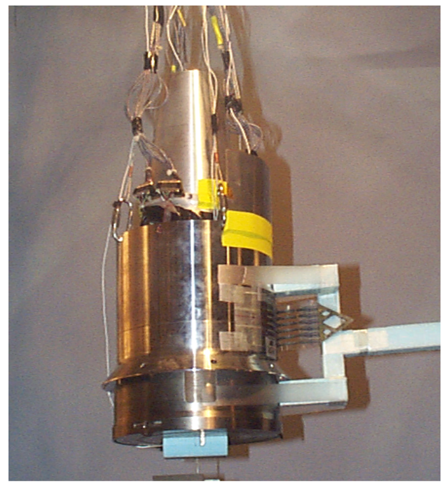

Figure 4. LANL forward mount test.

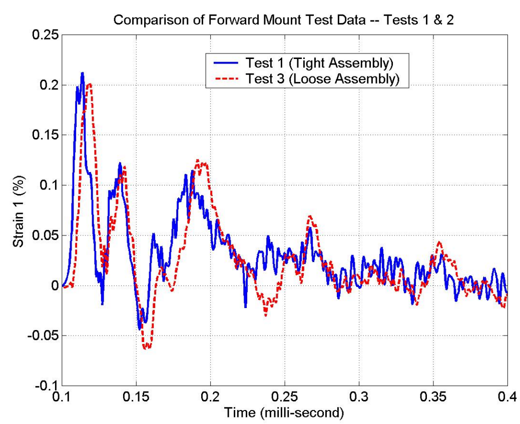

Figure 5. Strain responses (sensor 1) for two tests.

Full-scale, explosive testing has been performed and multiple test data sets are available for estimating the degree of repeatability of the experiment. Design of experiments, statistical analysis and model validation are being implemented to identify specific joint properties as well as the degree of variability of the assembly. Figure 4 pictures the system instrumented prior to detonation and Figure 5 illustrates the variability obtained from tests with different assembly tolerances. Future plans include scheduling more tests at higher impulse levels and assessing if the updated numerical model can predict the dynamics of the system at input levels different from those used during the updating. Preliminary results are presented in Reference [4].

The ultimate task of verifying that predictions of the optimized model are correct remains a challenging one. This is nothing less but the old mathematical dilemma between interpolation and extrapolation. Our opinion of this issue is that model validation does not exist. There is only model "invalidation" as demonstrated by Pearson's work on hypothesis testing [5], that is, a model may be considered correct as long as it can not be proved wrong. Practically, this implies that:

1) **Test-analysis correlation must be able to discriminate discrepancies caused by environmental variability, experimental and modeling uncertainty from those caused by parametric modeling errors;**

2) **The consistency between different models must be assessed when different features and metrics are used to define the optimization's objective functions;**

3) **Data sets not used during the updating step are required to assess the predictive quality of a model;**

4) **Probabilities must be assigned to each model developed to reflect the degree of confidence (or lack of confidence) in its prediction.**

6. DIRECTIONS OF THE RESEARCH

To conclude this discussion, five issues are briefly introduced that seem critical to the success of model updating, uncertainty quantification and model validation for linear and nonlinear dynamics.

1) Uncertainty quantification.

The success of any model validation depends on the ability to quantify uncertainty. The current approach in statistical sciences is to analyze the error of the model output. This is not efficient for identifying the sources of discrepancy between test and analysis results. Instead, the uncertainty should be built at the beginning of the analysis, then propagated through the forward resolution. One potential approach is Bayes inference [6] that estimates the posterior probability, that is, the probability of the model **{p}** given data **{y}**. What is therefore important is not necessarily that the correlated models reproduce the responses measured during a single test but that they predict the response levels with the same probability of occurrence as the one inferred from test data.

2) Sampling and fast probability integration.

The notions discussed here rely strongly on the capability to propagate uncertainty and/or variability throughout an analysis. For large-scale applications featuring nonlinear models, Monte Carlo simulations remain computationally too inefficient when it comes to predicting unlikely or catastrophic events, which is one of the main reasons for carrying out an analysis. Stochastic finite element techniques [7] and fast probability integration methods [8] must therefore be developed and interfaced with engineering codes. Accelerated sampling methods such as the Latin Hypercube sampling [9], Taguchi arrays and orthogonal array sampling [10] are efficient alternatives.

3) Generation of fast running meta-models.

Efficient numerical optimization requires that the objective functions be obtained at low computational cost. Therefore, fast running models or meta-models must be generated to replace the expensive, large-scale simulations. One difficulty of fitting meta-models is efficient sampling, that is, the generation of sufficient information in regions where the feature's joint probability density function is maximum. This issue has been the focus of recent advances in the statistics [11].

The second direction of open research is the implementation of probabilistic meta-models when the objective of model validation is to account for sources of variability in the experiment and the numerical model. Stochastic processes can also be included to propagate other sources of discrepancy between test and analysis data such as numerical and truncation errors or to bound the experiment's total uncertainty. This procedure, well-known in the geo-physics community, is progressively being tested and applied in structural dynamics [12].

