UNCERTAINTY ANALYSIS METHODS

Sankaran Mahadevan
Email: sankaran.mahadevan@vanderbilt.edu
Vanderbilt University, School of Engineering
Consortium for Risk Evaluation with Stakeholders Participation, III
Nashville, TN 37235

and

Sohini Sarkar
Email: sohini.sarkar@vanderbilt.edu
Vanderbilt University, School of Engineering
Consortium for Risk Evaluation with Stakeholders Participation, III
Nashville, TN 37235

November 2009
CBP-TR-2009-002, Rev. 0

CONTENTS **Page No.**

CONTENTS (contd) **Page No.**

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

ARIMA	Autoregressive Integrated Moving Average
CDF	Cumulative Distribution Function
FORM	First-order Reliability Method
GP	Gaussian Process
IID	Independent and Identically Distributed
KLE	Karhunen-Loeve Expansion
LHS	Latin Hypercube Sampling
MCMC	Markov Chain Monte Carlo
MPP	Most Probable Point
PA	Performance Assessment
PCE	Polynomial Chaos Expansion
PDF	Probability Density Function
RE	Richardson Extrapolation
SRSM	Stochastic Response Surface Method
SRV	Standard Random Variable
S-N	Stress vs. Number of Cycles
SORM	Second-order Reliability Method
V&V	Verification and Validation

LIST OF NOMENCLATURE

$\mathbf{B}$	Vector of Reliability Indices
B	Bayes Factor
B^i	Backward Operator of Order i
$C(\chi_1,\chi_2)$	Covariance Function of a Random Process
$N(\mu\sigma^2)$	Gaussian Distribution With Mean μ and Variance σ^2
$P(H_o),P(H_o)$	Prior Probabilities of Null and Alternative Hypotheses
$P(H_o\|D),P(H_1\|D)$	Probabilities of H_0 and H_a Given Observed Data D
$P(D\|H_o),P(D\|H_1)$	Probabilities of Observing Data D Given H_0 and H_a
P_f	Probability of Failure
$\mathbf{R}$	Correlation Matrix
$\mathbf{X}$	Vector of Input Random Variables
c,c_1	Constants
$f_i(x)$	Eigenfunctions of $C(x_1,x_2)$
$f_x(x)$	Joint Probability Density Function of $\boldsymbol{X}$
$\mathbf{g(X)}$	Model Output
k	Regulatory Requirement
y_{exp}	Experimental Observation
y_{true}	True Value of the Parameter
z_t,z_{t-i}	Observations at t^{th} and $(t\text{-}i)^{th}$ Time Steps
∇^d,∇^D	Backward Difference Operators of Degree d and D
$\Phi(\mathbf{B},\mathbf{R})$	Standard Normal Multivariate CDF
$\Phi_p()$	Polynomial of Order P
$\Theta_Q()$	Polynomial of Order Q
α	Unit Gradient Vector of the Limit State in Standard Normal Space
β	Reliability Index
ε_d	Discretization Error
ε_{exp}	Measurement Error
ε_h	Input Parameter Error
ε_{model}	Model Form Error
ε_{num}	Numerical Solution Error
ε_{obs}	Overall Prediction Error
ε_s	Stochastic Analysis Error
ε_t	Error at t^{th} Time Step
ε_{t-i}	Error at $(t\text{-}i)^{th}$ Time Step
η	Vector of Random Variables in Uncorrelated Standard Normal Space
λ_i	Eigenvalues of $C(x_1,x_2)$
ϕ_1	Coefficients of ARIMA Model
$\phi_p()$	Polynomial of Order p
θ_i	Coefficients of ARIMA Error Model
$\theta_q()$	Polynomial of Order p
$\boxtimes(x,\chi)$	Random Process Dependent on Spatial Coordinate x and an Event χ
$\boxtimes(x)$	Mean of the Random Process $\boxtimes(x,\chi)$
ξ_1	Sets of Uncorrelated Standard Normal Random Variables

UNCERTAINTY ANALYSIS METHODS

Sankaran Mahadevan
Sohini Sarkar
Vanderbilt University, School of Engineering
Consortium for Risk Evaluation with Stakeholders Participation, III
Nashville, TN 37235

ABSTRACT

This report surveys available analysis techniques to quantify the uncertainty in performance assessment (PA) arising from various sources. Three sources of uncertainty – physical variability, data uncertainty, and model error – are considered. The uncertainty quantification methods are described in the context of four types of analyses needed, namely, (1) quantification of uncertainty in the inputs to the PA models, (2) propagation of input uncertainty through the PA models, (3) model error quantified through verification and validation activities, and (4) probabilistic PA. Random variable and random process descriptions of physical variability are outlined. Methods for handling data uncertainty through flexible families of probability distributions, confidence bounds, interval analysis and Bayesian analysis are described. Useful surrogate modeling and sensitivity analysis techniques for efficient uncertainty propagation analysis are discussed, as well as methods to quantify the various sources of model error. Statistical hypothesis testing techniques (both classical and Bayesian) are discussed for the validation of PA models, and a Bayesian approach to quantify the confidence in model prediction with respect to field conditions is developed. First-order approximations as well as efficient Monte Carlo sampling techniques for probabilistic PA are described.

1.0 INTRODUCTION

Uncertainty quantification is important in assessing and predicting performance of complex engineering systems, especially in the absence of adequate experimental or real-world data. Simulation of complex physical systems involves multiple levels of modeling ranging from the material to component to subsystem to system. Interacting models and simulation codes from multiple disciplines (multiple physics) may be required, with iterative analyses between some of the codes. As the models are integrated across multiple disciplines and levels, the problem becomes more complex and assessing the predictive capability of the overall system model becomes more difficult. Many factors contribute to the uncertainty in the prediction of the system model including: variability in model input variables, modeling errors, assumptions and approximations, measurement errors, and sparse and imprecise data.

The overall goal of this report is to discuss possible methods and tools for quantifying uncertainty. Sources of uncertainty are listed below:

- Physical variability
- Data uncertainty
- Model error

Physical variability: This type of uncertainty, also referred to as aleatory or irreducible uncertainty, arises from natural or inherent random variability of physical processes and variables, due to many factors such as environmental and operational variations,

construction processes, and quality control. This type of uncertainty is present both in system properties (e.g., material strength, porosity, diffusivity, geometry variations, reaction rates) and external influences and demands on the system (e.g., concentration of chemicals, temperature, humidity, mechanical loads). As a result, in model-based prediction of system behavior, there is uncertainty regarding the precise values for model parameters and model inputs, leading to uncertainty about the precise values of the model output. Such quantities are represented in engineering analysis as random variables, with statistical parameters such as mean values, standard deviations, and distribution types estimated from observed data or in some cases assumed. Variations over space or time are modeled as random processes.

Data uncertainty: This type of uncertainty falls under the category of epistemic uncertainty (i.e., knowledge or information uncertainty) or reducible uncertainty (i.e., the uncertainty is reduced as more information is obtained). Data uncertainty occurs in different forms. In the case of a quantity treated as a random variable, the accuracy of the statistical distribution parameters depends on the amount of data available. If the data is sparse, the distribution parameters themselves are uncertain and may need to be treated as random variables. On the other hand, information may be imprecise or qualitative, and it is not easy to treat this type of uncertainty through random variables. In some cases, data regarding some variables may only be available as a range of values, based on expert opinion. Non-probabilistic representations such as fuzzy sets and evidence theory are available for describing such uncertainties. Measurement error (either in the laboratory or in the field) is another important source of data uncertainty.

