A FRAMEWORK FOR ASSESSING CONFIDENCE IN COMPUTATIONAL PREDICTIONS

Kenneth M. Hanson[1] and François M. Hemez[2]

Los Alamos National Laboratory
Los Alamos, NM 87545

INTRODUCTION

This article is the third in a series of papers concerning the importance of simulation code validation to the US Department of Energy Accelerated Strategic Computing Initiative (ASCI) program [1]. The series started with a review by John Garcia of the critical need for advanced validation techniques in the ASCI program, which was created to make up for the absence of nuclear testing through the use of simulation codes. Without testing, the simulation codes must be able to answer critical questions about the reliability of our aging stockpile of weapons. In the second paper, Bill Oberkampf gave an overview of validation concepts and described the requirements for a well-executed validation experiment. In this article we discuss the analysis of data obtained from validation experiments and motivate the use of uncertainties to quantify the accuracy of predictions made by simulation codes. This work represents merely a small fraction of the numerous verification and validation projects currently being conducted at the DOE National Laboratories and at several universities under the auspices of the ASCI program.

Engineers routinely use simulation codes to analyze and design critical structures and devices. Because public safety is often involved, confidence in the predictions made by simulation codes is clearly of paramount interest. An engineer needs to be confident that when used in an appropriate way, a simulation code will predict the behavior of the system under study to a specified degree of accuracy. The goal in validating a simulation code is to determine the degree to which the output of the code agrees with the actual behavior of a physical system in a specified situation. Because the criterion is real-world behavior, validation must involve comparison of the simulation code's output to experimental results.

Uncertainties in a quantity are described in terms of a probability density function (pdf) that specifies the probability of all possible values of that quantify. In this context, probability is used as the quantitative measure of our degree of belief, which summarizes our knowledge about a particular situation [2]. We use the Monte Carlo technique to make this kind of probabilistic analysis more tangible as well as to obtain quantitative estimates of uncertainties.

We will discuss the analysis of validation experiments, the role of uncertainties in material models, and in experimental conditions. From the viewpoint of uncertainties, inference about how well a simulation code can predict physical phenomena is limited not only by the uncertainties in the relevant measurements, but also by how well the conditions of the

[1] Methods for Advanced Scientific Simulation Group (CCS-2), 505-667-1402 (Voice), kmh@lanl.gov (E-mail)
[2] Engineering Analysis Group (ESA-EA), 505-665-7955 (Voice), hemez@lanl.gov (E-mail)

experiment are controlled, measured, and documented. We mention the possibilities for inferring unmeasured quantities, including the parameters in material models from comparisons to experiment. The ultimate goal of validation might be to have a simulation code assign uncertainties to its output.

MONTE CARLO CALCULATIONS

The basis of the Monte Carlo technique is to represent a defined probability density function (pdf) in terms of random samples drawn from the pdf. The underlying idea is to randomly pick values of a parameter such that the histogram of chosen values approximates the pdf. Suppose that a pdf is specified as a normal (Gaussian) distribution with a mean value of 100 and rms deviation of 10. Using a standard algorithm for drawing random numbers from such a distribution, we might get the following sequence of numbers, 104, 97, 89, 112, 94, etc. Figure 1 shows the histogram of the values from such a random sequence. As the number of random numbers increases, the histogram looks more and more like the specified normal distribution.
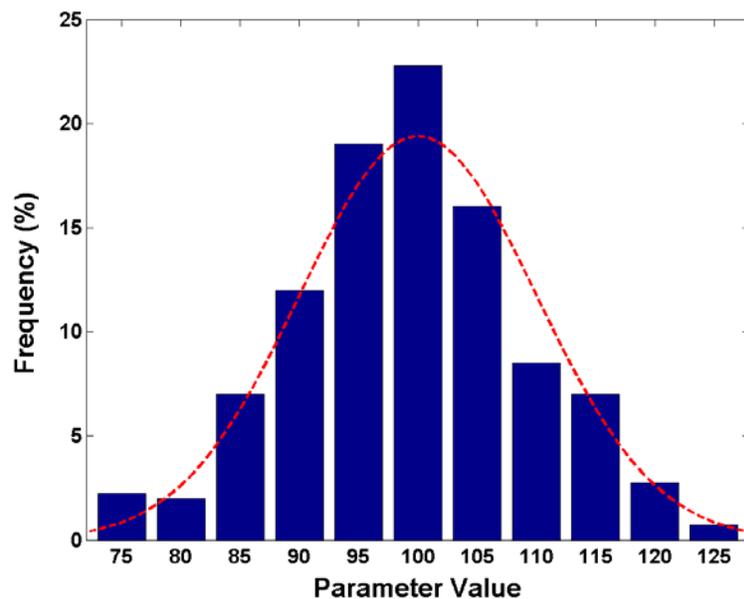
Figure 1. The frequency histogram of a set of 100 random numbers drawn from a normal pdf, shown as the dashed line.

