Principles of benchmark studies (Verification & Validation)

Lesław Kwaśniewski
Warsaw University of Technology

Plan for presentation

- Main question
- Spectacular example
- Some facts
- Predictive capabilities
- Verification & Validation
  - General aspects
  - Errors & Uncertainties
  - Definitions
  - Validation
    - Domains
    - Calibration
    - SRQ
    - Validation Metrics
  - Example
- Databases of benchmark problems
- Summary
What are the predictive capabilities of our computer simulations (in Structural Fire Engineering)?

Computational Science and Engineering (CS&E)
Computational Engineering and Physics (CE&P)

“Essentially, all models are wrong, but some are useful”


Spectacular example of a software bug

F-22 Squadron Shot Down by the International Date Line (2007)

Maj. Gen. Don Sheppard (ret.):
"...At the international date line, whoops, all systems dumped and when I say all systems, I mean all systems, their navigation, part of their communications, their fuel systems.

......

It was a computer glitch in the millions of lines of code, somebody made an error in a couple lines of the code and everything goes.”

http://www.defenseindustrydaily.com
Some facts

Number of articles according Google Scholar with FIRE in the title with FIRE + “FINITE ELEMENT” anywhere in the article

![Graph showing the number of articles from 2007 to 2012](image)

Moore’s law states that computer power increases by a factor of two every eighteen months.
Some facts

LS-DYNA® - a finite element (FE) based simulation software - had originally 50,000 lines of code and then approached 2 million lines in little more than a decade.

“...for many years the Journal of Applied Mechanics shunned papers on the finite element method because it was considered of no scientific substance.


Barriers to computability - smoothness and stability of the response, uncertainties, coupled physics, ...

The number of execution paths in a typical commercial code is often so large that some paths are never explored, even after years of service.

Belytschko T., Mish K., "Computability in nonlinear solid mechanics"
http://www.tam.northwestern.edu/tb/computability_w_figs.pdf

What are the predictive capabilities of our computer simulations?

- Modeling electronic systems e.g. Verilog
- Computational Science and Engineering (CS&E)
  Computational Engineering and Physics (CE&P)
- Numerical weather prediction
- Numerical models in economics

VERY GOOD

POOR
What are the predictive capabilities of our computer simulations (Computational Science and Engineering (CS&E))? 

**Linear FE Dynamic Analysis**

**Linear FE Static Analysis**

**Nonlinear Static (Stability)**

**Transient Dynamics (Crash Tests)**

**Computational Fluid Dynamics (CFD)**

**Structural Fire Engineering**

---

**Verification & Validation**

What are the recommended procedures?

- ASME Y14.5M-2006: Guide for Verification and Validation in Computational Solid Mechanics
- Concepts of Model Verification and Validation
- Guide for the Verification and Validation of Computational Fluid Dynamics Simulations
- Simulation Verification, Validation and Accreditation Guide

---

COST Action TU0904
Errors & Uncertainties

"Error of measurement (calculation) - the result of a measurement (calculation) minus the value of the measurand" (accurate solution),

"Uncertainty - a parameter associated with the result of a measurement (solution) that characterizes the dispersion of the values that could reasonably be attributed to the measurand" (accurate solution).

An error is the difference between our result and the value (given or imaginable) that is considered to be accurate.

The uncertainties can be thought as the bounds of the errors which can be irreducible (aleatory) or reducible (epistemic).

Errors in testing usually have two components, random (precision) and systematic (bias).

Errors & Uncertainties

Most of the literature point out at five primary sources of errors in computational solutions:

• insufficient spatial discretization,
• insufficient time discretization,
• insufficient iterative convergence,
• computer round-off,
• computer programming.


In the nonlinear computation there are also errors which are not the result of a programmer’s mistakes or improper use of the code but are an inherent part of the solution procedures.


Definitions of Verification & Validation

**Verification** is supposed to deliver evidence that mathematical models are properly implemented and that the numerical solution is correct with respect to the mathematical model.

**Verification** uses comparison of computational solutions with highly accurate (analytical or numerical) **benchmark** solutions and among themselves, whereas **validation** compares the numerical solution with the experimental data.

**Verification** should precede **validation**.

Experimental **validation** is the final check to reveal possible errors and to estimate the accuracy of the simulation.

**Validation** can be practically split into three tasks:
• to detect and separate the model’s significant discrepancies,
• to remove and reduce removable and unavoidable errors,
• to evaluate uncertainties in the results.

