Ross, Timothy; Kreinovich, Vladik and Joslyn, Cliff: (2001) "Assessing the Predictive Accuracy of Complex Simulation Models", in: Proc. 2001 Joint International Fuzzy Systems Association World Congress and North American Fuzzy Information Processing Society Int. Conf., pp. 2008-2012, https://doi.org/10.1109/NAFIPS.2001.944376

ASSESSING THE PREDICTIVE ACCURACY OF COMPLEX SIMULATION MODELS

Timothy Ross
Center for Autonomous Control Engineering, University of New Mexico, Albuquerque, NM 87131
ross@unm.edu

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968
vladik@cs.utep.edu

Cliff Joslyn
D-Division
Los Alamos National Labs
Los Alamos, NM 87545
joslyn@lanl.gov

Abstract

Predictive accuracy is the sum of two kinds of uncertainty–natural variability and modeling uncertainty. This paper addresses the quantification of predictive accuracy of complex simulation models from two perspectives. First, it recognizes that there is a difference between variability and modeling uncertainty; the former can not be reduced with more test information, while the latter can. We suggest that variability is a natural form of uncertainty that can be quantified with probability theory, but that modeling uncertainty is a form that is better addressed by a theoretical foundation that is not based on random variables, but rather random intervals. We suggest possibility theory as the formalism to address modeling uncertainty. The paper discusses the two different methods, and illustrates the power of their integration to address predictive accuracy with a recent case study involving the crushing load of axially loaded metallic spheres.

1. Introduction

Predictive accuracy can be defined as the degree to which a model of a complex system is able to foretell the state of that system under conditions for which the model has *not* been validated experimentally. In general, the lack of predictive accuracy that results from the random character of a variable, such as in games of chance or in the natural variability of things due to manufacturing processes, is often termed variability. Variability can not be reduced, but rather only quantified. The size of grains of sand or the specific shapes of a maple leaf are things that exhibit natural variability. If we want to predict either of these quantities we can only do so in an average sense for the population of grains or leaves. There have been many numerous, accurate characterizations of this form of uncertainty.

Another form of uncertainty is that due to a lack of specific information, and this has been generally called uncertain–to distinguish it from variability. There are various forms of uncertainty; uncertainty can arise from ignorance, from scare data, from misleading data, from unknown biases, or from our inability to understand complex systems. Collectively, we shall use the term *modeling uncertainty* to describe these various forms of non-random forms of imprecision, ambiguity, vagueness or *unknowingness*. Uncertainty and variability generally result in a loss of predictive accuracy. The question is how can the predictive accuracy of a model be quantified?

An interesting question arises: how can one contend with gaps in knowledge that cannot be represented probabilistically or statistically? Examples of the latter might include the degree of confidence placed in certain modeling assumptions before they can be validated experimentally, or the degree of confidence placed in extrapolating laboratory experiments to field conditions. Reasonably simplified assumptions are often made that render the problem tractable for engineering computations or simulation modeling. Assumptions are made with respect to the analyst's preferences, the available information, or other factors, which are generally approximate and embedded in the modeling uncertainty. In addition, engineering judgment is commonly involved in checking and modifying the predictions to make sure that the system behaves satisfactorily in some predetermined sense. Moreover, the engineering models are posed such that they adhere to necessary functions and to impose constraints such as boundary conditions or total computation time.

In this work we have proposed to use possibility theory as the mathematical form for characterizing modeling uncertainty. We do so for two very good reasons: first, this theory contains probability theory as a special case and second, the implementation of the theory is computationally simple and easy to understand. In the simplest sense a possibility distribution arises from the random selection of intervals, as opposed to a probability distribution which arises from the random selection and ranking of point-valued quantities.

To illustrate the utility of the possibility theory approach we discuss the prediction and modeling of the buckling load of metallic, spherical pressure vessels, i.e., the crushing capacity of axially loaded manufactured marine floats. In this approach we use a finite element code to predict the buckling load of the spheres in a numerical simulation environment, and then compare both a probabilistic and possibilistic assessment of the prediction to test results gleaned from testing numerous quantities of these floats.

