Reliability, Uncertainty, Estimates, Validation and Verification

Babuska, I.I.; Nobile, F.; Oden, J.T.; and Tempone, R.

Proceedings of the workshop on Elements of Predictability, held at John Hopkins University, in Baltimore, MD, November 13-14, 2003
RELIABILITY, UNCERTAINTY, ESTIMATES
VALIDATION AND VERIFICATION

Transcription of I. Babuška’s presentation at the workshop on the Elements of
Predictability, J. Hopkins Univ, Baltimore, MD, Nov. 13-14, 2003

I. I. Babuška, F. Nobile, J.T. Oden, R. Tempone *

Abstract
This lecture addresses the major features of the reliability of a mathemati-
cal model and related computer simulation for specific data of interest. It explains the basic notions of Validation and Verification, elaborat-
ing on a concrete example - the deformation of a loaded aluminum alloy
2024-T351 body. Further, it addresses the treatment of problems with
uncertain and possibly random data. Some comments on hierarchical
modeling are also included.

1 Introduction
Computational analysis, in general, and numerical treatment of PDE, in partic-
ular, are intended to provide reliable predictions of particular events, which are
used as the basis for crucial decisions. Reliability is related to Verification and
Validation. This subject is gaining increasing attention: the work by Roache,
Trucano, Oberkampf & al. and recent drafts written by Senseny, Thacker, An-
derson, represent the modern thinking in this area. In this talk we will address
these topics and give concrete examples and related mathematical details. For
definitions and comments we refer to [9].

- **Verification**: The process of determining if a computational model ob-
tained by discretizing a mathematical model of a physical event represents
the mathematical model with sufficient accuracy.

- **Validation**: The process of determining if a mathematical model of a phys-
ical event represents the actual physical event with sufficient accuracy.

Verification consist of a code verification and verification of the numerical
results. Verification of numerics is equivalent to a-posteriori error estimation.

*Ices, University of Texas at Austin
By error we mean the difference between the exact solution of the mathematical problem - the event of interest - and its approximation. Mathematical theory and practice of a-posteriori error estimates for PDE's was introduced by Babuska and Rheinboldt in 1979, together with the idea of adaptive approaches for mesh refinement. Today a-posteriori estimates and adaptive approaches are subject of hundreds of papers and monographs and are extensively used in engineering practice. For a survey on a-posteriori error estimation see the presentation of J. Stewart at this meeting and [1, 4]. For some recent results see for example F. Nobile [12, 11].

Validation is a much more complex problem. It relates theory and reality. This relation was at the center of philosophy for more than 2000 years. Aristotle, Kant, Hume, etc. and the philosophers of science of the 20th century, K. Popper, T. Kuhn addressed this relation.

The essence of the validation problem is the following: We wish to predict a complex event by solving a mathematical problem. We select a set of specific validation problems which are simpler than the specific problem leading to the prediction, and collect experimental data from them. We use then the mathematical theory under consideration, solve numerically these problems and compare the results with the experimental data. Based on a rejection criterion we can reject the theory. (We have to be sure, by verification, that the numerical results are close to the exact solution). If we do not reject the theory against what we consider to be a sufficient set of validation problems then we accept the model for the prediction problem.

- We emphasize that reliable experimental data are hard to obtain and hence various uncertainties are unavoidable.
- The model could be deterministic or stochastic with completely known data or data with uncertainties.
- The estimation of the model error is much more complex than that of the numerical solution (verification).

We will now address a specific example. Consider a body made of aluminum alloy 2024-T351 manufactured by Alcola, with the shape of a plate and thickness 3/8". The body is clamped at \( \Gamma_1 \) and subject to a time dependent load function \( g(t) \) at \( \Gamma_2 \), (see Figure 1). The physical event we are interested in is the displacement at a given point \( A \). The load \( g(t) \) is very slowly changing and resulting displacement \( u \) is small. We are interested in the prediction of \( u(A, t) \) using linear elasticity theory. Essential here is that the event of interest is clearly defined. The validation and verification are directly related to the event of interest.
2 The mathematical problem, specific mathematical problem, prediction and validation problems

A mathematical Problem, on which a given prediction is based, has two parts:

a) Functional Part: The structure, for example the type of differential equation including the mathematical definition of the event of interest (ex: Lamé elasticity equations). We denote this structure by $S$.

b) The set of admissible data $J$ for which solutions (events) in an output set $G$ exist.

Assuming that the functional part $S$, i.e. the structure, is fixed then the mathematical problem is the transformation of $J$ into $G$ i.e. $G = AJ$. Here $A$ is the operator defined by the structure of the problem and the goal of the analysis.