„**Verification** deals with mathematics; **validation** deals with physics”

Verification

A “posteriori” approach where the reasoning is based on the experience coming from repeated calculations.

**Benchmarking** and comparison with simplified models

- question about the sufficient mesh resolution
- the mesh size should be dependent of specific quantities of interest
- simple checks: mass, reactions (Equations of Equilibrium must be always satisfied)
- simple check: conservation laws for mass, momentum, and energy (non-physical energy components < 5%)
- simplify FE model: simple loading, simplified BC, materials
- Quasi-static loading as a special case for transient analysis
- importance of the databases collecting well-documented benchmark problems

### Mesh density study

- Discretization error

\[
E = f_h - f_{\text{exact}} = Ch^p + H.O.T.
\]

\( \ldots \ldots \)

- Order of convergence

\[
p = \frac{\ln \left( \frac{f_3 - f_2}{f_2 - f_1} \right)}{\ln(r)}
\]

- Estimate of the asymptotic solution

\[
f_h = 0 \approx f_1 + \frac{f_1 - f_2}{r^2 - 1}
\]

- \( E_1 \) is the estimator of the relative error

\[
E_1 = \frac{\varepsilon}{r^p - 1}, \quad \varepsilon = \frac{f_1 - f_2}{f_2}
\]

- Grid Convergence Index – GCI procedure (Richardson extrapolation)

\[
GCI = \frac{F_3 |\varepsilon|}{r^p - 1} \times 100\%
\]

Verification

Sensitivity study

- determine the crucial input parameters
- evaluate the possible range of their variation
- helps to identify the sources of errors
- can reveal if the considered problem is extremely sensitive to the input variation (imperfection sensitive structures or the on-off processes)
- imperfections can be applied to geometry, loading (as eccentricities), boundary conditions, and material properties

Validation Domains

Experimental validation is the final check to reveal possible errors and to estimate the accuracy of the simulation.

- differences between mathematical and physical systems
- differences between computerized and mathematical models
- distinction between a physical system (our concept of it) and the subject of an experiment used for validation

Application and validation domains

Application domain defines the intended boundaries for the predictive capability of the computational model.

Validation domain characterizes the representation capabilities of the experiment.

Validation

**Validation hierarchy**

The experiments for the considered system are usually divided into three or four levels (tiers) representing different degrees of complexity.

![Hierarchical verification and validation diagram]

**Validation and calibration**

The idea of the **calibration** procedure is to establish the quantities of modeling parameters that give the model’s response closest to the actual experimental data.

The calibration is performed through comparison between an experiment and repeated calculations with modified input parameters.

It can happen that due to superimposing of errors we can get good correlation between experimental and numerical results for a wrong model defined by incorrect input parameters.

**System response quantity SRQ**

Validation is based on the comparison between computational results and experimental data.

An experiment can provide much less information than the calculation.

Selection of the system response quantity (SRQ) is often limited by the experiment output.
Validation Metrics

- **Viewgraph Norm**
- **Deterministic**
- **Experimental Uncertainty**
- **Numerical Error**
- **Non-deterministic Computation**
- **Quantitative Comparison**


Example

Steel cube dropped on a cantilever thin-walled steel beam
- input variation vs. output variation
- system response quantity
- on-off processes (zero-one switching)

Drop heights 200.01 mm or 199.63 mm above the top flange.

„If the answer is highly sensitive to perturbations, you have probably asked the wrong question”

Example

Eigenmodes corresponding to the identified frequency peaks.

Time history (top) of relative displacement for top corners at the beam and corresponding frequency spectrum (bottom).
Databases of benchmark problems

- National Agency for Finite Element Methods and Standards (NAFEMS) ~ 280 verification benchmarks (thermal analysis-14, thermal stress analysis-8)
- ANSYS® - around 250 problems
- SAFIR – significant amount of evidence presented in publications
- VULCAN, Fluent, ....
- DIN EN 1991-1-2 Raul ZAHARIA
- One Stop Shop in Structural Engineering http://www.mace.manchester.ac.uk/project/research/structures/strucfire/default.htm

Summary

- For the non-linear problems there are unavoidable errors that are an inherent part of the solution procedures.

- Separation of all sources of errors is today impossible for many complex systems.

- Verification through the testing of different solution options is necessary.

- For the wide range of conditions found in practice, it is impossible to define general requirements guaranteeing satisfying accuracy.
Thank you for your attention!