Model error: This results from approximate mathematical models of the system behavior and from numerical approximations during the computational process, resulting in two types of error in general – solution approximation error, and model form error.

The performance assessment (PA) of a complex system involves the use of numerous analysis models, each with its own assumptions and approximations. The errors from the various analysis components combine in a complicated manner to produce the overall model error. This is also referred to as model bias.

The roles of several types of uncertainty in the use of model-based simulation for performance assessment can be easily illustrated with the following example. Consider the probability of an undesirable event denoted by $g(\mathbf{X}) < k$, which can be computed from

$$P(g(\mathbf{X}) < k) = \int_{g(\mathbf{X})<k} f_{\mathbf{X}}(\mathbf{x})\ d\mathbf{x} \qquad (1)$$

where:

$\mathbf{X}$ is the vector of input random variables, $f_X(\mathbf{x})$ is the joint probability density function of $\mathbf{X}$, $g(\mathbf{X})$ is the model output, and k is the regulatory requirement in performance assessment.

Every term on the right hand side of Equation (1) has uncertainty. There is inherent variability represented by the vector of random variables $\mathbf{X}$, data uncertainty (due to inadequate data) regarding the distribution type and distribution parameters of $f_X(\mathbf{x})$, and model errors in the computation of $g(\mathbf{X})$. Thus it is necessary to systematically identify the various sources of uncertainty and develop the framework for including them in the overall PA uncertainty quantification.

The uncertainty analysis methods covered in this report are grouped along four major steps of analysis that are needed for probabilistic PA:

- Input uncertainty quantification
- Uncertainty propagation analysis
- Model uncertainty quantification (calibration, verification, validation, and extrapolation)
- Probabilistic performance assessment

A brief summary of the analysis methods covered in the four steps is provided below:

Input uncertainty quantification: Physical variability of parameters can be quantified through random variables by statistical analysis. Parameters that vary in time or space are modeled as random processes or random fields with appropriate correlation structure. Data uncertainty that leads to uncertainty in the distribution parameters and distribution types can be addressed using confidence intervals and Bayesian statistics. Methods to include several sources of data uncertainty, namely, sparse data, interval data and measurement error, are discussed.

Uncertainty propagation analysis: Both classical and Bayesian probabilistic approaches can be investigated to propagate uncertainty between individual sub-models and through the overall system model. To reduce the computational expense, surrogate models can be constructed using several different techniques. Methods for sensitivity analysis in the presence of uncertainty are discussed.

Model uncertainty quantification (calibration, verification, validation, and extrapolation): Model calibration is the process of adjusting model parameters to obtain good agreement between model predictions and experimental observations (McFarland, 2008). Both classical and Bayesian statistical methods are discussed for model calibration with available data. One particular concern is how to properly integrate different types of data, available at different levels of the model hierarchy. Assessment of the "correct" implementation of the model is called verification, and assessment of the degree of agreement of the model response with the available physical observation is called validation (McFarland, 2008). Model verification and validation activities help to quantify model error (both model form error and solution approximation error). A possible Bayesian approach is discussed for quantifying the confidence in model extrapolation from laboratory conditions to field conditions.

Probabilistic performance assessment: Limit-state-based reliability analysis methods are discussed to help quantify the PA results in a probabilistic manner. Methods are also discussed to compute the confidence bounds in probabilistic PA results. Monte Carlo simulation with high-fidelity analyses modules is computationally expensive; hence surrogate (or abstracted) models are frequently used with Monte Carlo simulation. In that case, the uncertainty or error introduced by the surrogate model also needs to be quantified.

Figure 1 shows the four stages, within a conceptual framework for systematic quantification, propagation and management of various types of uncertainty. The methods discussed in this report address all the four steps shown in Figure 1. While uncertainty has been dealt with using probabilistic as well as non probabilistic (e.g., fuzzy sets, possibility theory, evidence theory) formats in the literature, this report will focus only on probabilistic analysis, mainly because the mathematics of probabilistic computation are very well established, whereas the non-probabilistic methods are still under development and generally result in interval computations that are expensive when applied to large problems with many variables.

The different stages of analysis in Figure 1 are not strictly sequential. For example, stage 3 (verification and validation – commonly denoted as V&V) appears after system analysis and uncertainty propagation. However, it is almost impossible to perform V&V on the system scale, because of extrapolation in time and space; therefore V&V is usually done for the sub-models. Also, several of the inputs to the overall system model may be calibrated based on the results of sub-model analysis, sensitivity analysis, and V&V activities. Thus the four stages in Figure 1 simply group together the different types of analysis, and might occur in different sequences for different problems and different sub-models.

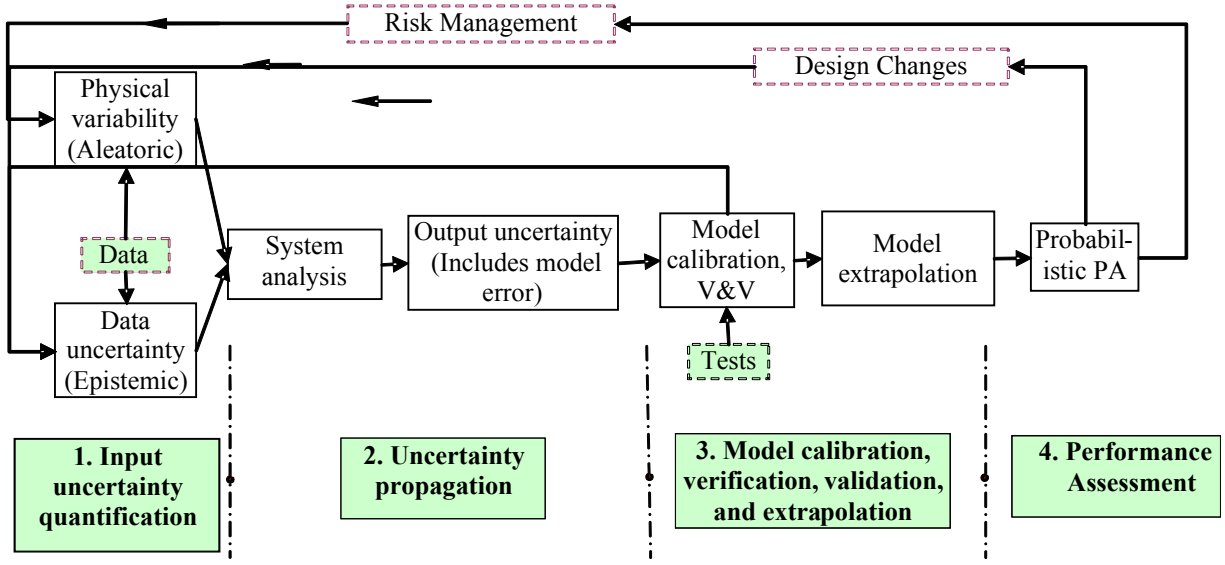

Figure 1. Uncertainty Quantification, Propagation and Management Framework

[1] The box Data in the input uncertainty quantification stage includes laboratory data, historical field data, literature sources, and expert opinion.
[2] The box Design Changes may refer to conceptual, preliminary, or detailed design, depending on the development stage.
[3] The boxes Design Changes and Risk Management are outside the scope of this report, although they are part of the overall uncertainty framework.