2. Case Study: Traditional Approach

As mentioned in the introduction, we discuss the prediction and modeling of the buckling load of metallic, spherical pressure vessels, i.e., the crushing capacity of axially loaded manufactured marine floats. In the traditional approach we use a finite element code to predict the buckling load of the spheres, along with a probabilistic simulation tool that is used to assess the degree of uncertainty in the buckling load as a function of the uncertainty in key parameters of the finite element model.

One hundred marine floats were purchased from a commercial vendor [1]. Though the vessels were intended to be spherical by the manufacturer, they possess variations in their geometry and material properties. The classic modeling approach would have been to generate a single nominal model of a sphere and assert that however this nominal model behaved in analysis, so would all the individual floats. Since the floats are manufactured units, subjected to specified levels of tolerance and quality control, they all deviate to some extent from the idealized model used in the analysis, and thus the uncertainty surrounding the analysis results needs to be modeled and quantified.

2.1 Probabilistic Simulation

In the numerical simulation of the buckling sphere problem the DYNA3D code, a nonlinear, three-dimensional dynamic finite element continuum code, to conduct the stochastic analysis. There are two primary approaches to a probabilistic simulation: a direct Monte Carlo simulation, or simpler sampling methods (of which there are many...mean value analysis, stratified sampling, response surface, Latin hypercube, and advanced mean value analysis to name a few). In a Monte Carlo simulation the DYNA3D code is run for a single combination of the input parameters, which have been randomly sampled from pdfs for each of the input variables. For the sphere, the input variables would include the sphere radius, thickness, modulus of elasticity, yield strength, and other material properties. This process is done 10^3 to 10^6 runs of the DYNA3D code and the output values (in this case, the failure load of the sphere when it buckles) are stored and plotted in histogram form to develop a pdf of the failure variable. Of course, if a single run of the DYNA3D code can take say one hour, then this method of producing the output, while very effective, is simply not practical, unless vast computational resources are available at low cost.

A more effective solution is to use one of the other sampling methods. For example, in a mean-value sampling the input distributions are sampled–not at random–but at pre-specified values. In this study, the values were the mean of each input parameter and 10% of the standard deviation above the mean, i.e., at μ_i and $\mu_i + 0.1\sigma_i$. If the standard deviation is assumed to be 10% of the mean, this sampling occurs at μ and 1.01μ. Since the deviation from the mean is so close, we can justify a linear mean-value analysis.

In our study, we looked at 6 parameters in our simulation model, such as Young's modulus, radius of the spherical float, and thickness of the spherical wall. All these parameters were modeled as random variables using lognormal distributions. In the mean-value analysis there were 7 simulation runs: one with all parameters being held to their mean values, one each where 5 of the 6 parameters were at their mean value and a sixth was at μ and 1.01μ. The output of this simulation is shown in Figure 1, which is a plot of the cumulative distribution function (CDF) of the maximum force seen by the sphere at buckling (i.e. the output variable). In another simulation the 6 input parameters were sampled at the mean (μ) and at $\mu \pm 1.01\mu$, for a total of 13 simulation runs. The result of these simulations is shown in Figure 1 as the light-curve CDF. The results in Figure 1 show that the variance in the output decreases as more points near the mean values of the parameters are sampled.

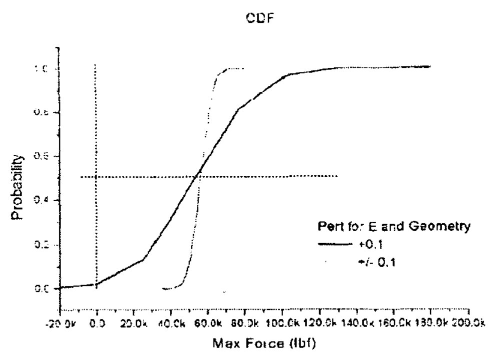

Figure 1. Cumulative distribution functions (CDF) for traditional solution to the variability in the sphere buckling load [2].