The specific problem is the problem to obtain the output $g \in G$ for a specific input $\xi \in J$. The prediction is then the solution of a specific problem.

The validation process is at the basis of the decision whether $g$ is a reliable prediction. The only tools for this decision are: the information obtained from the set of validation problems, the comparison of their solution with the experiments and the rejection criterion. The set of validation problems has to be designed for this purpose.

- A preliminary computational analysis of the prediction problem may help in the selection of the set of validation problems.

- A prediction cannot be validated. If there is a possibility to compare the prediction with an experiment then we speak about post audit and the "prediction" becomes just another validation problem.
- The estimation of the reliability of the prediction is based on an induction process, starting from the validation problems. Hence the selection of the set of validation problems and the rejection criterion is crucial.

- Very often the needed data (information) are unreliable or not available. Typically expert’s opinions are then used as the basis for the application of the fuzzy set theory.

Figures 2, 3, 4 show a schematic representation of a mathematical problem, specific mathematical problem and the validation problems.

As a concrete example we will address the deformation of a body from the aluminum alloy 2024-T351 mentioned earlier, modeled by linear elasticity. The material properties are given in: Materials and elements for aerospace vehicle structures: Mil-HDBK-5J, 31 Jan 2003 (see Table 1). Here, $E_t$ and $E_c$ are

<table>
<thead>
<tr>
<th>Thickness</th>
<th>$E_t$ (ksi)</th>
<th>$E_c$ (ksi)</th>
<th>$\nu$</th>
<th>$G$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t \in (0.01^\circ - 0.249^\circ)$</td>
<td>10.5</td>
<td>10.7</td>
<td>0.33</td>
<td>4.0</td>
</tr>
<tr>
<td>$t \geq 0.250^\circ$</td>
<td>10.7</td>
<td>10.9</td>
<td>0.33</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Table 1: Modulus of Elasticity and Poisson ratio reported for tension and compression

moduli of elasticity in tension and compression, respectively, $\nu$ is the Poisson ratio and $G$ is the shear modulus. The material is assumed to be isotropic.

It is not said if these values are mean values and what are the corresponding standard deviations. Moreover, information about heterogeneity, covariances, correlation between $E$ and $\nu$, and the differences in material due to rolling are not given. Experts opinion is that the standard deviation in $E$ and $\nu$ are 3–6% of the mean and that these two are uncorrelated (Hence the Lamé constants are correlated). Need of validation of material properties is one important aspect for the reliability of the prediction. We will use two validation problems to this purpose.

2.1 Tension dog bone validation problem

In Figure 5 we show the scheme of the dog bone experiment.

The bone is clamped on both ends and cyclically loaded and unloaded in the axial direction (see Figure 6). The strains $\varepsilon_i(t)$ $i = 1, 2$, are measured on both sides of the bone. Precisely, the measured quantity is the averaged strain over the gage. The value

$$\varepsilon(t) = \frac{1}{2}(\varepsilon_1(t) + \varepsilon_2(t))$$

is compared with numerical computations. The average of the strains on both sides is used to compensate possible bending effects due to the grasps at the ends. The transversal and axial strains are measured. Quantification of the reproducibility of the experiment is essential. The definition of the reproducibility
Figure 2: Mathematical Problem. Here $J$ is the set of admissible input data, $S$ is the structure of the problem and $G$ is the admissible output. In our concrete example $J$ consists of a set of possible domains $\Omega$, boundary conditions and the coefficients (material properties) of the partial differential equation (Lamé equations).

Figure 3: Specific Mathematical problem. Here $\xi \in J$ and $g \in G$ are singletons. The prediction is a specific mathematical problem.

Figure 4: Validation Problems. Validation problems can be and usually are very different from the prediction problem. The criterion can reject some input data, as for example the coefficients values in the PDE, or the structure itself (linear Lamé equations).
Figure 5: Dog bone validation problem

Figure 6: Loading and unloading cycles
has to be related to the rejection criterion used below. We define
\[
\theta^R = \frac{\max_{j,k} \| \varepsilon^j(t) - \varepsilon^k(t) \|_{L^\infty(0,T)} }{\max_{j,k} \frac{1}{2} \| \varepsilon^j(t) + \varepsilon^k(t) \|_{L^\infty(0,T)}}.
\]  

Here, \( \varepsilon^j(t) \), respectively \( \varepsilon^k(t) \), are the measured strains of the sample bones \( j \) and \( k \), respectively. The maximum is taken over all available combinations of \( (j,k) \). In the denominator of (1), we can use also \( \min \frac{1}{2} \| \varepsilon^j(t) + \varepsilon^k(t) \|_{L^\infty(0,T)} \). Yet, this change will practically not affect the value \( \theta^R \), because \( \theta^R \) is small.

