Science Based Nuclear Energy Systems Enabled by Advanced Modeling and Simulation at the Extreme Scale

White Paper on Verification, Validation, and Uncertainty Quantification

Richard I. Klein[1] and Paul J. Turinsky[2]

[1] University of California, Berkeley Department of Astronomy and Lawrence Livermore National Laboratory

[2] North Carolina State University, Department of Nuclear Engineering

Introduction

The purpose of this White Paper is to provide a framework for understanding the role that Verification and Validation (V&V), Uncertainty Quantification (UQ) and Risk Quantification, collectively referred to as VU, is expected to play in modeling nuclear energy systems. First, we provide background for the modeling of nuclear energy based systems. We then move to a brief discussion that emphasizes the critical elements of V&V as applied to nuclear energy systems but is general enough to cover a broad spectrum of scientific and engineering disciplines that include but are not limited to astrophysics, chemistry, physics, geology, hydrology, chemical engineering, mechanical engineering, civil engineering, electrical engineering, and nuclear engineering material science, etc. Finally, we discuss the critical issues and challenges that must be faced in the development of a viable and sustainable VU program in support of modeling nuclear energy systems.

Background for Modeling Nuclear Energy Systems

Nuclear energy systems and their associated fuel cycles involve complex, interacting subsystems whose modeling requires expertise that span many scientific and engineering disciplines. The key stages of a nuclear fuel cycle include mining and milling, conversion, enrichment, fuel fabrication, power production, temporary spent fuel storage, separations, and nuclear waste disposal. Whether the fuel cycle is open or closed, fuel type and reactor type will determine which of these stages come into play and the specific nature of each stage. A heavy reliance on computational simulation of each of the fuel cycle stages has historically played a major role in the advancement of nuclear energy, in particular the stages of power production, separations and nuclear waste disposal. The validation of these computational simulations has been based upon an extensive array of experiments, ranging from basic physics experiments, e.g. nuclear data, to single effects experiments, to integral system level experiments. Collectively these experiments have cost tens of billions of dollars and been executed over several decades. However, there still remain gaps in the experimental basis in areas such as severe accidents and aging.

Using the experimental base, it has been possible to improve simulation capabilities via model, numeric and input data enhancements. In more recent years, it has also been possible to estimate the uncertainties in best estimate predictions. This has been done by mathematically propagating the uncertainties in input data, initial conditions and sub-models through simulators. This process has been accepted by the USNRC as indicated in the Code Scaling, Applicability and Uncertainty (CSAU) methodology. However, since current nuclear system simulation most times are not based upon micro-scale science based models but macro-scale models, i.e. heat transfer sub-model, the number of parameters for which uncertainties must be treated is limited, e.g. fifty.

As simulation capability is developed to be more science based, a heavier reliance on micro-scale sub-models will evolve. Given that designers require macro-scale responses to make informed design decisions, this implies that increased usage of multiscale modeling will be necessary. Further, since nuclear energy systems like any complex system involves the interaction of many physical phenomena, tightly coupled, multiphysics modeling will also play a more significant role in the future. To support the development of multiphysics and multiscale modeling, new VU capabilities will be necessary. A portion of this new capability required will be possible based upon advances in computational power, but others will require advances in the mathematical and algorithmic foundations of VU.

For nuclear energy systems, the motivation for completing VU is two fold. The obvious reason is to provide users of simulation packages confidence in the system responses predicted and knowledge of prediction uncertainties. However, for nuclear energy systems VU is also performed because it is required by the licensing body, specifically the US Nuclear Regulatory Commission, built upon the premise of an extensive experimental data base regarding system attributes of interest.

VU has as its objective to be able to predict with ***confidence***, using models captured in computer simulation, the ***best estimate*** values and their associated ***uncertainties*** of complex system attributes, accounting for all source of error and uncertainty, i.e.

- modeling,
- numerical treatment,
- epistemic uncertainties (e.g. data including correlations),
- aleatory uncertainties (random phenomena), and

  - Initial, e.g. state condition, and boundary, e.g. domain decomposition, conditions.