In Monte Carlo, the degree of variability in the values of a set of random numbers reflects the width of the underlying pdf. Since the random samples mimic the pdf, its characteristics may be estimated from the samples. Thus, the mean and rms deviation of the pdf are approximately given by the mean and rms deviation of the set of random samples.

As we shall see in the next section, the pdf that describes the uncertainty in the quantity of interest may not be explicitly known. In fact, the real power of the Monte Carlo technique is that it can provide quantitative estimates of the uncertainties in a complex function of uncertain variables, so long as the function can be computed.

UNCERTAINTIES IN SIMULATION PREDICTIONS

Figure 2 shows how we think about simulations of dynamic phenomena. The simulation code requires that the situation be fully and explicitly specified. Thus, the code requires as input the properties of the materials involved over the range of conditions encountered in the experiment and the initial state of the system. Even though the diagram shows input from only one model of material behavior, more complicated situations could obviously require the input of models for several materials. Boundary conditions may be thought of as being included in the initial-state specification. The goal of the simulation code is to predict the behavior of the system at later times.

For simplicity, we consider only deterministic phenomena in which the state of the system at later times is uniquely determined by the initial state and material behavior. However, chaotic or stochastic phenomena can also be handled by the present approach.
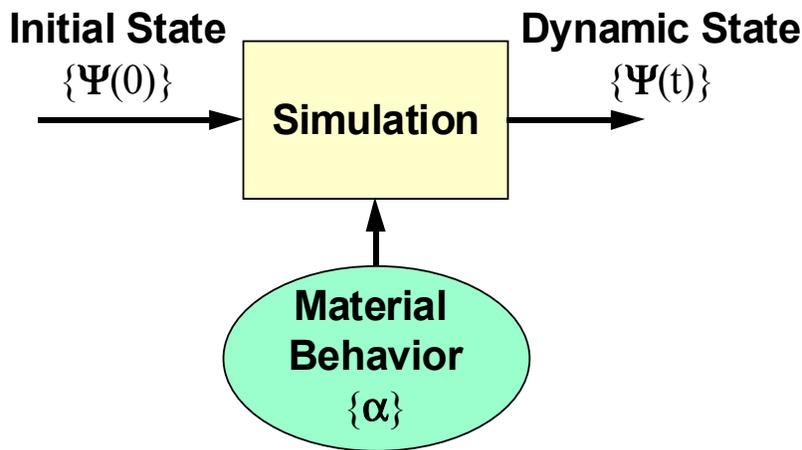
Figure 2. The purpose of a simulation code is to calculate the dynamic state of a system from a specified initial state and the defined behavior of the materials involved, represented by the parameter vector α. Uncertainties in the initial state and material models, signified here by braces, may be propagated through the simulation code to obtain uncertainties in the code predictions.

Validation is based on comparing the experimental measurements with the predictions made by a simulation code for that specific experiment. If we think about the experimental counterpart to the simulation depicted in Fig. 2, we realize that each input to the code is to some degree uncertain. Therefore, we should associate with each input variable a pdf that describes our uncertainty in its value. The properties of the material may be uncertain for a number of reasons, for example, because the exact composition and preparation of the material used in the experiment are not known. Uncertainties can also arise because of variability in the material characteristics, such as variations in grain structure that arise in specimen preparation, which may not be adequately controlled or measured. Also important in specifying an experimental set up is the degree of uncertainty in the initial state, described in terms of its geometry, initial velocities, materials, etc.

The Monte Carlo technique provides a way to estimate the uncertainties in the simulation output that arise from uncertainties in the inputs to the simulation code. To do this, the parameters that describe the initial conditions and the material behavior are drawn from the pdfs that describe our uncertainties in them. Braces are used in Fig. 2 to indicate that the quantities are uncertain. The variability in the output states of the simulation obtained for the set of random input values represents the uncertainty in simulation predictions. In this way, it is relatively easy to propagate uncertainties in inputs through a simulation code, a process that is called forward propagation of uncertainties.