Uncertainty analysis methods currently used in PA activities are discussed in another Cementitious Barriers Partnership report. The quantification of uncertainty in current PAs is limited to quantifying the probability distributions of key parameters. A more comprehensive implementation of uncertainty quantification for environmental PAs has been hampered by the numerous sources of uncertainty and the long time durations considered in the PAs. The methods presented in this report provide a basis for advancing the current state of the art in uncertainty quantification of environmental PAs.

The remainder of this report is organized as follows: Section 2 discusses methods to quantify the uncertainty in the inputs to the system analysis model, addressing both physical variability and data uncertainty. Model error is addressed in Sections 3 and 4.

2.0 INPUT UNCERTAINTY QUANTIFICATION

2.1 Physical Variability

Examples of cementitious barrier model input variables with physical variability (i.e., inherent, natural variability) include:

- Material properties (e.g., mechanical, thermal, porosity, permeability, diffusivity)
- Geometrical properties (e.g., structural dimensions, concrete cover depth)
- External conditions (e.g., mechanical loading, boundary conditions, physical processes such as freeze-thaw, chemical processes such as carbonation, chloride or sulfate attack)

Many uncertainty quantification studies have only focused on quantifying and propagating the inherent variability in the input parameters. Well-established statistical (both classical and Bayesian) methods are available for this purpose.

2.1.1 Modeling Variability in System Properties

In probabilistic analysis, the sample–to–sample variations (random variables) in the parameters are addressed by defining them as random variables with probability density functions (PDFs). This assumes that the system/material is homogeneous on a macroscale. For example, chloride ion diffusivity has been modeled using a lognormal distribution (Hong, 2000; Gulikers, 2006; Rafiq et al., 2004; Chen, 2006) and water–cement ratio has been modeled using a normal distribution (Chen, 2006) and uniform and triangular distributions (Kong et al., 2002).

Some parameters may vary not only from sample to sample (as is the case for random variables), but also in spatial or time domain. Parameter variation over time and space can be modeled as *random processes* or *random fields*. For example, concrete cover depth and compressive strength have been modeled as random fields using squared exponential correlation functions (Stewart and Mullard, 2007).

Some well known methods for simulating random processes are spectral representation (SR) (Gurley, 1997), Karhunen-Loeve expansion (KLE) (Ghanem and Spanos, 2003, Huang et al., 2007; Mathelin et al., 2005), and polynomial chaos expansion (PCE) (Huang et al., 2007; Mathelin et al., 2005; Red-Horse and Benjamin, 2004). The PCE method has been used to represent the stochastic model output as a function of stochastic inputs.

Consider an example of representing a random process using KLE, expressed as

$$\varpi(x,\chi) = \varpi(x) + \sum_{i=1}^{\infty} \sqrt{\lambda_i}\,\xi_i(\chi) f_i(x) \qquad (2)$$

where:

$\varpi(x)$ is the mean of the random process $\varpi(x, \chi)$, λ_i and $f_i(x)$ are eigenvalues and eigenfunctions of $C(x_1,x_2)$, and $\xi_i(\chi)$ is a set of uncorrelated standard normal random variables (x is a space or time coordinate, and χ is an index representing different realizations of the random process).

Using Equation (2), realizations of the random process $\varpi(x,\chi)$ can be easily simulated by generating samples of the random variables $\xi(\chi)$, and these realizations of $\varpi(x,\chi)$ can be used as inputs to PA.

2.1.2 Modeling Variability in External Conditions

Some boundary conditions (e.g., temperature and moisture content) might exhibit a recurring pattern over shorter periods and also a trend over longer periods. An example of variability in an external condition, i.e., rainfall, is illustrated in Figure 2. It is evident from the figure that the rainfall data has a pattern over a period of 1 year and a downward trend over a number of years. These can be numerically represented by a seasonal model using an autoregressive integrated moving average (ARIMA) method generally used for linear[1] nonstationary[2] processes (Box et al., 1994). This method can be used to predict the temperature or the rainfall magnitudes in the future so that it can be used in the durability analysis of the structures under future environmental conditions.

[1] The current observation can be expressed as a linear function of past observations.
[2] A process is said to be non-stationary if its probability structure varies with the time or space coordinate.

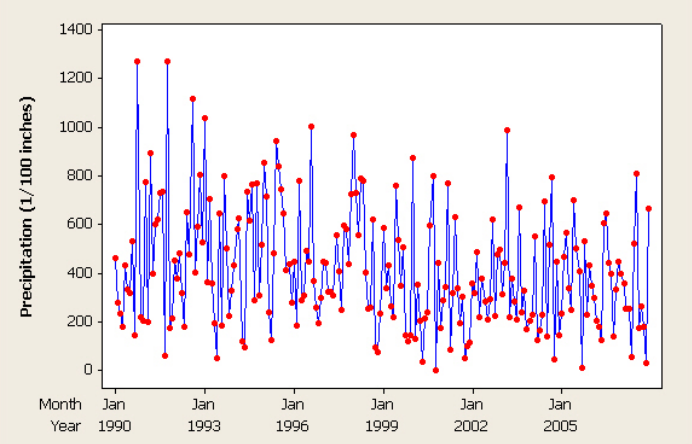

Figure 2. Precipitation Data for Aiken, SC (National Oceanic and Atmospheric Administration)

2.1.3 Stationary External Processes

For a stationary process[3], the ARIMA method expresses the observation at the t^{th} time step in terms of the observations at previous time steps as

$$z_t = c + \sum_{i=1}^{p} \phi_i z_{t-i} + \varepsilon_t \quad (3)$$

where:

z_t and z_{t-i} are observations at the t^{th} and $(t-i)^{th}$ time steps, c is a constant, ϕ_is are coefficients and ε_t is the error between the observed and the predicted values at t^{th} time step.

Assuming that the error at t^{th} time step is also dependent on the errors at previous time steps, ε_t can also be expressed as

$$\varepsilon_t = c_1 + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \quad (4)$$

where:

c_1 is a constant and θ_i's are coefficients.

Using a backward operator B such that $B^i z_t = z_{t-i}$ and combining Eqs. (3) and (4), results in Equation 5.

$$\phi_p(B) z_t = \theta_q(B) \varepsilon_t \quad (5)$$

where:

ϕ_p(B) and θ_q(B) are polynomials of pth and qth order. The coefficients of the polynomials can be determined using the least-squares method.

2.1.4 Non-Stationary External Processes

A random non-stationary process fluctuates about a mean value that exhibits a specific pattern. If the differences in levels of fluctuation are considered, the process can be simulated using the same method as for stationary processes. For example, differentiating a second order polynomial twice will result in a constant. Thus, a non-stationary process of d^{th} degree can be expressed as

$$\phi_p(B) \nabla^d z_t = \theta_q(B) \varepsilon_t \quad (6)$$

[3] A process is said to be stationary if its probability structure does not vary with the time or space coordinate.

where:

∇ is called the *backward difference operator* of the d^{th} degree.