If this objective is satisfied, it will support the following favorable outcomes:

- To make confident, risk-informed decisions when considering alternative designs and operations, and nuclear safety.
- More specifically to support
  - the identification of code development needs,
  - the identification and design of required validation experiments,
  - design decision making in regard to managing margins, and
  - presenting the risk-informed safety case with the regulatory body.

A challenge when VU is applied to nuclear systems is that it must be able to predict high-impact consequences of low probability events with high confidence, factoring in aging effects, with limited experimental data at the macro-scale. This challenge is noted to be similar to that associated with nuclear weapons' stewardship.

Nuclear System Models' Attributes

Given the diversity between the stages of the nuclear fuel cycle, there is considerable diversity in the associated simulation models. An example using the power production stage of the nuclear fuel cycle will serve to indicate the complexity of nuclear systems simulation models. To model a nuclear power plant, which includes the mechanical, electrical and nuclear components and systems, structures, and external environment, the following must be modeled: thermal-hydraulic behavior of fluid circuits including fluid-structure interactions, thermal behaviors of components making up the system, material behaviors factoring in radiation, temperature, pressure, and chemistry effects, structural responses, instrumentation responses, control and protection systems logic, reactor physics and radiation fields. It is recognized that weak to strong coupling exists between these effects due to natural or engineered feedback effects. Today, with capabilities reflected in such simulation packages as TRACE, TRAC, RELAP and SASSYS codes, one is limited in not only modeling detail, but also the degree of coupling that can be represented. Introducing science-based multiphysics and multiscale modeling will only make these challenging modeling problem orders of magnitude more challenging with regard to best estimate calculations.

Overview of Critical Elements of V&V and Uncertainty Quantification

It is useful to state the definitions.

> **Verification**: Verification is the process of confirming that a computer code correctly implements the algorithms that were intended. This is the process of confirming that the equations are numerically solved accurately.

Validation: Validation is the process of confirming that the predictions of a code adequately represent measured physical phenomena. This is the process of confirming that the equations are (physically) accurate.

Verification

Code verification is the most general component of V&V. It answers, or seeks to answer, three specific questions: (1) Are the equations represented by a code *mathematically* (not physically) correct? (2) Are the algorithms that provide the numerical solution of the mathematical equations themselves mathematically correct? (3) Is the software implementation of these algorithms correct (that is, free of faults)?

It is useful to think about verification testing and test problems in three ways. The first of these is the ***structure*** of the chosen test problems, that is, the logical principles underlying them. This addresses the important question of why given test problems are chosen, and how they are organized.

The second is the specific ***construction*** of the test suite, or the specific means chosen by the code team for populating the test problem suite. Finding or developing new test problems that fully address the complexities of multiphysics codes is a tremendous challenge.

The third aspect of importance for verification test problems is that of ***assessment***, specifically the criteria that are applied for deciding whether or not the code has passed or failed a given test problem. Verification test problems are intended to be strong tests of the code. Therefore, assessment must be objective and rigorous, and well-documented.

Solution Verification

Solution verification is quantification of the numerical error in a presented calculation. This answers a direct question: What is the error in a given calculation? Unfortunately, this is all but impossible to perform completely and rigorously for complex calculations. However, it can be partially and practically addressed by explicit discretization robustness and convergence studies, formal error estimation procedures, inference from test problem suites, and – possibly with some danger – inference from previous experience (i.e. judgment). Past experience can count for much *if properly understood and presented*.

Code verification can be completely achieved, and calculations can still be inaccurate, due to poor discretizations (lack of converged calculations). More generally, verification of the correct functioning of algorithms cannot be partitioned as cleanly as we would like. It may be impossible to determine that algorithms are failing only on available test problems; the failures may appear only on large-scale problems for which there is no referent solution. In validation, explicit solution verification must be performed. It targets the numerical errors present in any comparison of a calculation with experimental data. The fundamental question that must be recognized, if not completely answered, is “Does the numerical error fatally corrupt the comparison with experimental data?” In the absence of acknowledgment of this problem, comparison with experimental data is irrelevant.

Solution verification, in the absence of completely rigorous mathematics applicable to the full scope of the mathematical equations being solved, is essentially empirical. The key procedures that offer promise are: (1) *a posteriori* error estimation; (2) convergence studies; (3) numerical error models; (4) uncertainty quantification methods treating the numerical error as an epistemic (lack-of-knowledge) uncertainty.