EXAMPLE – THE TAYLOR IMPACT TEST

The Taylor impact test consists of impacting a cylindrical sample of material against a fixed, rigid surface, as depicted in Fig. 3. Taylor tests are often performed to investigate material behavior at high stress and high strain rates. Extremely high plastic strains develop at the crushed end of the rod, resulting in severe local deformation. The type of experimental measurements typically performed range from simply measuring the initial and final radii or lengths of the deformed cylinder, to a full specification of its profile. When the cylinder is composed of an anisotropic medium, the deformation may not be axially symmetric and these measurements would be made at several polar angles. In some cases, the strain on the surface of the deformed cylinder might be measured. These measurements can be used for validating material and fracture models developed by physicists and material scientists [3, 4].
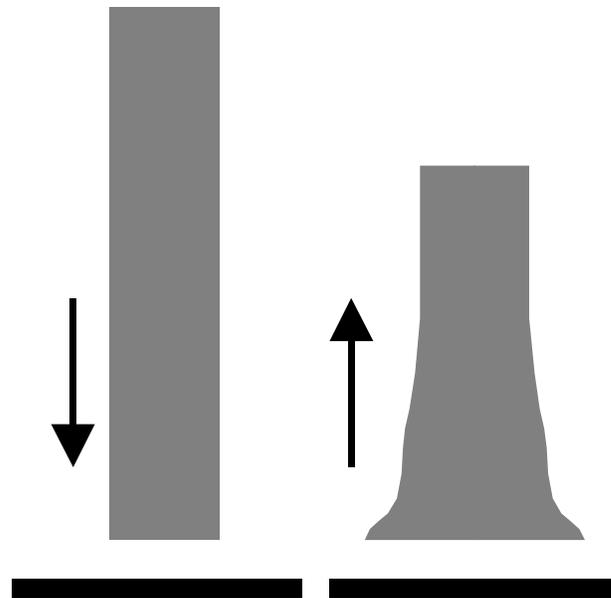
Figure 3. In the Taylor impact test, a cylinder is thrown with high velocity into a fixed, rigid plate (left) producing a significant permanent deformation in the rebounded cylinder (right).

For our numerical example, the cylinder is assumed to be made of a high-strength steel, 15 mm in diameter and 38-mm long. The impact velocity is 350 meters-per-second. Because large plastic deformations and strain rates in excess of 10^{+5} second^{-1} are expected, plasticity and strain-rate dependence must be included in the stress-strain model. We use the Johnson-Cook model for rate-dependent plasticity:

$$\sigma = (\alpha_1 + \alpha_2 \varepsilon_p^N)\left[1 + \alpha_3 \log\left(\frac{\partial \varepsilon_p}{\partial t}\right)\right] , \qquad\qquad (1)$$

where ε_p and σ denote the equivalent plastic strain and resulting stress, respectively. The parameters α_1, α_2, α_3, and N are material specific. For simplicity, we are not including anisotropies in the material behavior and have assumed that the test is conducted at ambient temperature. Furthermore, although the Johnson-Cook model can incorporate fracture, that is not included in this simulation.

