EDITORIAL

Verification and Validation

Randall J. Allemang, Contributing Editor

Scientists and engineers working in all areas of engineering frequently use the nomenclature of verification or validation to reflect that a theory or model has been confirmed through some sort of test procedure involving experimental data. Just as frequently, we have been using the nomenclature somewhat incorrectly in light of the rigor of the emerging field of *verification and validation*, or V&V.

This emerging field of verification and validation is the subject of a number of books over the last 10-15 years and recent guidelines published by ASME and AIAA. This topic is regularly part of several sessions in this year's IMAC conference sponsored by the Society of Experimental Mechanics (SEM) where there is a technical division (Model Validation and Uncertainty Quantification) that focuses on the area. Likewise, there is a new ASME symposium on verification and validation that will be held for the second time in May of this year.

Certainly, all of us in the sound and vibration community need to be aware of this development and begin to understand the importance and impact of this topic. The end goal of a more formal statement and understanding of verification and validation is the need to quantify the uncertainty in our models and experimental data. The verification and validation framework yields a structure for such quantification, which is always desirable based on increasing requirements for some level of risk mitigation.

While words written here cannot completely explain the rigor of the V&V process, perhaps I can trigger more interest and an initial understanding of the topic. Much of the development of the V&V concept and structure has evolved from significant research programs of the last 50 years in the aerospace and automotive areas as well as the need to give structure to research topics where modeling and experimental data are more limited, such as environmental contamination of groundwater, weather prediction and effects of catastrophic events.

In the area of sound and vibration, this would include modeling and testing in extreme environments where multiple inputs include conventional loads at low and high levels, temperature, fluid pressure, fluid flow and humidity involving nonlinear and coupling effects that are not well understood due to gaps in knowledge or limitations of testing facilities.

Ultimately, in a world that increasingly relies on modeling rather than testing, and with these gaps and limitations in mind, it is envisioned that the end requirement will be validation of models and that such validation will be conducted within the more rigorous framework of V&V. Specifically, verification refers to numerical experiments designed to detect errors in software (code) or numerical limitations in software (*Are we solving the equations correctly?*) and validation refers to the physical experiments designed to evaluate the results of the simulations (*Are we solving the correct equations?*)

While this framework has become increasingly more structured since 1990, this field has evolved considerably even in the last 10 years. The V&V framework now has evolved to focus more on *uncertainty quantification* (UQ) and more recently has moved to include *quantification of margins and uncertainties* (QMU). Within the nomenclature, *margin* refers to the difference (distance) between the design requirement and the mean of the expected or measured result, and *uncertainty* refers to the variation about the mean of the expected or measured result, often with respect to the plus/minus two sigma information.

Within these definitions, many experimentalists feel that their primary concern is, therefore, validation. While my personal interest is mostly from the experimental side of the V&V issue, it is important to recognize that parameter estimation models used to extract information from experimental data and the verification of those models is a major concern. Also, many modern experimental sensing methods, such as digital image correlation (and to a certain extent, laser interferometry), are software based and require verification. Within the published V&V guidelines and textbooks, there seems to be a noticeable lack of recognition of this sort of embedded verification requirement in experimental data.

To fully understand the V&V concept, it is important to consider that V&V methodology includes a hierarchical procedure that uses a building-block structure to modeling and testing. This is often misunderstood as a bottom-up process that begins with material, moves to components, further moves to subsystems, and then finally moves to a total system. Instead, current V&V methodology involves the same building blocks but is driven by a top-down process. In fact, planning for validation begins before almost any other step in the V&V process, with *project validation* methodology decisions initially, proceeding to *scientific validation* methodology as the process develops.

While this may at first seem backward, it is exactly the process used informally in most engineering and scientific projects today. The end result of the project is always the starting point. V&V ideas just require that more distinct recognition of the experimental and analytical methods and metrics be decided before the process begins rather than during or after the process has started.

For experimentalists, while all aspects of V&V are important, validation is the last step in the V&V process and possibly the most important. It is significantly tied to modeling and not just experimental testing. Validation is the culmination of a process that requires that numerous criteria first be identified. These criteria include: model use and purpose, validation experiments, conceptual models, mathematical models, computational models, response measures of interest, validation metrics, comparison domains, calibration experiments, and adequacy of validation requirements.

One key part of the validation process is the design of relevant experiments. These experiments must be identified early in the validation process so that the models developed can be sufficiently detailed to have the needed fidelity at the applicable temporal and physical scales. However, since this is a case of needing to know what you do not know in order to proceed, a structured, hierarchical approach to this problem is a mechanism for providing an initial strategy for the validation plan. Fundamental to the V&V process is the need to review and possibly restart the process as new information is gained at each step of the process.