Validation

Experimental Validation

Validation is fundamentally an experimental challenge. The equations that are solved in the codes used for nuclear energy system design are determined to be physically accurate (for a given application) through confrontation with experimental data having quality suitable for achieving the goals of validation. Because of limited resources, it is important to prioritize validation tasks. The logical desire to achieve a complete validation of a complex code for a predictive complex multi-physics application must be balanced against these constrained resources. Key elements (mainly necessary, but not claimed to be sufficient) of experimental validation, are inevitably:

1. Precise specification of the needed validation tests to optimize the alignment of validation calculations with executed experiments. This requires sophisticated two-way communication between those who execute validation experiments and those who perform validation calculations. Validation is weakened when experimental data are not validation quality. The expectation is that the experiments themselves have been subjected to verification and validation to provide the highest quality data. That is, **experiment verification** confirms that the experiment was executed correctly; **experiment validation** confirms that the correct experiment was executed.

2. Performance of calculation verification for all validation calculations.

3. Quantification of measurement/computational prediction comparisons, including quantified uncertainty. This *requires* (a) experimental error bars that encompass experimental uncertainty and (b) calculation error bars that encompass calculation uncertainty determined by a program of simulation uncertainty quantification (UQ)

Validation calculations are calculations that are compared with validation quality experimental data for the purpose of inferring physical accuracy of the associated calculations. Validation calculations have the specific purpose of enabling an assessment of the physical quality/physical accuracy/predictive capability of the code for the application represented by the chosen validation data. The ***experimental data*** that validation calculations are compared with must have specific characteristics in order to be effective in enabling validation. These characteristics include quantified experimental uncertainty, reproducibility and robustness of experimental data, and as directly comparable with calculations as possible.

Experimental Error Bars

Experimental "***Error bars***" is a euphemism for "experimental uncertainty quantification." This is another problem that is unlikely to be completely and rigorously solved for complex experiments. The components of error bars are experimental bias and variability, and various factors enter into these components. The presentation of experimental error bars can literally be error bars of experimental data. To perform validation, some approximation to experimental "error bars" must be accomplished and presented to serve as a starting point for inference about the experimental-computational comparisons. Gross contributions to experimental uncertainty are diagnostic fidelity, experimental variability, and experimental bias. The more we expect to rigorously infer from a validation comparison, the more we need to understand about experimental error bars as quantifications of experimental uncertainty. For example, is an experimental error bar a central tendency of an underlying Gaussian distribution, a statistical confidence interval, a representation of a uniform distribution, a possibility interval, or something else again?

Uncertainty Quantification

The quantification of uncertainty (UQ) in large scale simulations is playing an increasingly important role in the process of code verification and validation. If a simulation is to be quantitatively validated against the results from an experiment, it is crucial to understand the expected uncertainty in the output metrics of the calculation and also have a quantitative determination of the error bars associated with the output metrics from the experiment. In practice, it is possible to assess the true accuracy of a simulation when the experimental uncertainty is less than the predicated uncertainty of the simulation. Error estimates of uncertainty for the experiment usually require that one performs an ensemble of experiments with controlled parameters and understands known systematic errors.

A broad definition of UQ includes risk quantification. For risk quantification, one is not only interested in the uncertainty in the response metrics of some system, but also the impact of the response metrics on risk. Risk may be economic risk, human health risk, and other types of risk that concern an enterprise. Being able to complete UQ is a necessary but not sufficient condition to complete risk quantification, which in addition requires a model that takes system response metrics and their uncertainties as input and produces risk metrics and their uncertainties as the output. In all likelihood, the risk model is itself uncertain, e.g. the impact of a radiation dose on human heath, so convolution of the probability distributions of the system response metrics with the probability distributions of the risk metrics is called for. A quite different example of risk quantification concerns probabilistic risk assessment (PRA), where one is concerned with the likelihood of a given sequence of events occurring, including the uncertainty associated with the stated likelihood. In the following discussion of UQ, we implicitly include risk uncertainty for both of the instances noted above.