Our analysis is performed with Abaqus, a general-purpose finite-element modeling and analysis package, which employs explicit time integration [5]. The nominal material constants are chosen to be α_1 = 760 MPa, α_2 = 400 MPa, α_3 = 0.011, and N = 0.26, representative of a high-strength steel. The modulus of elasticity is E = 310 GPa and the material density, ρ = 7,750 kg/m^3. Under these conditions, Abaqus predicts the shape and strain values shown in Fig. 4. The result at 50 microseconds represents the final deformed state of the cylinder.
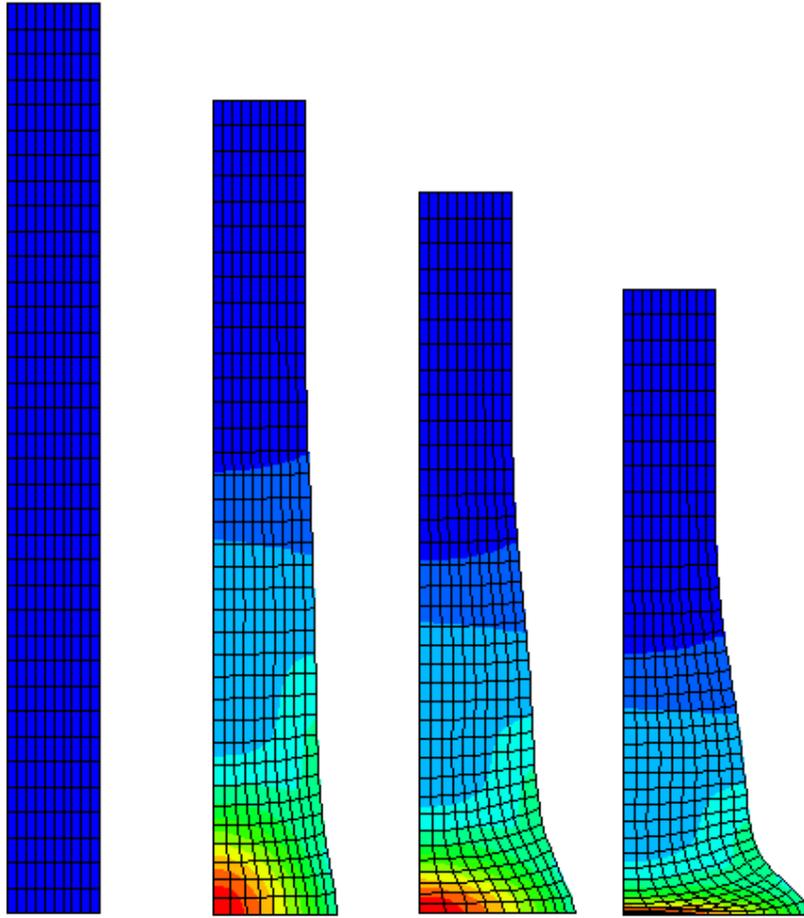
Figure 4. Simulation of the Taylor test described in the text. The figures show the half-cylinder shapes at times of 0 µs, 17 µs, 33 µs, and 50 µs after impact (from left to right). Their equivalent-strain patterns vary from 0% (dark blue) to maximum strains of 83%, 165%, and 248% (red) for the last three shapes, respectively.

UNCERTAINTIES IN THE SIMULATION OF THE TAYLOR TEST

We use the Monte Carlo technique to illustrate the forward propagation of uncertainty through the simulation code. We consider a hypothetical situation in which the only parameters that are uncertain are those in the above stress-strain model (1). We assume that the uncertainty distribution in each parameter is given by a normal distribution with the mean values set equal to the nominal values given above. We arbitrarily choose the relative standard deviations in α_1, α_2, α_3, and N to be 20%, 20%, 30%, and 40%, respectively.

As indicated above, the Monte Carlo process consists of sampling parameter values from the assumed uncertainty distributions and running the simulation code for each set of values. This procedure for uncertainty propagation can therefore be viewed as nothing more than multiple runs of a deterministic computer program. Figure 5 shows the results of employing this process to generate six final profiles. The results from only six Monte Carlo runs are

shown to avoid the confusion of plotting too many profiles on the same figure. The observed variation in shape can be interpreted as indicative of the uncertainty in profile that is produced by the assumed uncertainty in the plastic stress-strain relation. We observe that the largest variability in radius, and hence largest uncertainty, seems to occur at the contact end of the cylinder. Of course, a quantitatively accurate characterization of the uncertainties in simulation output requires many more than six realizations.
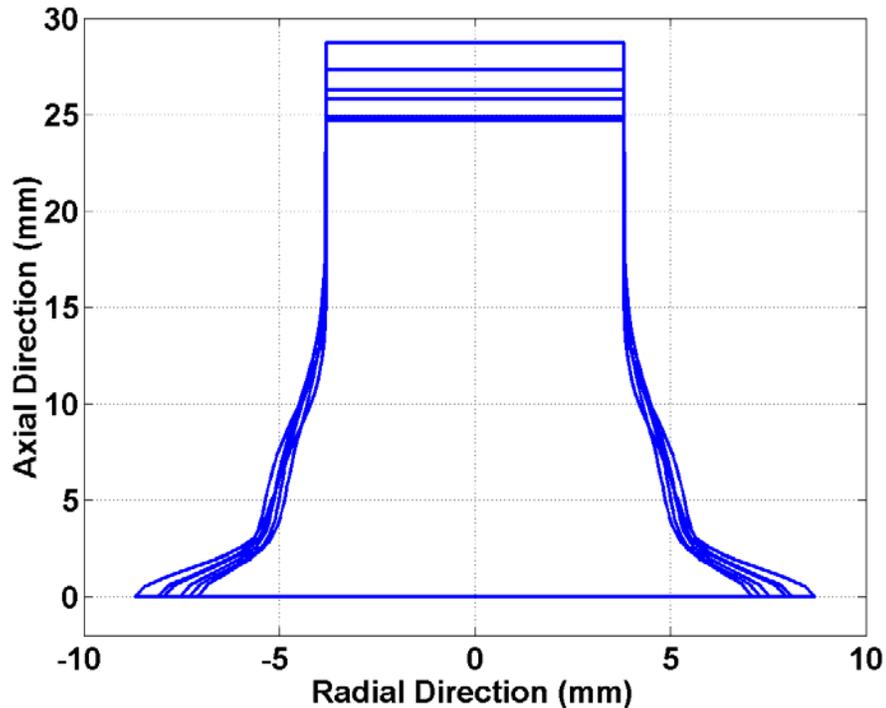
Figure 5. Six representative Taylor-test profiles for the steel cylinder predicted in a short Monte Carlo run in which there is a high degree of uncertainty in its plastic behavior. The variability in these profiles indicates the degree of their uncertainty arising from the uncertainties in the stress-strain curve for the material.