If the process exhibits patterns over a shorter period (s) and a trend over a longer period, the process can be expressed as

$$\Phi_P(B^s)\nabla_s^D z_t = \Theta_Q(B^s)\varepsilon_t \quad (7)$$

where:

$\Phi_P(B^s)$ and $\Phi_Q(B^s)$ are polynomials of order P and Q, $B^s z_t = z_{t-s,}$ and D is the order of differentiation.

A similar model may be used to relate the current error (error between observation and model prediction at t^{th} time step) to the previous errors (errors between observations and model predictions at previous time steps) as

$$\varphi_p(B)\nabla^d \varepsilon_t = \theta_q(B)a_t \quad (8)$$

where:

$\varphi_p(B)$ and $\theta_q(B)$ and are polynomials of order p and q, d is the order of differentiation and a_t is a white noise process.

The final model is obtained by combining Eqs. (7) and (8) as

$$\varphi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D z_t = \theta_q(B)\Theta_Q(B^s)a_t \quad (9)$$

Eq. (9) is referred to as a *general multiplicative model* of order $(p \times d \times q) \times (P \times D \times Q)_s$. This method can be used to simulate a seasonal process.

It may also be important to quantify the statistical correlations between some of the input random variables. Many previous studies on uncertainty quantification simply assume either zero or full correlation, in the absence of adequate data. A Bayesian approach may be pursued for this purpose, as described in subsection 2.2.

2.2 Data Uncertainty

A Bayesian updating approach is described below to quantify uncertainty due to inadequate statistical data and measurement errors (ε_{exp}). This is consistent with the framework proposed in Figure 1, and is used to update the statistics of different physical variables and their distribution parameters. The prior distributions are based on available data and expert judgment, and these are updated as more data becomes available through experiments, analysis, or real-world experience.

2.2.1 Sparse Statistical Data

For any random variable that is quantitatively described by a probability density function, there is always uncertainty in the corresponding distribution parameters due to small sample size. As testing and data collection activities are performed, the state of knowledge regarding the uncertainty changes, and a Bayesian updating approach can be implemented. For example, suppose we decide that an input variable X follows a Gaussian distribution $N(\mu,\sigma^2)$ with μ and σ estimated from the data.

There is uncertainty in the normal distribution assumption, as well as in the estimates of the distribution parameters μ and σ, depending on the sample size. In the Bayesian approach, μ and σ are also treated as random variables, and their statistics are updated based on new data. However, we do not know the distribution of μ and σ *a priori*, so we may assume Gaussian for μ and Gamma distribution for $\phi = \sigma^{-2}$ as an initial guess for example, and then do a Bayesian update after more data is collected.

The Bayesian approach also applies to joint distributions of multiple random variables, which also helps to include the uncertainty in correlations between the

variables. A prior joint distribution is assumed (or individual distributions and correlations are assumed), and then updated as data becomes available.

Instead of assuming a well known prior distribution form (e.g., uniform, normal) for sparse data sets, either empirical distribution functions, or flexible families of distributions based on the data can be constructed. A bootstrapping[4] technique can then be used to quantify the uncertainty in the distribution parameters. The *empirical distribution function* is constructed by ranking the observations from lowest to highest value, and assigning a probability value to each observation.

Examples of flexible distribution families include the: Johnson family, Pearson family, gamma distribution, and stretched exponential distribution. The use of the Johnson family distribution has been explored by Marhadi et al., 2008, and extended to quantify the uncertainty in distribution parameters by McDonald et al., 2009. In constructing the Johnson family distribution, the available data is used to calculate the first four moments, and then the distribution form is chosen based on the values of the four moments. A jack-knife procedure is used to estimate the uncertainty in the distribution parameters, based on repeated estimation by leaving out one or more data points in each estimation.

2.2.2 Measurement Error

The measured quantity y_{exp} usually deviates from the unknown true value y_{true} due to the uncertainties in the test setup, equipment, environment, and operator. For example, large errors in the measurement of expansion due to sulfate attack can be seen in the experiments performed by Ferraris et al., 1997. The measurement error ε_{exp} can be expressed as $y_{exp} = y_{true} + \varepsilon_{exp}$. The measurement error in each input variable in many studies (e.g., Barford, 1985) is assumed to be independent and identically distributed (IID) with zero mean and an assumed variance, i.e., $\varepsilon_{exp} \sim N(0,\sigma^2_{exp})$. Due to the measurement uncertainty, the distribution parameter σ_{exp} cannot be obtained as a deterministic value. Instead, it is a random variable with a prior density $\tau(\sigma_{exp})$. Thus, when new data is available after testing, the distribution of σ_{exp} can be easily updated using the Bayes theorem.

Another way to represent measurement error ε_{exp} is through an interval only, and not as a random variable. In that case, one can only say the true value y_{true} lies in the interval $[y_{exp} - \varepsilon_{exp}, y_{exp} + \varepsilon_{exp}]$ without any probability distribution assigned to ε_{exp}. Methods to include data in interval format are discussed next.

2.2.3 Data Available in Interval Format

Some quantities in the system model may not have probabilistic representation, since data may be sparse or may be based on expert opinion. Some experts might only provide information about a range of possible values for some model input variable. Representations such as fuzzy sets, possibility theory, and evidence theory have been used. This report is focused on probabilistic methods to include interval data.

Transformations have been proposed from a non-probabilistic to probabilistic format, through the maximum likelihood approach (Langley, 2000; Ross et al., 2002). Such transformations have attracted the criticism that information is either added or lost in the process. Two ways to address the criticism are: (1) construct empirical distribution functions based on interval data collected from multiple experts or experiments (Ferson et al., 2007); or (2) construct flexible families of distributions with bounds on distribution

[4] Bootstrapping is a data-based simulation method for statistical inference by re-sampling from an existing data set (Efron et al., 1994).

parameters based on the interval data, without forcing a distribution assumption (McDonald et al., 2008). These can then be treated as random variables with probability distribution functions and combined with other random variables in a Bayesian framework to quantify the overall system model uncertainty. The use of families of distributions will result in multiple probability distributions for the output, representing the contributions of both physical variability and data uncertainty.

3.0 PROPAGATION UNCERTAINTY METHODS

In this section, methods to quantify the contributions of different sources of uncertainty and error as they propagate through the system analysis model, including the contribution of model error, are discussed, in order to quantify the overall uncertainty in the system model output.

This section will cover two issues: (1) quantification of model output uncertainty, given input uncertainty (both physical variability and data uncertainty), and (2) quantification of model error (due to both model form selection and solution approximations).

Several uncertainty analysis studies, including a study with respect to the Yucca Mountain high-level waste repository, have recognized the distinction between physical variability and data uncertainty (Helton and Sallaberry, 2009a & 2009b). As a result, these methods evaluate the variability in an inner loop calculation and data uncertainty in an outer loop calculation. Another example is provided by Holdren et al., 2006 in a baseline risk assessment study with respect to the Idaho Cleanup Project, where contributions of different sources of uncertainty are separately analyzed, such as from inventory, infiltration, sorption characteristics, model calibration, and simulation periods.