The quantification of uncertainty in large scale simulations becomes particularly important when the simulation is used as a predictive tool in describing phenomena in a regime that is outside of the bounds of previous experimental tests or known observations. Examples of this circumstance for nuclear energy systems include accident analysis of nuclear power plants, predicting the effects of aging on materials in a hostile environment, and predicting long term high-level waste repository performance. Without experiments to check against code predictions in such regimes, it becomes essential to quantitatively evaluate the expected uncertainty in code output. This task of UQ is complex in its undertaking for any simulation code that has non-linearly coupled multi-physics algorithms as a representation of the underlying partial differential equations.

In a complex multi-physics simulation code, many aspects of the physics may have a parametric representation or a choice of physics models each with their own degree of approximation. The range or bounds of parametric settings in physical models and the choice of physics models represent a span of uncertainty in the simulation. Typically, simulation codes are used with a particular choice of input physics models and perhaps a typical choice of parametric settings without any exploration of the full uncertainty in the simulation outcome. Occasionally, a few different models are run in a few large scale simulations to uncover an estimation of the range or dispersion of output results and this gives some measure of the uncertainty, but it is usually woefully inadequate for determination of the full uncertainty in the simulation. The problem of determination of uncertainty quantification is complex and is a topic for current research.

To start, one must first identify the known sources of uncertainty in the simulation. This may involve uncertainties associated with approximate models for the underlying physics, approximations in the numerical algorithms used; uncertainties associated with the settings of parameters that are used in physical models; settings that individual algorithms may have to work in a stable fashion; uncertainties associated with various levels of opacity tables, equation of state tables, and of course uncertainties associated with performing the simulation at a given spatial resolution when this resolution is not converged. Considering that a multi-physics code embodies many components of coupled physics, the list of possible sources of simulation uncertainty can be quite large. Moreover, the uncertainties associated with these sources do not combine linearly, but may take on combinatorics of all possible settings. Furthermore, uncertainties associated with various physics models within the code may cancel giving compensating effects. In a realistic multi-physics, multi-dimensional code, the number of parameters whose values may be bounded may be large and the problem of examining the full possible uncertainty resultant from all possible non-linear interactions among the uncertain components becomes exponentially complex. The problem of Uncertainty Quantification becomes one of reducing the computation of the full uncertainty space by a huge factor to become computationally tractable.

The first step in an approach to Uncertainty Quantification is to identify all avenues of certainty for the simulation code. Once this is established, some approach to the development of a sensitivity analysis must be developed to determine which components of uncertainty (algorithmic approximation, parameters, etc.) are the dominant drivers of the output metrics. This is likely to be an iterative process that cannot be determined a priori. In order to perform a sensitivity study to filter out those components of uncertainty that may not dominate the total output uncertainty, one must know the physically or mathematically reasonable bounds of any set of parameters that represent a physical model. The determination of physically reasonable bounds may require a considerable research effort and the quantification of such bounds may be possible with knowledge gained from experiments, analytic analysis and scientific judgment. Given a first estimate of the sensitive drivers of the code's response to parametric and physical model variation, the problem can now be viewed as navigating the uncertainty of these dominant drivers in an N dimensional space where each dimension is representative of a parameter, physical model, and degree of approximation etc. to the underlying code physics. In doing this, correlations within the N dimensional space must be accounted for, which has the potential of reducing the dimensionality. It becomes essential to sample the full N-dimensional space with a set of simulations that are representative of all dimensions of uncertainty within the bounds of those dimensions. Thus, the problem of uncertainty quantification becomes one in which all identifiable uncertainties and their interaction with one another are run through the simulation code to give a predictable total output uncertainty in the code's response to variation over acceptable bounds of all the components. The uncertainty in code response to uncertainties in all the key components of the code can be expressed as the total uncertainty in the main metrics of code output that are objectives of the simulation.

Critical Issues and Challenges in V&V and UQ

Verification

Verification of computational science codes is dominated by testing. Testing remains the most essential contributor to the collection of verification evidence. Sufficient confidence in verification of software firmly rests upon the idea of sufficient testing. Inadequate testing increases the risk of malfunctioning software in important circumstances.

Testing first and foremost depends upon having well-defined tests that a code passes or fails. While simple tests directed at individual code components can be devised that have strong assessment criteria, more complex tests that integrate larger parts of the physics and have greater numerical complexity are very difficult to devise, and it can be extremely difficult to determine related assessment criteria. It is a critical problem in verification to devise such tests, as well as strong assessment criteria that create the verification consequences associated with the use of the test.