To illustrate the ability of the Monte Carlo technique to obtain quantitative uncertainty estimates, we extend the above Monte Carlo calculation to include 1,000 simulation runs. We assume that the only measurement made is the radius at the base of the deformed cylinder. Figure 6 shows the histogram of the ratio of the final radius to initial radius, R/R_o. The mean value of this distribution is 2.06. Because the distribution in Fig. 6 represents the uncertainty in R/R_o in our hypothetical situation, the rms deviation of the distribution, 0.14, is our estimate of the standard deviation in the uncertainty of our prediction for R/R_o. We observe that this histogram does not appear to be normally distributed, which should come as no surprise because the numerical simulation of this phenomenon is highly nonlinear.
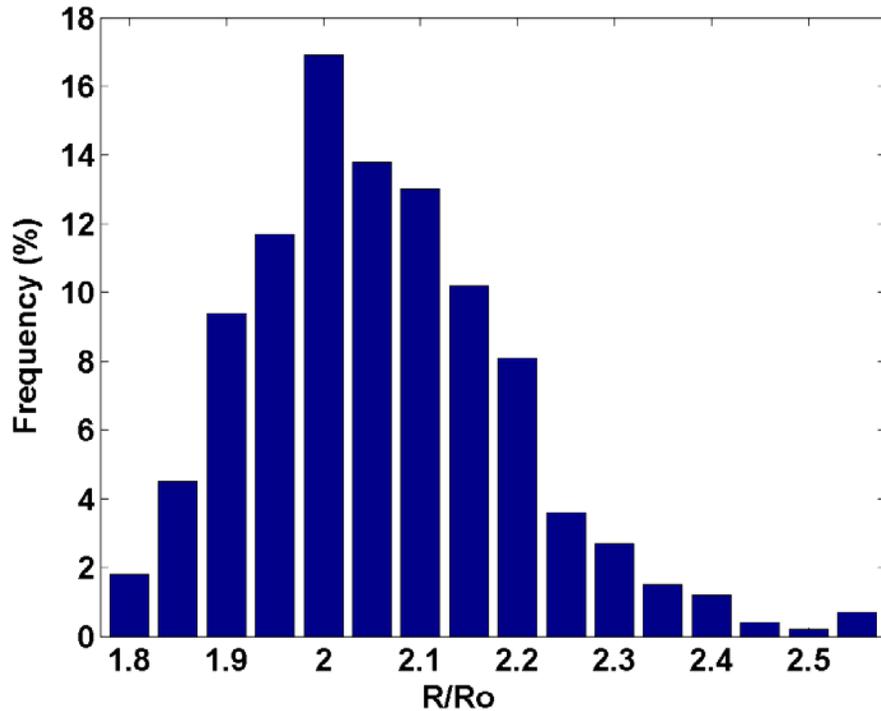
Figure 6. Histogram of the ratio of the radius at the base of the deformed cylinder to its initial value obtained from a Monte Carlo run of 1,000 simulations. This histogram represents the uncertainty in R/R_o coming from the assumed uncertainties in the plastic stress-strain behavior.