3.1 Propagation of Physical Variability

Various probabilistic methods (e.g., Monte Carlo simulation and first-order or second-order analytical approximations) have been studied for the propagation of physical variability in model inputs and model parameters, expressed through random variables and random process or fields. Stochastic finite element methods (e.g., Ghanem and Spanos, 2003; Haldar and Mahadevan, 2000) have been developed for single discipline problems in structural, thermal, and fluid mechanics. An example of such propagation is shown in Figure 3. Several types of combinations of system analysis model and statistical analysis techniques are available:

- Monte Carlo simulation with the deterministic system analysis as a black-box (e.g., Robert and Cesalla, 2004) to estimate model output statistics or probability of regulatory compliance;
- Monte Carlo simulation with a surrogate model to replace the deterministic system analysis model (e.g., Ghanem and Spanos, 2003; Isukapalli et al., 1998; Xiu and Karniadakis, 2003; Huang et al., 2007), to estimate model output statistics or probability of regulatory compliance;
- Local sensitivity analysis using finite difference, perturbation or adjoint analyses, leading to estimates of the first-order or second-order moments of the output (e.g., Blischke and Murthy, 2000); and
- Global sensitivity and effects analysis, and analysis of variance in the output (e.g., Box et al., 1978).

These techniques are generic, and can be applied to multi-physics analysis with multiple component modules as in the PA of cementitious barriers. However, most applications of these techniques have only considered physical variability. The techniques need to include the contribution of data uncertainty and model error to the overall model prediction uncertainty.

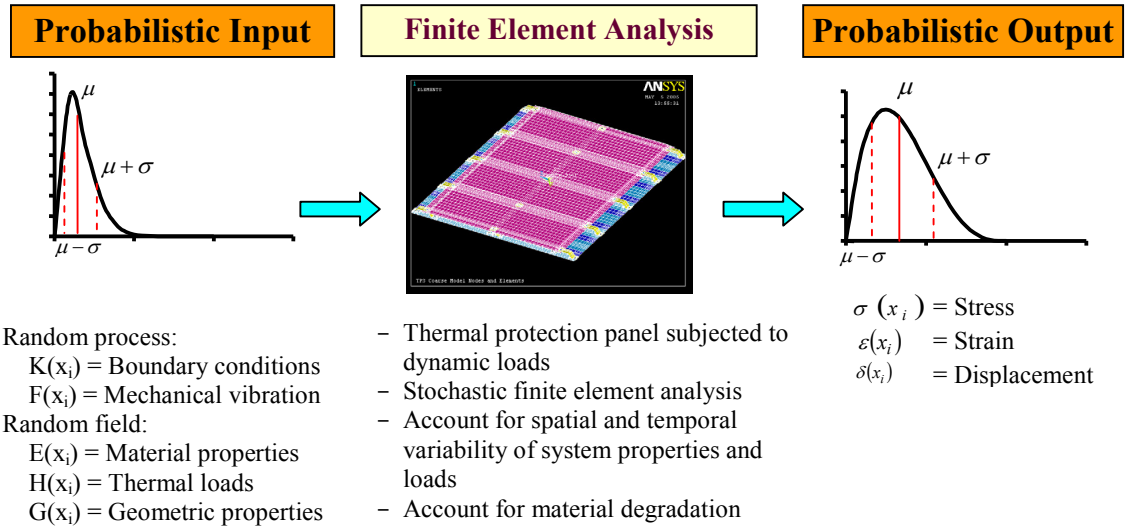

Figure 3. Example of Physical Variability Propagation

Computational effort is a significant issue in practical applications, since these techniques involve a number of repeated runs of the system analysis model. The system analysis may be replaced with an inexpensive surrogate model in order to achieve computational efficiency; this is discussed in Section 3.2. Efficient Monte Carlo techniques have also been pursued to reduce the number of system model runs, including *Latin hypercube sampling* (LHS) (Mckay et al., 1979; Farrar et al., 2003) and *importance sampling* (Mahadevan and Raghothamachar, 2000; Zou et al. 2003).

3.2 Propagation of Data Uncertainty

Three types of data uncertainty were discussed in Section 2. Sparse point data results in uncertainty about the parameters of the probability distributions describing quantities with physical variability. In that case, uncertainty propagation analysis takes a nested implementation. In the outer loop, samples of the distribution parameters are randomly generated, and for each set of sampled distribution parameter values, probabilistic propagation analysis is carried out as in Section 3.1. This results in the computation of multiple probability distributions of the output, or confidence intervals for the estimates of probability of non-compliance in PA.

In the case of measurement error, choice of the uncertainty propagation technique depends on how the measurement error is represented. If the measurement error is represented as a random variable, it is simply added to the measured quantity, which is also a random variable due to physical variability. Thus a sum of two random variables may be used to include both physical variability and measurement error in a quantity of interest. If the measurement error is represented as an interval, one way to implement probabilistic analysis is to represent the interval through families of distributions or upper and lower bounds on probability distributions, as discussed in Section 2.2.3. In that case, multiple probabilistic analyses, using the same nested approach as in the case of sparse data, can be employed to generate multiple output distributions or confidence intervals for the model output. The same approach is possible for interval variables that are only available as a range of values, as in the case of expert opinion.

Propagation of uncertainty is conceptually very simple, but computationally quite expensive to implement, especially when both physical variability and data uncertainty are to be considered. The presence of both types of uncertainty requires a nested implementation of uncertainty propagation analysis (simulation of data uncertainty in the outer loop and simulation of physical variability in the inner loop). If the system model runs are time-consuming, then uncertainty propagation analysis could be prohibitively expensive. One way to overcome the computational hurdle is to use an inexpensive surrogate model to replace the detailed system model, as discussed next.

3.3 Surrogate Models

Surrogate models (also known as response surface models) are frequently used to replace the expensive system model, and used for multiple simulations to quantify the uncertainty in the output. Many types of surrogate modeling methods are available, such as linear and nonlinear regression, polynomial chaos expansion, Gaussian process modeling (e.g., Kriging model), splines, moving least squares, support vector regression, relevance vector regression, neural nets, or even simple look-up tables. For example, Goktepe et al., 2006 used neural network and polynomial regression models to simulate expansion of concrete specimens under sulfate attack. All surrogate models require training or fitting data, collected by running the full-scale system model repeatedly for different sets of input variable values. Selecting the sets of input values is referred to as statistical design of experiments, and there is extensive literature on this subject. Two types of surrogate modeling methods are discussed below that might achieve computational efficiency while maintaining high accuracy in output-uncertainty quantification. The first method expresses the model output in terms of a series expansion of special polynomials such as Hermite polynomials, and is referred to as a *stochastic response surface method* (SRSM). The second method expresses the model output through a Gaussian process, and is referred to as *Gaussian process modeling*.

3.3.1 Stochastic Response Surface Method

The common approach for building a surrogate or response surface model is to use least squares fitting based on polynomials or other mathematical forms based on physical considerations. In SRSM, the response surface is constructed by approximating both the input and output random variables and fields through series expansions of standard random variables (e.g. Isukapalli et al., 1998; Xiu and Karniadakis, 2003; Huang et al., 2007). This approach has been shown to be efficient, stable, and convergent in several structural, thermal, and fluid flow problems. A general procedure for SRSM is as follows:

- Representation of random inputs (either random variables or random processes) in terms of Standard Random Variables (SRVs) by K-L expansion, as in Equation (2).

- Expression of model outputs in chaos series expansion. Once the inputs are expressed as functions of the selected SRVs, the output quantities can also be represented as functions of the same set of SRVs. If the SRVs are Gaussian, the output can be expressed a Hermite polynomial chaos series expansion in terms of Gaussian variables. If the SRVs are non-Gaussian, the output can be expressed by a general Askey chaos expansion in terms of non-Gaussian variables (Ghanem and Spanos, 2003).