Benchmarks for code verification are needed for a wide range of physics and engineering applications with special emphasis on coupled multi-physics. Important areas where solutions to semi-analytic verification test problems are sorely needed include, but are not limited to:

> Component physics semi-analytic test problems and solutions in 1-, 2-, and 3-D. Examples include: hydraulics for single-phase/single component, single-phase/multi-component, two-phase/single-component and two-phase/multi-

component fluids; heat conduction through structures; structural response of structures to applied loads; isotopic composition with irradiation; neutron and gamma spatial interactions rates; and material thermal conductivity with applied irradiation and temperature.

Coupled physics semi-analytic test problems and solutions in 1-, 2-, and 3-D. Examples of interest areas are: thermal/hydraulics, thermal/neutronics; thermal/materials; thermal/structures; neutronic/materials; hydraulic/neutronics; hydraulics/structures; hydraulics/materials; structures/materials; thermal/hydraulics/neutronics; thermal/hydraulics/materials; and thermal/hydraulics/neutronics/structures/materials.

Semi-analytic test problems and solutions for neutron transport beyond flux-limited diffusion, characterized by angle dependent transport solutions. Radiation transport is among the limited class of physics problems where Monte Carlo simulation can provide meaningful test problem solutions.

Research topics in the area of solution verification include:

- Practical methods for estimating or bounding numerical errors associated with spatial and/or temporal discretizations,
- Methods for estimating numerical errors associated with parameters that control the performance of numerical algorithms (e.g., artificial viscosity or hourglassing parameters), particularly in conjunction with other discretization errors, and
- Practical methods for making validation or application decisions with under-resolved models.
- Approaches for solution methods based upon parallel asynchronous algorithms.

Validation

Well characterized validation experiments lie at the heart of simulation and model development. It is through these experiments that model accuracy can be assessed. Experiments can be generally classified into two types: (1) component (i.e. single physics phenomena); and (2) integrated (i.e. coupled physics phenomena). The types of component and integrated validation experiments will vary from application to application.

Component experiments:

High quality experiments for component physics are needed for multi-scale, multi-physics, multi-dimensional codes for nuclear energy systems. Due to the highly non-linear interactions that occur between physical processes, it is important to ensure that the isolated physical process under consideration be assessed for its accuracy. With integrated experiments, it is difficult to distinguish an error in the coupling between component physics from an error in the individual components themselves. This is called a compensating error. Therefore, component physics experiments form a critical part of any validation process. A sample of component validation experiments that are desirable, include but are not limited to: single effect thermal/hydraulic; zero power reactor criticals; materials stress-strain, load deformation of structures; and chemical separations unit components.

Even with the large number of existing experimental data that helps validate the computational physics models of nuclear energy systems, there is a real need for new experimental data that addresses component performance under severe accident conditions, aging of components, and fundamental parameters related to predictions of fuel performance.

Integrated experiments:

Ultimately, the applications under consideration tend to be multi-physics in nature. This means the validation of **coupled/integrated** physics models is of critical importance. In most codes, modularity of physics models means some type of operator splitting must be performed. Therefore, high quality, well diagnosed experiments of integrated physics are needed for multi-scale, multi-physics, multi-dimensional codes. It is this class of experiments that ultimately any multi-physics code must be able to simulate. Examples include but are not limited to: integral thermal/hydraulics for natural circulation systems; fuel performance in power reactors, integral thermal/hydraulics/neutronic/structural/materials under degraded core conditions.

Validation Methodology

Beyond interest in the validation of specific phenomena, there are a number of methodological needs to support validation in a manner that allows quantification of uncertainties in non-linear, coupled multi-physics nuclear energy system applications:

- Advanced statistical methods for making quantitative measurement/prediction comparisons, particularly in the presence of non-negligible variabilities and uncertainties in diagnostics, initial conditions, boundary conditions, and other model inputs.
- Tools to automate the process of quantitative validation.
- Methodologies for validation inference through a hierarchy of validation experiments ranging from simple material characterization test through a series of experiments of increasing complexity.
- Extrapolation inference from a validation parameter space to an application parameter space that is significantly outside the validation database
- Statistical methods for validation when there is only a single well instrumented test.