By restricting our attention to a scalar quantity, R/R_o, we have greatly simplified how we display and quote the uncertainties involved. The same 1,000 simulations have associated with them uncertainties in every aspect of their time-dependent predictions. Thus, we could have chosen to measure the deformed cylinder profiles. Then the prediction uncertainties would be similar to those presented in Fig. 5, but now for 1,000 samples. The uncertainties in the profile could be presented in terms of the mean values and rms deviation in radius as a function of distance along the axis, z. A new complication is introduced, however, because the uncertainties at one z position are statistically correlated with those at another position. The underlying statistical correlation is evidenced by the smoothness in the profiles in Fig. 5, which essentially indicates that the predicted radius cannot change abruptly from one z value to the next. It is critically important to take into account this correlation in the statistical analysis of this experiment [6]. Even more difficult would be to try to characterize the uncertainties in the full strain field, which is predicted by the simulation, as shown in Fig. 4.

Despite these complications, the Monte Carlo technique provides an easy way to visualize the nature of the uncertainties in the code predictions. One only has to view the relevant features in the code output for a sequence of Monte Carlo simulations.

These considerations, as well as practical issues, have led researchers to try to reduce the number of variables that need to be considered, both in terms of inputs and outputs from a simulation code. The use of metamodels or response surfaces represents this approach, which is often useful in dealing with complex modeling situations [7, 8]. However, as we will discuss below, it may be useful in some situations to think about uncertainties in the broader context of the full simulation.

We have ignored the potential uncertainties in the experimental set up. These are readily included in the Monte Carlo technique. For example, to include the uncertainty in the impact velocity, a fifth random variable v representing the impact velocity can be included in the input parameter vector ($\alpha_1, \alpha_2, \alpha_3, N, v$). By drawing random samples for v from its pdf, the effects of uncertainties in v can easily be included in the Monte Carlo uncertainty calculation. Other uncertainties in the experimental situation can be similarly incorporated. This approach to assessing the consequences of uncertainties in the experimental conditions provides a good basis for designing validation experiments, for example, as discussed in Ref. [9] for structural dynamics applications.

There are other sources of uncertainties regarding the material that ought to be kept in mind. For example, the material properties might not be uniform. If the density were thought to be variable throughout a specimen, that uncertainty could be included in the Monte Carlo calculation. We assumed the material properties to be isotropic, which may not be the case. For a new material for which we have no information about its anisotropic characteristics, this uncertainty ought to be included in the Monte Carlo calculations, if indeed the simulation code can take them into account.

Another source of uncertainty in the output from a simulation code is the inadequacy of the code to include certain kinds of details in the experimental situation. For example, the finite elements may not be small enough to permit the engineer to include features such as welds or fasteners in the computer representation of the structure being modeled. These details might be approximated in the model or neglected all together. Mesh size is often thought of as an issue of verification, that is, the degree to which the code solves the physics equations correctly. However, since this, and other verification issues can be an important source of uncertainties in code predictions, their effects need to be considered and accounted for in the validation process.

For this example, the Monte Carlo calculation is handled using a scripting language that allows one to run a simulation code (such as Abaqus) with controlled input parameters and to summarize the output data. The scripting language used here is Python [10]. Another possibility would be to use a commercial code such as NESSUS [11], which has been developed for reliability analysis of structural mechanics applications. Other general-purpose applications are being developed for the forward propagation of uncertainty, characterization of output probability information, and inference concerning the parameter values. One example is the DAKOTA toolkit under development at the Sandia National Laboratories [12].

COMPARISON OF SIMULATION CODE WITH EXPERIMENT

The basis of validation is the comparison of experimental measurements with a simulation code. In that comparison, it is important that the conditions put into the simulation code accurately match those of the experiment. The degree of accuracy that one can quote for the simulation is limited by the combined uncertainties in the measurements and the uncertainties in the predictions made by the code that arise from uncertainties in the experimental set up. If the prediction matches the measurements to within this combined uncertainty, we can say the simulation code has been validated to the combined uncertainty. The fact that the uncertainties in the experimental conditions are important was alluded to by Oberkampf in the previous article in this series [1]. In his Guideline 2 he cautioned that in a properly conducted validation experiment, the experimental conditions should be well controlled and documented.

Because the goal to determine the degree to which the output of the code agrees with reality, the comparison of the simulation code results to experimental measurements needs to be done in a quantitative manner. The first question to be answered is, "How well does the prediction agree with the measurements?" Then we want to state to what degree the accuracy of the code prediction has been confirmed. Both of these questions can be addressed in a quantitative way with the t test or some other appropriate statistical means for comparing two probability distributions [13]. As indicated in the preceding paragraph, the uncertainties in code output that arise from uncertainties in the experimental conditions and set up must be included in this comparison.