- Estimation of the unknown coefficients in the series expansion. The improved probabilistic collocation method (Isukapalli et al., 1998) is used to minimize the residual in the random dimension by requiring the residual at the collocation points equal to zero. The model outputs are computed at a set of collocation points and used to estimate the coefficients. These collocation points are the

roots of the Hermite polynomial of a higher order. This way of selecting collocation points would capture points from regions of high probability (Tatang et al., 1997).

- Calculation of the statistics of the output that has been cast as a response surface in terms of a chaos expansion. The statistics of the response can be estimated with the response surface using either Monte Carlo simulation or analytical approximation.

3.3.2 Kriging or Gaussian Process Models

Gaussian process (GP) models have several features that make them attractive for use as surrogate models. The primary feature of interest is the ability of the model to "account for its own uncertainty." That is, each prediction obtained from a Gaussian process model also has an associated variance, or uncertainty. This prediction variance primarily depends on the closeness of the prediction location to the training data, but it is also related to the functional form of the response. For example, see Fig. 4, which depicts a one-dimensional Gaussian process model. Note how the uncertainty bounds are related to both the closeness to the training points, as well as the shape of the curve.

The basic idea of the GP model is that the output quantities are modeled as a group of multivariate normal random variables. A parametric covariance function is then constructed as a function of the inputs. The covariance function is based on the idea that when the inputs are close together, the correlation between the outputs will be high. As a result, the uncertainty associated with the model prediction is small for input values that are close to the training points, and large for input values that are not close to the training points. In addition, the GP model may incorporate a systematic trend function, such as a linear or quadratic regression of the inputs (in the notation of Gaussian process models, this is called the *mean function*, while in Kriging it is often called a *trend function*). The effect of the mean function on predictions that interpolate the training data is small, but when the model is used for extrapolation, the predictions will follow the mean function very closely.

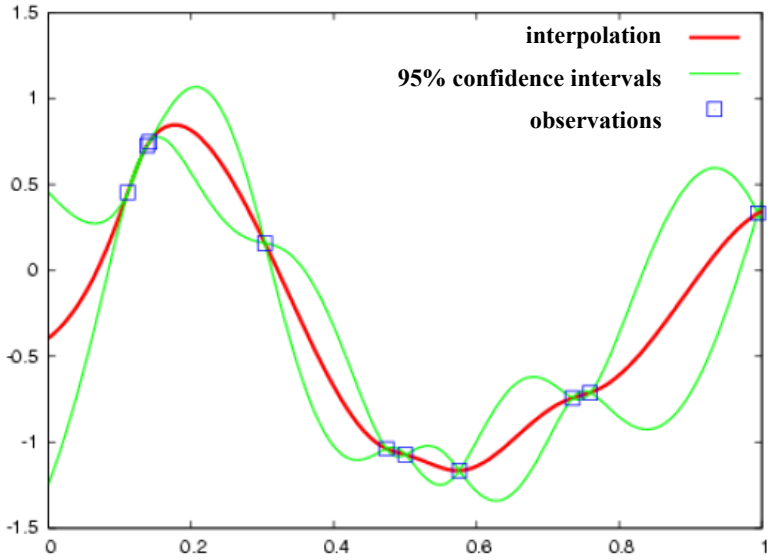

Figure 4. Gaussian Process Model With Uncertainty Bounds

Within the GP modeling technique, it is also possible to adaptively select the design of experiments to achieve very high accuracy. The method begins with an initial GP model built from a very small number of samples, and then one intelligently chooses where to generate subsequent samples to ensure the model is accurate in the vicinity of the region of interest. Since the GP model provides the expected value and variance of the output quantity, the next sample may be chosen in the region of highest variance, if the objective is to minimize the prediction variance. The method has been shown to be both accurate and computationally efficient for arbitrarily shaped functions (Bichon et al., 2007).

3.4 Sensitivity Analysis Methods

Sensitivity analysis serves several important functions: (1) identification of dominant variables or sub-models, thus helping to focus data collection resources efficiently; (2) identification of insignificant variables or sub-models of limited significance, helping to reduce the size of the problem and computational effort; and (3) quantification of the contribution of solution approximation error. Both local and global sensitivity analysis techniques are available to investigate the quantitative effect of different sources of variation (physical parameters, models, and measured data) on the variation of the model output. The primary benefit of sensitivity analysis to uncertainty analysis is to enable the identification of which physical parameters have the greatest influence on the output (Campolongo et al., 2000; Saltelli et al., 2000). An analysis of the impact of the parametric uncertainty is conducted to weed out those that have an insignificant effect upon the system output. For example, Chen (2006) performed sensitivity analysis to identify the important parameters affecting the service life of the concrete structures.

Three sensitivity analysis methods are factor screening, local-, and global-sensitivity analysis approaches. Factor screening determines which parameters have the greatest impact on the system output variability, by evaluating the output at the extreme values within the ranges of the parameters. Local sensitivity analysis utilizes first-order derivatives of system output quantities with respect to the parameters. It is usually performed for a nominal set of parameter values. Global sensitivity analysis typically uses statistical sampling methods, such as Latin Hypercube Sampling, to determine the total uncertainty in the system output and to apportion that uncertainty among the various parameters. Classical and Bayesian statistical analysis techniques, including the analysis of variance and differential sensitivity analysis, can be pursued to assess the global influence of an input parameter on an output variable by sampling from each input parameter's probability density function or from intervals of possible values.

3.5 Multi-Physics Models

In the past decade, different approaches have been proposed to quantify the uncertainty for individual physical models or simulation codes (e.g. see, Glimm and Sharp, 1999; Hanson, 1999; Devolder et al., 2002; Bae et al., 2003; Hanson and Hemez, 2003; Oberkampf et al., 2003; Millman et al., 2006; Witteveen and Bijl, 2006). For example, Hanson (1999) proposed a Bayesian probabilistic method for quantifying uncertainties in simulation predictions. Bae et al. (2003) used evidence theory to handle epistemic uncertainty about a structural system. Mathelin et al. (2004) and Witteveen and Bijl (2006) applied a polynomial chaos-based stochastic method for uncertainty propagation in numerical simulations. However, these existing approaches have not accounted for the uncertainty quantification in multiple modules of the system model, where the challenge is to combine data (available from different sources, in different formats) and model predictions regarding different physical phenomena (e.g., diffusion, chemical reaction, and mechanical damage), thus using all available information to quantify the overall prediction uncertainty. Urbina and Mahadevan (2009) have recently proposed a Bayes network approach to uncertainty quantification in multi-physics models.

3.6 Model Error Quantification

Model errors may relate to governing equations, boundary and initial condition assumptions, loading description, and approximations or errors in solution algorithms (e.g., truncation of higher order terms, finite element discretization, curve-fitting models for material damage such as S-N curve). Overall model error may be quantified by comparing model prediction and experimental observation, properly accounting for uncertainties in both. This overall error measure combines both model form and solution approximation errors, and so it needs to be considered in two parts. Numerical errors in the model prediction can be quantified first, using sensitivity analysis, uncertainty propagation analysis, discretization error quantification, and truncation (residual) error quantification. The measurement error in the input variables can be propagated to the prediction of the output. The error in the prediction of the output due to the measurement error in the input variables is approximated by using a first-order sensitivity analysis (Rebba et al., 2006). Then the model form error can be quantified based on all the above errors, following the approach illustrated for a heat transfer problem by Rebba et al. (2006).