From our measurement we have \( \theta^R = 0.045 \).

The value \( \varepsilon(t) \) in (1) involves also errors in the measurements, caused by the apparatus. In the experiments we have to consider the analogs of validation and verification. In this case, validation is related to the confidence that the experiment addresses the event of interest. Verification relates, instead, to the confidence that the measurements are not significantly influenced by the errors of the measurements apparatus.

Validation based on the dog bone problem

A) Validation of published data: We have \( E = 10.7 \) and \( \nu = 0.33 \). Denote by \( \varepsilon(t) \) the computed strain using these data, assuming isotropic homogeneous material, and by \( \varepsilon^j(t) \) the measured strain in the \( j \)-th experiment. Define
\[
\theta(E, \nu) = \frac{\max_j \| \varepsilon(t) - \varepsilon^j(t) \|_{L^\infty(0,T)} }{\max_j \| \varepsilon^j(t) \|_{L^\infty(0,T)}}
\]
where the max is taken over all used samples. We have \( \theta = 0.1009 \).

The rejection could be based on

a) Comparison of \( \theta^R \) and \( \theta \). We reject the model whenever \( \theta > a \theta^R \) where \( a = 2 \) for example.

b) The value of \( \theta \). We reject the model when \( \theta > \beta \) with \( \beta = 0.06 \) for example.

This means that we reject the published data using either criterion a) or b).

We emphasize that it is essential to verify the computed data - i.e. to have sufficient accuracy. *Verification comes before validation.*

B) Validation of the guessed data: We use \( E = 9.85, \nu = 0.3256 \) and get \( \theta = 0.03212 \).

Hence we have no reason to reject \( E, \nu \) based on the dog bone validation problem using either criterion a) or b).

Why did we use the rejection criterion \( \theta \)? Because the solution of the Lamé equation is stable with respect to perturbations in the coefficients when measured in \( L^\infty \) norm.

What can we conclude from the Dog Bone problem and what can we ask:
1. What $E$ and $\nu$ can be used to avoid rejection of the homogeneous isotropic model?

Consider the sets $E \in \mathcal{I}_E = (9.68, 10.12)$ and $\nu \in \mathcal{I}_\nu = (0.318, 0.338)$ and define $\theta_H = \sup \theta(E, \nu)$ where the sup is taken over all constants $(E, \nu) \in \mathcal{I}_E \times \mathcal{I}_\nu$. We get $\theta_H = 0.0478$ and hence any $(E, \nu) \in \mathcal{I}_E \times \mathcal{I}_\nu$ cannot be rejected.

2. Will be some of $(E, \nu) \in \mathcal{I}_E \times \mathcal{I}_\nu$ rejected when the material is non-homogeneous, i.e. $E$, $\nu$ are not constant?

To this end, define $\theta_{NH} = \sup \theta(E(x), \nu(x))$ where the sup is taken over all measurable $(E(x), \nu(x)) \in \mathcal{I}_E \times \mathcal{I}_\nu$. We get $\theta_{NH} = 0.0941$. Hence, considering rejection criterion a) with $\alpha = 2$, we have no reason to reject the model of non homogeneous isotropic material. An anisotropic model can be analyzed in a very similar way.

- The determination of $\theta_H$ and $\theta_{NH}$ requires the solution of an optimization problem. We have solved it numerically and verified the corresponding accuracy.

- Based on the dog bone validation test we reject $E$ and $\nu$ given in the mentioned report. Any $E \in \mathcal{I}_E$ and $\nu \in \mathcal{I}_\nu$, constant or only measurable cannot be rejected.

2.2 Non restricted tension validation problem

We consider the problem shown in Figure 7. The data of interest (events) are $Q_1 = \frac{1}{|A|} \int_A u_y \, dx \, dz$ (in plane deformation) and $Q_2 = \frac{1}{|A|} \int_A u_x \, dx \, dz$ (out of plane deformation). The main difference between the dog bone problem and this one is that the right end is free to move laterally in the latter case. As we will see below, there is a large difference in the two events. Functional $Q_1$ is not sensitive to possible inhomogeneities while $Q_2$ is. To show this, assume that $E = 9.87 + \delta_E$, $|\delta_E| \leq \Delta_E = 0.22$, $\nu = 0.3308 + \delta_\nu$, $|\delta_\nu| \leq \Delta_\nu = 0.012$ and compute the range of $Q_1$ and $Q_2$ when

a) $E$ and $\nu$ are constant in $\Omega$.

b) $E$ and $\nu$ are arbitrary measurable functions in the given range.