2.2 Some problems in Predictive Accuracy

The purpose of simulation is obviously to avoid the high cost of physical testing and to provide an environment where many exploratory iterations and tradeoff studies can be conducted. However, in order to assess the predictive accuracy of a simulation there must be a way to assess the fidelity of its output. Experience in comparing simulations with physical tests represents the essential knowledge for an analyst in being able to assess this fidelity. But, there are problems associated with any simulation, that even experts have trouble foreseeing. The analyst has to make many assumptions to conduct the simulations according to intuition and desired economies. For example, in our case with the marine floats the simulation assumes a loading speed of 5 m/sec. This is a tradeoff resulting from the cost of simulation on the one hand and the computational noise in the resulting data on the other. The actual load speed in the testing of the spheres is 5 cm/sec. Unfortunately, if a simulation were to use this loading speed, the resulting computation would consume many months of computer time. Even a load speed of 1 m/sec takes one week of computer time. Alternatively, loading speeds greater than 5 m/sec produce simulation results whose noise obliterates the output of interest, hence 5 m/sec is seen as a compromise by the analyst. But what is the degradation in simulation accuracy when this tradeoff is made? Such an assessment can't be made probabilistically.

3. Case Study: Possibility Distribution Approach

Previous efforts on the derivation of possibility distributions are few, especially in the derivation of empirical possibility distributions [3]. Possibility theory has long been confused with fuzzy set theory, in that possibility distributions were considered to be membership functions [4]. Possibility distributions were also viewed as resulting from consonant crisp sets in fuzzy measure theory. This perception arises from Dempster and Shafer's evidence theory when the evidence focuses on consonant support functions [5].

Previous methods for deriving possibility distributions [3] do not assist modeling empirical interval data. In the case of deriving empirical possibility distributions it is more natural to consider random sets (or intervals) and build a distribution based on the original set of interval data. This interpretation is more realistic as experimental observations are usually recorded as ranges of numbers. Joslyn [6] has developed a method that calculates a possibility histogram from random sets. Joslyn's method, however, derives possibility measures based on only *consistent* random sets rather than *consonant* sets. It is important to note that the property of *consonance* (i.e., nesting of sets) is essential in not only calculating possibility measures but also for combining two or more possibility distributions using interval arithmetic. Donald [7] developed a new method to take random sets that are *consistent* and derive from this a *consonant* set of empirical intervals.

3.1 Possibility Distributions using the New Method

Many of the assumptions made in the traditional approach can be addressed in a less computationally expensive, and perhaps more epistemologically appropriate, environment using possibility theory. In this approach, modeling assumptions like the EOS selected, the size of the finite element grid, the type and extent of boundary conditions, and the loading speed can be implemented into the theory as upper and lower bound judgments on the typical assumptions. The upper and lower bound levels are represented simply as intervals in the theory. For purposes of illustration of our method, we considered the following ranges for input variables which are not normally modeled as random pdfs in conventional reliability analysis, but which nonetheless are very important to the prediction of the crushing load of the spherical vessels:

1. Mesh Density: 7,500 elements – 15,000 elements
2. Static coefficient of friction – 0.10 – 0.35
3. Material Model - #24, piecewise linear strain hardening - #18 power law isotropic elastic plastic
4. Shell Thickness – varies from 0.57" to 0.43" at the pole and equator of the sphere
5. Loading speed of Platens – 10m/s – 50 m/s

If we consider the peak crushing load data as non-consonant intervals, expressing the imprecision in the data from the outputs of the finite element code for the various choices for variables listed above, we can determine the possibility distribution that would quantify this *modeling* uncertainty. Information such as this are common in the real world wherein they are presented as a range of *possible* numbers given within a certain error value. Through the new method developed by Donald [7] we compute a possibility distribution from non-consonant information, for our sphere buckling problem:

Figure 2. Possibility distribution for output intervals relating to the 5 modeling variables

3.2 Comments about the Possibility Approach

What is most interesting in Figure 2 is that the region represented by $\pi(A)=1$, under which a probability assessment of the same variables should exist, is approximately the same region bounded by the probabilistic results shown in the Figure 1, within ± 2-sigma bounds, even though the same variables were not assessed in both methods (i.e., both methods assessed variations in thickness, but only the probabilistic assessment looked at variations in Young's modulus).