The t test is valid only then the pdfs are normal distributions. We saw in Fig. 6 that the uncertainty in simulation output may not be normally distributed. In that case the t test cannot be legitimately used and a suitable alternative needs to be employed. As mentioned before, another issue to be considered when there are two or more measurements is whether there are correlations among the uncertainties in the simulation predictions for those measurements. Correlations may also exist among the experimental measurements, for example, through systematic uncertainties. These would also have to be quantified and accounted for [13]. These caveats indicate that one must carefully assess the appropriateness of the t test in any specific application.

Under suitable conditions, it should be possible to use the comparison of the experimental results of a validation experiment with the output from a simulation code to gain information about the models used in the code. When this process is employed to set the values of the model parameters, it is often called calibration, which is different from how we are using the term validation [1]. In that case, it is questionable whether the same experiments may be employed to subsequently validate the code, in the sense of assessing the accuracy of its predictions. However, if a full uncertainty analysis is employed in this parameter-updating process, the process becomes one of inference [14]. In a sense the Monte Carlo technique for estimating uncertainties in simulation-code output described above is reversed and the uncertainties in the parameters are determined from the combined uncertainties in the measurements and the effects on the simulation of uncertainties in experimental set up. Since the inference process involves determining the uncertainties in model parameters, we contend that it offers the possibility of predicting the uncertainty in simulation output. Thus, inference brings us closer to reaching the ultimate goal of validation.

INFERENCE ABOUT PHYSICS MODELS FROM EXPERIMENTS

If a simulation code has been validated with a series of experiments, how much can we say about the physical variables predicted by the simulation that have not been directly measured? For example, in the Taylor test, how well can we rely on the maximum stress predicted by the simulation? Since we have not directly measured the stress, we might conclude that the simulation could not be relied on. On the other hand, if the code did predict the final shape to some degree of accuracy, it seems reasonable that the predicted stress might also be believable to some degree. It is clearly important to understand to what extent this type of inference can be made. Structural engineers need to know how much they can rely on the stresses calculated by FEM codes in order to be confident in the integrity of their designs. These concerns may be addressed using a model-based approach to understanding simulation uncertainty.

Model-based simulation codes are based on our view of how things work. Such codes are usually constructed on the basis of physical principles or laws and incorporate the behavior of materials through constitutive models. When experimental data are found to agree with code predictions, we tend to view that as a confirmation that we understand the phenomenon involved. Our hope is that the simulation code is robust to changes in the experimental conditions with respect to changes in geometry, materials, and operating conditions. This hope forms the foundation of simulation science, but is perhaps the most difficult aspect with which to come to grips and say anything about quantitatively.

Consider the hierarchy of experiments shown in Fig. 7. The level of integration is based on the number of physics models needed to describe each experiment. The nomenclature is slightly different from that used by Oberkampf [1], but the intent is similar. Ideally, one learns about the individual physics models used in a simulation code through basic experiments, which are designed to isolate and characterize each physics model. Then by doing more and more complicated experiments, one can learn more about the physics models by extending the range of physical conditions probed and also about possible interactions between different individual models, e.g., involving different materials. Experiments 1, 3, and 4 are considered to be basic experiments because they involve only the individual material models represented by the parameter vectors α, γ, and δ, respectively. Experiment 2 is partially integrated because it involves two models, represented by α and β. Experiment 5 is fully integrated in this scenario because it involves all four models [6].