3.6.1 Solution Approximation Error

Several components of prediction error, such as discretization error (denoted by ε_d) and uncertainty propagation analysis error (ε_s) can be considered. Several methods to quantify the discretization error in finite element analysis are available in the literature. However, most of these methods do not quantify the actual error; instead, they only quantify some indicator measures to facilitate adaptive mesh refinement. The Richardson extrapolation (RE) method comes closest to quantifying the actual discretization error (Richards, 1997). (In some applications, the model is run with different levels of resolution, until an acceptable level of accuracy is achieved; formal error quantification may not be required.)

Errors in uncertainty propagation analysis (ε_s) are method-dependent, i.e. sampling error occurs in Monte Carlo methods, and truncation error occurs in response surface methods (either conventional or polynomial chaos-based). For example, sampling error could be assumed to be a Gaussian random variable with zero mean and variance given by σ^2/N where N, is the number of Monte Carlo runs, and σ^2 is the original variance of the model output (Rubinstein, 1981). The truncation error is simply the residual error in the response surface.

Rebba et al. (2006) used the above concept to construct a surrogate model for finite element discretization error in structural analysis, using the stochastic response surface method. Gaussian process models may also be employed for this purpose. Both options are helpful in quantifying the solution approximation error.

3.6.2 Model Form Error

The overall prediction error is a combination of errors resulting from numerical solution approximations and model form selection. A simple way is to express the total observed error (difference between prediction and observation) as the sum of the following error sources:

$$\varepsilon_{obs} = \varepsilon_{num} + \varepsilon_{model} - \varepsilon_{exp} \tag{10}$$

where:

ε_{num}, ε_{model}, and ε_{exp} represent numerical solution error, model form error, and output measurement error, respectively.

However solution approximation error results from multiple sources and is probably a nonlinear combination of various errors such as discretization error, round-off and truncation errors, and stochastic analysis errors. One option is to construct a regression model consisting of the individual error components (Rebba et al., 2006).

The residual of such a regression analysis will include the model form error (after subtracting the experimental error effects). By denoting ε_{obs} as the difference between the data and prediction, i.e., $\varepsilon_{obs} = y_{exp} - y_{pred}$, we can construct the following relation by considering a few sources of numerical solution error (Rebba et al., 2006):

$$\varepsilon_{obs} = f(\varepsilon_h, \varepsilon_d, \varepsilon_s) + \varepsilon_{model} - \varepsilon_{exp} \quad (11)$$

where:

ε_h, ε_d, and ε_s represent output error due to input parameter measurement error, finite element discretization error, and uncertainty propagation analysis error, respectively, all of which contribute to numerical solution error.

Rebba et al. (2006) illustrated the estimation of model form error using the above concept for a one-dimensional heat conduction problem, assuming a linear form of Eq. (11). However, the function $f(\varepsilon_h, \varepsilon_d, \varepsilon_s)$ is nonlinear, and may be approximated through a response surface with respect to the three error variables, using a polynomial chaos expansion. The quantity $\varepsilon_{model} - \varepsilon_{exp}$ is simply the residual error of such a response surface. Thus the distribution of model error ε_{model} is quantified by knowing the distributions of residual error and measurement error.

Note that the above approach to quantifying model form error is only within the context of model validation—where actual data is available from targeted validation experiments—and compared with corresponding model predictions. In the context of PA, however, the concern is with extrapolation in time and space, and no direct comparison is possible between prediction and observation (at the time when the PA is done). Quantifying the model errors during extrapolation is difficult, and a Bayesian methodology might need to be pursued within restrictive assumptions (e.g., no change in physics). The Bayesian approach is discussed in Section 4.

4.0 MODEL CALIBRATION, VALIDATION AND EXTRAPOLATION

After quantifying and propagating the physical variability, data uncertainty, and model error for individual components of the overall system model, the probability of meeting performance requirements (and our confidence in the model prediction) needs to be assessed based on extrapolating the model to field conditions (which are uncertain as well), where sometimes very limited or no experimental data is available. Rigorous verification, validation, and calibration methods are needed to establish credibility in the modeling and simulation. Both classical and Bayesian statistical methodologies have been successfully developed during recent years for single physics problems, and have the potential to be extended to multi-physics models of cementitious barrier systems. The methods should have the capability to consider multiple output quantities or a single model output at different spatial and temporal points.

This section discusses methods for (1) calibration of model parameters, based on observation data; (2) validation assessment of the model, based on observation data; and (3) estimation of confidence in the extrapolation of model prediction from laboratory conditions to field conditions.

4.1 Model Calibration

Two types of statistical techniques may be pursued for model calibration uncertainty, the least squares approach, and the Bayesian approach. The least squares approach estimates the values of the calibration parameters that minimize the discrepancy between model prediction and experimental observation. This approach can also be used to calibrate surrogate models or low-fidelity models, based on high-fidelity runs, by treating the high-fidelity results similar to experimental data.

The second approach is Bayesian calibration (Kennedy and O'Hagan, 2001). This approach is flexible and allows different forms for the calibration factor, and it has been illustrated for a heat transfer example problem (McFarland and Mahadevan, 2007, McFarland, 2008).

In the literature, several researchers have calibrated their models using experimental results, especially if the phenomenon being modeled is complicated and the model is based on simplifying assumptions. For example, Tixier and Mobasher (2003) calibrated two parameters (reaction rate constant and fraction of porosity available for solid product deposition), and Krajcinovic et al. (1992) calibrated one parameter (reaction rate constant), while modeling the degradation of concrete structures under sulfate attack.

4.2 Model Validation

Model validation involves comparing prediction with observation data (either historical or experimental) when both have uncertainty. Since there is uncertainty in both model prediction and experimental observation, it is necessary to pursue rigorous statistical techniques to perform model validation assessment rather than simple graphical comparisons, provided data is even available for such comparisons. Statistical hypothesis testing is one approach to quantitative model validation under uncertainty, and both classic and Bayesian statistics have been explored. Classical hypothesis testing is a well-developed statistical method for accepting or rejecting a model based on an error statistic (see e.g., Trucano et al., 2001; Hills and Trucano, 2002; Paez and Urbina, 2002; Hills and Leslie, 2003; Rutherford and Dowding, 2003; Dowding et al., 2004; Chen et al., 2004; Oberkampf and Barone, 2006). Validation metrics have been investigated in recent years based on Bayesian hypothesis testing (Zhang and Mahadevan, 2003; Mahadevan and Rebba, 2005; Rebba and Mahadevan, 2006), reliability-based methods (Rebba and Mahadevan, 2008), and risk-based decision analysis (Jiang and Mahadevan, 2007 & 2008).