Possibility distributions, such as the specific one illustrated in Figure 2, or the generic one illustrated in Figure 3, relate to probability distributions in the sense that a region of unit possibility spans the space of a non-zero probability distribution (e.g. a probability density function or pdf), while outside of that interval some possibility may still exist in the face of conflicting (or dissonant) evidence. As more data are acquired, the dissonance (represented by the sloping regions of the possibility distribution) diminishes and the side boundaries of the possibility distribution become steeper. One possible use of possibility distributions might be, for example, to test whether the predictive accuracy of a model based on generic uncertainty data is valid for a model of a newly designed component or system.

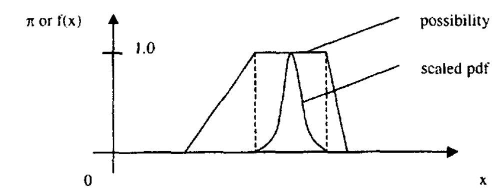

Figure 3. Relationship between probability and possibility distributions

4. Conclusions

In this paper we propose the use of the possibility distribution approach to affect three main objectives in the assessment of the predictive accuracy of simulation codes:

- to be used initially for all assessments to determine the regions into which more focus should be placed by subsequent probability computations
- to be used to quantify all variables in the simulation for which little or no data exists, or for modeling assumptions for which a probabilistic evaluation simply is not warranted since the underlying structure of the variable is non-random
- to assess those regions of the output where dissonance, or disagreement exists in previous data or existing analytic judgments or knowledge.

Using these three objectives we believe that the assessment of predictive accuracy can be streamlined in terms of cost savings and the efficient use of valuable historical data. It also allows for the judgments and knowledge of the analyst's making the predictions more flexibility in embedding all their knowledge–not just the numeric information–into their analyses. The use of historical data to guide our analytic judgments has been used primarily in establishing a sort of classification of the appropriate methods and models to apply to any physical system.

As a final note, the possibility distribution can ultimately be used as a guide in determining how well an analyst understands the extent of the relationship between modeling uncertainty and variability. We surmise that a probability density function (pdf) reflects the amount of variability in a simulation. In contrast, the possibility distribution reflects the amount of *predictive uncertainty* (which, again, is the sum of variability and modeling uncertainty) in the simulation. In Figure 3 we see two different distributions that can be used to assess the differences between modeling uncertainty and variability. On the one hand, the predictive uncertainty and the variability could be almost the same if the aprons on the possibility distribution function are have a near vertical slope (the pdf is a large part of the possibility distribution). On the other hand, the predictive uncertainty and the variability could be vastly disparate if the aprons on the possibility distribution function are have very low slopes (the pdf is a small part of the possibility distribution). Hence, in a sort of graphical way, the difference in the regions mapped by the variability (pdf) and the possibility distribution is a quantitative assessment of the modeling uncertainty in a problem. In this sense, the boundary regions of the possibility distribution can be used as a guide about where, specifically, we need more information in any planned future testing to reduce total uncertainty.

5. References

[1] Treml, C., R. Dolin, and N. Shah (2000),"Using Bayesian Belief Networks and an As-Built/As-Is Assessment for Quantifying and Reducing Uncertainty in Prediction of Buckling Characteristics of Spherical Shells", Proceedings 8th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability, Notre Dame, IN, 24-26 July.

[2] Pepin, J. (2000), private communication on finite element results.

[3] Ross, T. (2000) "Possibility Intervals in the Characterization of Dynamic Model Uncertainty", proceedings of the 19th Int. Conference of NAFIPS, PeachFuzz '2000, Atlanta, GA, July 2000, IEEE Publ., T. Whalen (ed), ISBN 0-7803-6274-8, pp. 353-357.

[4] Zadeh, L. (1978), Fuzzy Sets as a Basis for a Theory of Possibility, Fuzzy Sets and Systems, 1, 3-28.

[5] Shafer, G. (1976), A Mathematical Theory of Evidence, Princeton University Press.

[6] Joslyn, C. (1992), Possibility Measurement and Set Statistics, proceedings of the Conference of the North American Fuzzy Information Processing Society, December, Puerto Vallarta.

[7] Donald, S. (2001), An alternative method for the empirical development of possibility distributions, PhD Dissertation, Department of Civil Engineering, University of New Mexico, Albuquerque, NM.