4) Feature extraction.

Large computer simulations tend to generate enormous amounts of output that must be synthesized into a small number of indicators for the analysis. This step is referred to as data reduction or feature extraction [13]. These features are typically used to define the test-analysis correlation metrics optimized to improve the predictive accuracy of the model. The main issue in feature extraction is to define indicators that provide meaningful insight regarding the ability of the model to capture the dynamics investigated. Features that are used to analyze nonlinear, transient data sets include: the RMS error of time series; the principal component decomposition; the shock response spectrum; AR, ARX and ARMA-based features; the power spectral density; higher-order statistical moments; and probability density functions.

5) Statistical hypothesis testing.

Another issue of open research is the problem of establishing a correlation between multiple data sets. By this we mean, "assessing the degree to which two populations are consistent with each other." Such statistical consistency can be assessed using the Mahalanobis distance and a standard, multivariate Hotelling's T^2 test. These statistics, however, can only compare the mean of two distributions. One of the only possibility available for testing both mean and variance is to calculate Kullback-Leibler's relative entropy defined as the expected value of the ratio between the probability density functions of the two populations. These statistics are attractive because they are independent of the parent distribution.

The computational requirements associated with this procedure may become very important because the probability distribution of each feature considered for test-analysis correlation must be assessed for each candidate design evaluated during the optimization. This, however, is the only possibility to guarantee at a given confidence level that the numerical simulation is validated in the context of uncertainty propagation.

ACKNOWLEDGEMENTS

The authors are grateful to the many colleagues who have expressed interest and helped them organize this session. Among them, we recognize Ken Alvin, Sandia National Laboratories; Mark Anderson, Los Alamos National Laboratory; Tim Hasselman, ACTA Inc.; Michael McKay, Los Alamos National Laboratory; John Schultze, Los Alamos National Laboratory; Hoon Sohn, Los Alamos National Laboratory; and Keith Worden, University of Sheffield.

REFERENCES

[1] Lieven, N.A.J., and Ewins, D.J., "A Proposal For Standard Notation and Terminology in Modal Analysis," *10th IMAC, International Modal Analysis Conference,* San Diego, CA, Feb. 2-5, 1992, pp. 1414-1419.

[2] Mottershead, J.E., Friswell, M.I., "Model Updating in Structural Dynamics: A Survey," *Journal of Sound and Vibration*, Vol. 162, No. 2, 1993, pp. 347-375.

[3] Hemez, F.M., Doebling, S.W., "Inversion of Structural Dynamics Simulations: State-of-the-art and Orientations of the Research," *25th ISMA, International Conference in Noise and Vibration Engineering*, Leuven, Belgium, Sep. 13-15, 2000, pp. 403-413.

[4] Butler, T., Doebling, S.W., Hemez, F.M., Sohn, H., "Model Validation For a Complex Jointed Structure," *19th IMAC, International Modal Analysis Conference,* Feb. 5-8, 2001, Kissimmee, FL.

[5] Neyman, J., Pearson, E.S., "On the Problem of the Most Efficient Tests of Statistical Hypotheses," *Philosophical Transactions of the Royal Society,* Series A, Vol. 231, 1933, pp. 289-337.

[6] Hanson, K.M., Cunningham, G.S., Saquib, S.S., "Inversion Based on Computational Simulations," *Maximum Entropy and Bayesian Methods,* Kluwer Academic, Dordrecht, Germany, 1998, pp. 121-135.

[7] Red-Horse, J., Paez, T.L., Field, R.V., Romero, V**., Non-deterministic Analysis of Mechanical Systems**, Report #SAND2000-0890, Sandia National Laboratories, Albuquerque, NM, April 2000.

[8] **NESSUS**, User's Manual, Version 2.3, Southwest Research Institute, San Antonio, TX, 1996.

[9] McKay, M.D., Beckman, R.J., Conover, W.J., "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code," *Technometrics,* Vol. 21, No. 2, 1979, pp. 239-245.

[10] Hedayat, A.S., Sloane, N.J.A., Stufken, J., **Orthogonal Arrays: Theory and Applications**, Springer-Verlag, New York, NY, 1999.

[11] McKay, M., "Sampling Variability of Measures of Input-Variable Importance in Computer Models," *3rd DOE/MICS Workshop on the Predictability of Complex Phenomena,* Los Alamos, NM, Dec. 6-8, 1999.

[12] Rutherford, B., "A Re-sampling-based Approach to Optimal Experimental Design for Computer Analysis of a Complex System," *3rd DOE/MICS Workshop on the Predictability of Complex Phenomena*, Los Alamos, NM, Dec. 6-8, 1999.

[13] Bishop, C.M., **Neural Networks for Pattern Recognition**, Clarendon Press, Oxford University Press Inc., New York, NY, 1998.