Denoting by $\Delta_E^\infty = \Delta_E/9.87$ and $\Delta_\nu^\infty = \Delta_\nu/0.3308$, we have

- The case a): $Q_1 = Q_1^0(1 \pm \Delta_E^\infty \pm 0.0035\Delta_\nu^\infty + \text{higher order terms})$, $Q_2 = Q_2^0$. Here $Q_1^0 = 1.76 \times 10^{-4}$, $Q_2^0 = 2.43 \times 10^{-8}$. The higher order term was estimated and is negligible.

- The case b): $Q_1 = Q_1^0(1 \pm 1.082\Delta_E^\infty \pm 0.0055\Delta_\nu^\infty + \text{higher order terms})$ and $Q_2 = Q_2^0 \pm 1.5778 \times 10^{-3} \Delta_E^\infty \pm 1.506 \times 10^{-5} \Delta_\nu^\infty + \text{higher order terms}$.
Hence, perturbations of the order of 2% in $E$ and $\nu$ lead to a change in $Q_2$ which is about 50% of the value of $Q_1$. We see that the assumption on inhomogeneities and possible discontinuities in $E$ and $\nu$ leads to large changes in $Q_2$. In Figure 8 we see the functions $E(x)$ and $\nu(x)$ leading to the largest range in $Q_1$ and $Q_2$. Experimental values will possibly allow to reject the model with non homogeneous $E$ and $\nu$ and accept the homogeneous one. Unfortunately we do not have access to the corresponding experimental data. We have seen that the unrestricted tension problem gives more information than the dog bone problem.

The approach we have addressed so far, corresponding to the case where the input data are only known to range in a given set, is called worst scenario approach.

In the same spirit, validation problems can be designed for other formulations (structures) as, for example, for a stochastic formulation.

Figure 7: Cantilever beam diagram

We mentioned only two validation problems which are important but possibly not sufficient for the reliability of the prediction. For example, of importance could be also the validation of the boundary conditions, determination of the range of admissible load vectors, influence of nonlinear materials behavior in the area where the linear theory leads to the large stresses, etc.

2.3 Conclusions

1. Only the validation problems can be validated and not the predictions (otherwise it is a "post-audit")

2. Validation problems should be selected in relation to the prediction problem. This is not a trivial issue; mathematics and numerics have to be
used.

3. A model is validated for the prediction if it is not rejected on the basis of the validation problems. Rejection criterion has to be related to the prediction problem.

4. The accuracy control of the numerical solution and of the experiments is essential i.e. verification applies both for numerics and experimentation. Without it, validation results can be misleading. Mathematics, error estimation and exact statements about these topics are in paper [5], in preparation. Applications, experimentation, concrete examples and interpretations are in paper [6], in preparation.
3 UNCERTAINTIES, FUZZINESS, STOCHASTICS AND FUZZIFICATION

As it was said before, a mathematical problem has a structure $S$ and an input $J$. Assuming that the structure is known, then $S$ maps $J$ into the output set $G$. In the classical deterministic problem $J$ consists of vectors of numbers. These vectors are never completely known. Hence, they are fuzzy vectors (fuzzy set theory). In the simplest way, we associate to every $x \in J$ a membership function, constructed from the experts’ opinions, which characterize the likelihood of the input data. The goal is to determine the membership function of the output.

There are a few important special cases.

a) The membership function is constant. This is the worst case scenario we addressed earlier.

b) The membership function are constructed from probability.

Let us now address the case when the uncertainties in $E$, $\nu$ are described in terms of stochastic functions, assuming their probability field is perfectly known. Let us consider the problem

$-\nabla \cdot (a(x,\omega)\nabla u(x,\omega)) = f(x,\omega), \quad \text{in } \Omega$

$u(x,\omega) = 0 \quad \text{on } \partial \Omega.$

Here $a(x,\omega)$ is a stochastic function described by the Karhunen-Loeve expansion

$a(x,\omega) = \mathbb{E}[a](x) + \sum_{m=1}^{\infty} \lambda_m^{1/2} \psi_m(x) X_m(\omega),$

where $X_m(\omega)$ are uncorrelated random variables with mean value zero and variance one. Then the problem can be transformed into a deterministic problem on $\Omega \times \Gamma$ with $x \in \Omega$ and $y \in \Gamma$. Its solution $u = u(x,y)$ is defined by the weak formulation

$\int_{\Omega} \int_{\Gamma} \rho(y) \nabla_x u \nabla_y v \, dx \, dy = \int_{\Omega} \int_{\Gamma} f(x,y) \rho(y) v \, dx \, dy, \quad \forall v \in H_0^1(\Omega) \times L^2(\Gamma).$

For detailed mathematics of this problem and its numerical solution we refer to the papers [7] and [8].