Figure 7 is meant to capture the way knowledge about the physics models flows from the analysis of experiments at one level to the next. Reference [6] suggests that this knowledge may be quantitatively specified by the uncertainty distributions in the parameter vectors α, β, γ, and δ. Furthermore, Bayes law may be used in each bubble to update our knowledge of the physics models. Thus, Fig. 7 is a probabilistic network in which probabilities flow between nodes and are updated in each node on the basis of new experimental information. From this perspective, it appears feasible to systematically incorporate the information from all the experiments into a consistent set of parameters for the models involved. Because this approach is grounded in quantitative uncertainty assessment, the uncertainties in simulation predictions may be calculated through the Monte Carlo procedure described above. Additionally, we should be able to effectively address concerns about the uncertainties in inferred quantities that not been directly measured.
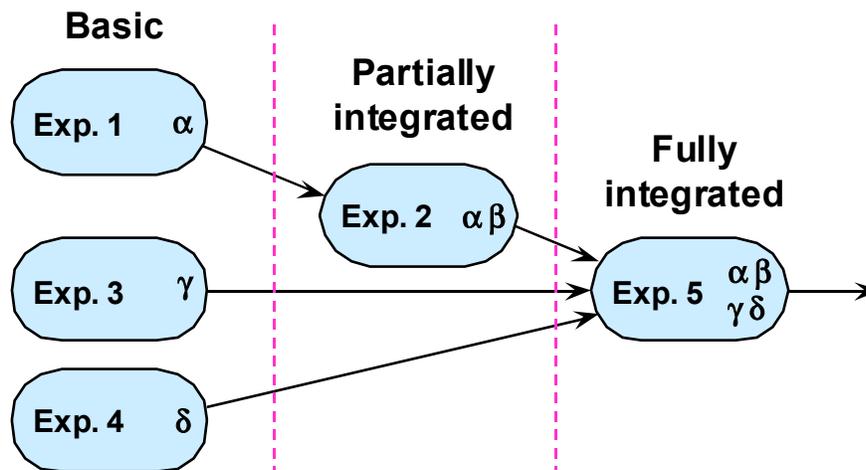
Figure 7. Diagram of the conceptual process for combining information from several experiments conducted for a hierarchy of experiments involving various levels of integration. Four underlying physics models, represented here by the parameter vectors α, β, γ, and δ, are required to account for the behavior of the fully integrated system.

CONCLUSION

We have presented the viewpoint that uncertainties in predictions made by simulation codes should be the basis for code validation. The first place to look for uncertainties in simulation code output are the uncertainties in the physics models that are incorporated in the simulation code. The Monte Carlo technique may be used to propagate uncertainties in the models into uncertainties in the code's predictions. Other sources of output uncertainties to consider include the numerical implementation of the physics models, especially the finite size of the finite elements, and aspects of the physics that are not accounted for. We point out that validation experiments may be used in an inference process to improve our understanding of the physics models, which is legitimately captured in terms of model-parameter uncertainties. This line of reasoning underscores the importance of conducting validation experiments that are thoughtfully designed to provide results that can be quantitatively compared to simulation codes.

These kinds of validation issues have been raised by the ASCI program, which was created to make up for the absence of nuclear testing through the use of computational models. However, they apply to all application areas in which people rely on simulations for decisions or design.

REFERENCES

[1] W. L. Oberkampf, "What are validation experiments?," *Exp. Tech.* **25**, pp. 35-40 (2001)

[2] D. S. Sivia, *Data Analysis – A Bayesian Tutorial* (Clarendon, Oxford, 1996)

[3] G. T. Gray III et al., "Bridging the experimental-modeling gap - the Taylor test: an integrated test used to validate constitutive models," to appear in *AMPTIAC Newsletter*

[4] P. J. Maudlin et al., "On the modeling of the Taylor cylinder impact test for orthotropic textured materials: experiments and simulations," *Int. J. Plasticity* **15**, pp. 139-166 (1999)

[5] *Abaqus™/Explicit User's Manual* (Hibbitt, Karlsson & Sorensen, Pawtucket, RI, 1998)

[6] K. M. Hanson, "A framework for assessing uncertainties in simulation predictions," *Physica D* **133**, pp.179-188 (1999)

[7] R. H. Myers and D. C. Montgomery, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments* (Wiley Inter-science, New York, 1995)

[8] T. A. Butler et al., "Model validation for a complex jointed structure," *Proc. 19th SEM Int. Conf. on Modal Analysis*, pp. 1318-1324 (2001)

[9] F. M. Hemez et al., "Design of computer experiments for improving an impact test simulation," *Proc. 19th SEM Int. Conf. on Modal Analysis*, pp. 977-985 (2001)

[10] M. Lutz, *Programming Python* (O'Reilly & Associates, 1996)

[11] *NESSUS User's Manual* (Southwest Research Institute, San Antonio, TX, 1996)

[12] S. F. Wotjkiewicz et al., "A toolkit for uncertainty quantification in large computational engineering models," *Proc. 42nd Joint Conf. on Structures, Structural Dynamics, and Materials* (2001)

[13] A. C. Rencher, *Methods of Multivariate Analysis* (Wiley, New York, 1995)

[14] K. M. Hanson et al., "Inversion based on computational simulations," in *Maximum Entropy and Bayesian Methods*, G. J. Erickson et al., eds., pp. 121-135 (Kluwer Academic, Dordrecht, 1998)

This and other articles available at http://www.lanl.gov/home/kmh/publications/