Uncertainty Quantification

Intelligent statistical sampling techniques will be necessary to sample the full domain of an N-dimensional space of possible outcomes if other methods, e.g. adjoint method and automatic differentiation, are not appropriate. If the dimensionality is high (N>10) then standard sampling techniques (e.g. Monte Carlo) will not be nearly efficient enough to cover the full domain of uncertainty with a number of sample calculations (likely in 2-D) that are computationally feasible. Adaptive sampling procedures will have to be developed that will sample the full N-dimensional domain in an efficient enough way that clusters of sample simulations in those regions of the domain will capture where the sensitivity of the simulation response to variation in parameters, models, approximations etc. is highest. Examining the code response to the full variation of all parameters in the physical models comprising the code by intelligent sampling of the N-dimensional parameter space will provide a total output certainty, but the full ensemble of models consisting of the combinations of the parameters and their variations may not satisfy data from available experiments. Thus, it is necessary to find the ensemble of models and the parametric settings that comprise them that at least satisfy available data. This requires an intelligent filtering of the full ensemble of models that cover all of the uncertainty space of the simulation. Once such a filtered set of parametric settings becomes available that give rise to an ensemble of output calculations that satisfy known experimental data from different experimental regimes, techniques must then be developed to propagate this set of models to regimes for which no experimental data exits and use the ensemble set in this regime to predict the total uncertainty of output metrics for those regimes.

The entire process of Uncertainty Quantification has important challenges that must be addressed. The study of these issues is critical to any UQ component of a V&V program plan. Many of these issues are under current exploration in the laboratories V&V programs.

1. What approaches can be developed that allow for the determination of the dominant sensitivities in the code that drive the uncertainty in the output of a large scale simulation (particularly when the outputs are highly non-linear functions of the inputs).

2. How do these approaches compare with one another in determining the dominant sensitivity drivers of output uncertainty?

3. What approaches can be developed to propagate the uncertainty associated with a large number of uncertain parameters (N>>10) through the simulation to determine a prediction of the total uncertainty in the output metrics of large scale simulations. How can this be accomplished in a computationally efficient way when dimensionality of uncertainty space is high (i.e. N>>10) and the computation cost of a code run is very high?

4. What approaches can be developed to reduce the dimensionality of a high dimension UQ space (e.g. "The Curse of High Dimensionality")?

5. How sensitive is the final uncertainty of output code metrics to the input probability density functions of the settings of code parameters in the physical models?

6. How many sample calculations are required to obtain the output uncertainty in a code simulation for an arbitrary number of uncertainty dimensions N? How can the accuracy of the output uncertainty for a given number of sample simulations covering an arbitrary number of uncertainty dimensions be quantitatively determined?

7. What techniques can be developed to determine if the ensemble of models that all fit known experimental data is complete?

8. How do we V&V a UQ methodology?

9. How do quantitatively determined output uncertainties compare when determined by different UQ methodologies?

10. What experiments can be designed that can be used to test a UQ methodology?

11. How can the confidence in a UQ methodology in the determination of the uncertainty in code output metrics be measured when possible experiments that could potentially test the methodology are not in the desired regime of code simulations?

12. Are there benchmark problems that can be developed that represent a fair test of competitive methodologies for UQ and sensitivity analysis?

13. What are practical methodologies for the aggregation and propagation of aleatory uncertainties, epistemic uncertainties, and combined aleatory and epistemic uncertainties?

14. How is UQ to be completed when parameter and other sources of uncertainties state condition dependent for transient problems?

15. How to efficiently propagate uncertainties through loosely coupled physics packages typical of operator splitting approaches?

16. How to propagate uncertainties through scales for multiscale problems?

17. How to gain computational efficiency advantage of multiphysics problems composed of a mixture of linear and nonlinear individual physics responses?

18. How to complete UQ for PRA when using dynamic event sequences?

19. What special problems present themselves with extending UQ analysis to exascale architectures where 10^6 ensemble simulations become feasible and enormous data sets results from the UQ analysis?

20. How do we develop codes of the future that intrusively propagate uncertainty as the simulation evolves in time?