Up to now, it was assumed that the Karhunen-Loeve expansion is perfectly known. The question is whether the needed information is available. It can be obtained only by experiments. The important question is how accurately the expansion can be obtained from the experiments. An insight in this question can be obtained by the virtual experimentation. This is addressed in the paper [3]. It is shown that to obtain a reasonable accuracy in the Karhunen-Loeve expansion a large number of experiments is needed. Practically, it is impossible to have such a large number of experiments. Hence, a fuzzy set approach is necessary.
Considered another problem

\[-\Delta u(x, \omega) = f(x, \omega)\]
\[u = 0.\]

Here the goal is to get the covariance function of \(u\). The work [2] proves that the desired covariance function, \(\text{Cov}[u]\), satisfies

\[\Delta_x \Delta_y \text{Cov}[u](x, y) = \text{Cov}[f](x, y) \text{ in } \Omega \times \Omega,\]
\[\text{Cov}[u](x, y) = 0, \text{ on } \partial\Omega \times \partial\Omega.\]

Here \(\text{Cov}[f](x, y)\) is covariance function of \(f(x, \omega)\). We can use for example the worst scenario approach mentioned earlier when for every \((x, y)\) the covariance \(\text{Cov}[f](x, y)\) is a fuzzy number.

Let me mention an example. We have asked various our friends - material scientists - about covariance length \(L\) for aluminum alloy. For stationary fields \(L\) can be taken as the distance where the covariance is 1/2 of the value at zero. We received very different guesses in the interval \((20^\circ, 100^\circ)\).

3.1 Conclusions

1. The information needed in the mathematical model for the prediction has always uncertainties i.e. it is fuzzy.
2. More complex problems need usually more information which is more fuzzy.
3. The predictions have to address uncertainties.
4. Stochastic formulation needs information which is practically impossible to obtain and some fuzzy approach is unavoidable.

4 On Hierarchic Modeling

Let us consider three models, \(B\), \(\Sigma_1\) and \(\Sigma_2\) indicated in the Figures 9, 10 and 11 respectively.

4.0.1 The basic model \(B\)

\[\begin{array}{c}
J \\
\rightarrow
\end{array} \begin{array}{c}
B \\
\rightarrow
\end{array} \begin{array}{c}
G
\end{array}\]

Figure 9: 3 boxes diagram. The basic model \(B\).
4.0.2 The surrogate model $\Sigma_1$

![Diagram of surrogate model $\Sigma_1$]

Figure 10: The surrogate model $\Sigma_1$.

4.0.3 The surrogate model $\Sigma_2$

![Diagram of surrogate model $\Sigma_2$]

Figure 11: The surrogate model $\Sigma_2$.

For instance, the basic problem $B$ could be the three-dimensional elasticity problem, and $\Sigma_1$ and $\Sigma_2$ the corresponding one- and two-dimensional problems approximating $B$. Every problem i.e. $B$, $\Sigma_1$, $\Sigma_2$ can be understood as a separate prediction problem. In this case, the validation should be done for all these models.

Alternatively, $\Sigma_2$ can be understood as an approximation of $\Sigma_1$ and $\Sigma_1$ as approximation of $B$ respectively. In this second case the problem $B$ is validated while $\Sigma_1$ (resp. $\Sigma_2$) is verified, i.e. the a-posteriori error estimation of $\Sigma_1$ with respect to $B$ is computed.

It is essential that the "lower" model, i.e. the surrogate, is easier to solve. A posteriori estimate of the accuracy of the surrogate problem - if the data of the basic problem are known - can show that the prediction based on $\Sigma_1$ is as reliable as the one based on the basic model. Then we have to solve the simpler surrogate problem (in various contexts this is called Occam's Razor Principle$^1$).


---

$^1$William Occam (ca 1290-1349), an English philosopher and theologian. He was strict nominalist and a critic of toism. He became famous for his "razor principle": *Essentia non sunt multiplicanda praeter necessitatem* (Essential terms should not be multiplied beyond necessary).
5 Conclusions

1. The bottleneck in the reliability of computer predictions is in the mathematical model. The validation is essential.

2. The selection of the sequence of validation problems is essential. The mathematics of the selection is at its very beginning.

3. The Validation step needs close relation between experimentalists and computational scientists.

4. The verification is essential part of the validation process and of the predictions.

5. Much more theoretical work is needed.

References