This need to revisit the previous work as new information comes to light is troublesome to many scientists and engineers who are used to a fixed schedule (probably somewhat of an illusion anyway). The V&V process is considered to be a re-entrant process requiring a review and revisit of previous modeling and physical testing when new or unexplained information is discovered. This may require repeating earlier activity in the V&V process. Naturally, as gaps in knowledge, science and technology are addressed, the overall V&V plan will need to be updated and refined. This may require that some very expensive or time-consuming experimental testing be repeated. Some situations may be resolved through a substantially increased use of numerical experiments. In the end, rather than going back and repeating earlier experiments as required by the re-entrant structure, this may be addressed by changing the acceptable levels of uncertainty (QU and QMU).

From an experimental point of view, you may be wondering where the experiments fit into the V&V structure. While much of the emerging V&V structure has been driven from the analytical side, experiments are integral to the process. Note that with respect to the rigor of the evolving V&V nomenclature, the aspect of using different models or analytical methods to validate a model, these analytical procedures and comparisons are also referred to as experiments. If we restrict our discussion to experiments that involve physical testing and sensors, the lowest levels of physical experimentation are *exploratory (discovery) experiments* that are designed to assist in determining what physics model is most appropriate for the system in light of the required environments.

The next level of physical experimentation includes *calibration experiments*, which are designed to develop correct model order, verify the parameters in the models, and assist in the quantification of uncertainty associated with the probable environments, models and also with the physical experiments. Another possible level of physical experimentation includes *qualification experiments,* which are physical experiments that are required to measure whether certification/qualification standards are met, if these standards exist. The final level of experimentation includes *validation experiments*, which are designed to compare results between the analytical model predictions and the measured data.

This series of experiments must be planned and conducted at each of the levels of the hierarchical structural system, and each series of experiments will need to satisfy a predetermined and refined set of the criteria (validation metrics). It should be clear that validation experiments are separate from the other physical experiments, and should be blind comparisons with model simulations, if possible, and may or may not include all of the application environments. For this reason, the terminology of *application domain* and *validation domain* has evolved to reinforce that it may not be possible to design physical validation experiments that reflect the entire application domain.

Focusing on the experimental side of V&V even further, utilizing traditional physical experiments, is a desirable and worthy goal for the validation aspects of V&V. But, the limitations of existing test facilities and the financial outlays associated with building new test facilities may preclude comprehensive testing beyond some level. Nevertheless, validation will need to be performed wherever possible at material, part/component, subassembly and full-assembly levels to manage risk and quantify margins and uncertainty. Current verification and validation strategies do include methodology for assessing overall risk in situations where validation experiments do not span the complete application domain through the utilization of *expert panel elicitation*. Expert panel elicitation is being used to establish the scientific validation metrics from project validation metrics. More recently, the definition of required statistical risk functions are derived from truth/plausibility/belief techniques (using Bayesian methods).

Some of the open issues with current validation procedures involve the practical application of the V&V methodology to realistic, complicated problems. Current validation methodology usually requires an *a priori* definition of validation metrics. The validation problems associated with large and complicated systems, limits the ability to plan discretely for validation experiments early in the activity. However, initial validation plans will focus on project validation metrics to define where effort needs to be concentrated (largely where the largest unknowns or limitations exist). This will allow the scientific validation plans to be developed at lower levels. As much of the science and technology is developed, the validation plan will need to be updated correspondingly at each level. The scientific validation will concentrate on multiphysics validation strategies that recognize the ability to incorporate model changes as the structure evolves, either through scale or fidelity during development or through degradation during service.

These types of realistic problems usually require multiple physical experiments with many measurements in each experiment. Defining a specific set of validation metrics under a specific set of environments is quite complicated. Current validation methodology usually involves more than one validation metric and may involve hundreds or thousands when all hierarchical levels are accounted for.

Validation tests and metrics should be jointly determined by both simulation and physical test experts although the actual process should be conducted independently (blind). An acceptable match between the model simulation and the physical experiment is preferably defined in terms of engineering units. It is common to define metrics based on the areas (2D), volumes (3D) or generalized space (ND) characterized by the difference between the simulation and physical experiment, but the question of how to combine metrics involving the same and different engineering units into a higher level metric remains a problem. Lack of experience in defining, using and refining these validation metrics is a significant problem.

Finally, although validation guidelines have been structured for solid mechanics models and for fluid dynamics models, the integration of multidiscipline models has not been specifically addressed in the current professional society guidelines. Excellent references have recently been published or updated concerning V&V, and recent reports document the concerns and details of uncertainty quantification and quantification of margins and uncertainties. Be prepared for even more nomenclature as you jump into the area – things like epistemic and aleatory uncertainty. It is clear that the V&V methodology will be part of our future, but much remains to be defined. Don't be surprised if your future work becomes part of the evolving definitions and procedures.

I hope I have not offended too many of the rigorous verification and validation practitioners while trying to summarize what the V&V field means and has to offer those of us in the sound and vibration community. I hope it gives you something interesting to think about and, as always, I value your comments on verification and validation techniques. If you have comments, please feel free to contact me at: randall.allemang@uc.edu. SV