In Bayesian hypothesis testing, prior probabilities were assigned for the null and alternative hypotheses; $P(H_0)$ and $P(H_a)$ respectively, such that $P(H_0) + P(H_a) = 1$. Here H_0 : model error < allowable limit, and H_a: model error > allowable limit. When data D is obtained, the probabilities are updated as $P(H_0 \mid D)$ and $P(H_a \mid D)$ using the Bayes theorem. Then a Bayes factor (Jeffreys, 1961) B is defined as the ratio of likelihoods of observing D under H_0 and H_a; i.e., the first term in the square brackets on the right hand side of

$$\frac{P(H_0 \mid D)}{P(H_a \mid D)} = \left[\frac{P(D \mid H_0)}{P(D \mid H_a)}\right]\frac{P(H_0)}{P(H_a)} \tag{12a}$$

If $B > 1$, the data gives more support to H_0 than H_a. Also the confidence in H_0, based on the data, comes from the posterior null probability $P(H_0 \mid D)$, which can be rearranged from Eq. (12a) as

$$\frac{P(H_0)B}{P(H_0)B + 1 - P(H_0)} \tag{12b}$$

Typically, in the absence of prior knowledge, equal probabilities may be assigned to each hypothesis and thus $P(H_0) = P(H_a) = 0.5$. The posterior null probability can then be further simplified to $B/(B+1)$. Thus a B value of 1.0 represents 50% confidence in the null hypothesis being true.

The Bayesian hypothesis testing is also able to account for uncertainty in the distribution parameters, as mentioned in Section 2.2. For such problems, the validation metric (Bayes factor) itself becomes a random variable. In that case, the probability of the Bayes factor exceeding a specified value can be used as the decision criterion for model acceptance/rejection.

Notice that model validation only refers to the situation when controlled, target experiments are performed to evaluate model prediction, and both the model runs and experiments are done under the same set of input and boundary conditions. The validation is done only by comparing the outputs of the model and the experiment. Once the model is calibrated,

verified and validated, it may be investigated for confidence in extrapolating to field conditions different from laboratory conditions. This is discussed in the next section.

4.3 Confidence Assessment in Extrapolation

The Bayesian approach can also be used for assessing the confidence in extrapolating model prediction from laboratory conditions to field conditions, from lower resolution to higher resolution analysis, and from the lower level to the higher level in system analysis, through the construction of the Bayes network (Jensen and Jensen, 2001). Bayes networks are directed acyclic graphical representations with nodes to represent the random variables and arcs to show the conditional dependencies among the nodes. Data in any one node can be used to update the statistics of all other nodes. This property makes the Bayes network a powerful tool to extrapolate model confidence from laboratory conditions to field conditions (Mahadevan and Rebba, 2005). After computing the posterior distribution of the output under field conditions, through the Bayes network, the confidence in the prediction can be calculated similar to Section 4.2, using the Bayes factor.

Markov Chain Monte Carlo (MCMC) simulation is used for numerical implementation of the Bayesian updating analysis. Several efficient sampling techniques are available for MCMC, such as Gibbs sampling, the Metropolis algorithm, and the Metropolis-Hastings algorithm (Gilks et al., 1996).

Figure 5 shows an illustrative Bayes network for confidence extrapolation. An ellipse represents a random variable and a rectangle represents observed data. A solid line arrow represents a conditional probability link, and a dashed line arrow represents the link of a variable to its observed data if available. The probability densities of the variables Ω, z, and y are updated using the validated data **Y**. The updated statistics of Ω, z, and y are then used to estimate the updated statistics of the decision variable d (i.e., assessment metric). In addition, both model prediction and predictive experiments are related to input variables **X** via physical parameters Φ. Note that there is no observed data available for d; yet the confidence in the prediction of d, can be calculated by making use of observed data in several other nodes and propagation of posterior statistics through the Bayes network.

The Bayes network thus links the various simulation codes and corresponding experimental observations to facilitate two objectives: (1) uncertainty quantification and propagation and (2) extrapolation of confidence assessment from validation domain to application domain.

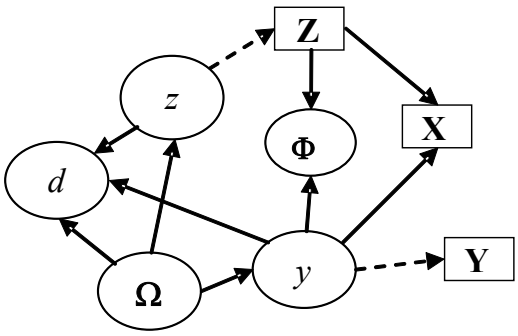

Figure 5. Bayes Network

5.0 PROBABILISTIC PERFORMANCE ASSESSMENT

Several methods are available in the reliability methods literature to efficiently perform probabilistic performance assessment, as fast alternatives to expensive Monte Carlo simulation. Performance assessment can be conducted with respect to single or multiple requirements. Efficient reliability analysis techniques that are based on first-order or second-order approximations or adaptive importance sampling can be used for this purpose. When multiple requirements are defined, computation of the overall probability of satisfying multiple performance criteria requires integration over a multidimensional space defined by unions and intersections of individual events (of satisfaction or violation of individual criteria).

An important observation here is that the same methods that are described here for reliability analysis can also be used to compute the cumulative distribution function (CDF) of the output, which may be of more general interest with respect to uncertainty quantification of model output. The term reliability analysis here refers only to computing the probability of exceeding or not meeting a single threshold value, which is a special case of constructing the entire CDF.

This section will discuss methods for probabilistic performance assessment with respect to individual criteria (5.1) and multiple criteria (5.2).

5.1 Individual Criteria

Probabilistic performance assessment can be based on the concept of a limit state that defines the boundary between success and failure for a system (Haldar and Mahadevan, 2000). The limit state function, g, is derived from a system performance criterion and formulated such that $g < 0$ indicates failure. If the input parameters in the system analysis are uncertain, so will be the predicted value of g. The probability of system failure, i.e. $P(g < 0)$ may be obtained from the volume integral under the joint probability density function of the input random variables over the failure domain as

$$P_f = \int \dots \int_{g \le 0} f_X(x_1, x_2, \dots, x_n)\ dx_1 dx_2 \dots dx_n \quad (13)$$

where:

P_f is the probability of failure, fx is the joint probability density of a random variable vector $\boldsymbol{X}$ with n elements; vector $\boldsymbol{x}$ represents a single realization of $\boldsymbol{X}$. Note that the integral is taken over the failure domain, or where $g \le 0$, so $P_f = P(g \le 0)$.

The basic Monte Carlo simulation method evaluates the above integral by drawing random samples from the distributions of the variables $\boldsymbol{X}$, and by evaluating whether $g \le 0$ in each run. Then the failure probability is simply the number of samples with $g \le 0$ divided by the total number of samples. While this technique is very simple to implement, it is also very expensive for problems with low failure probability.

The First Order Reliability Method (FORM) approximately estimates the failure probability as $P_f = \Phi(-\beta,)$ where β is the minimum distance from the origin to the limit state in the space of uncorrelated standard normal variables[5], as shown in Figure 6 (Hasofer and Lind, 1974). The minimum distance point on the limit state is referred to as the most probable point (MPP), and β is referred to as the reliability index. Finding the MPP is an optimization problem:

$$\text{Minimize } \beta = \|\eta\| \text{ , subject to } g_\eta(\eta) = 0 \quad (14)$$

[5] In general, a set of random variables x may be non-normal and correlated, but these may be transformed to an uncorrelated standard normal space (i.e. the space of random normal variables with 0 mean and unit standard deviation) via a transformation T, i.e $\eta = T(x)$.

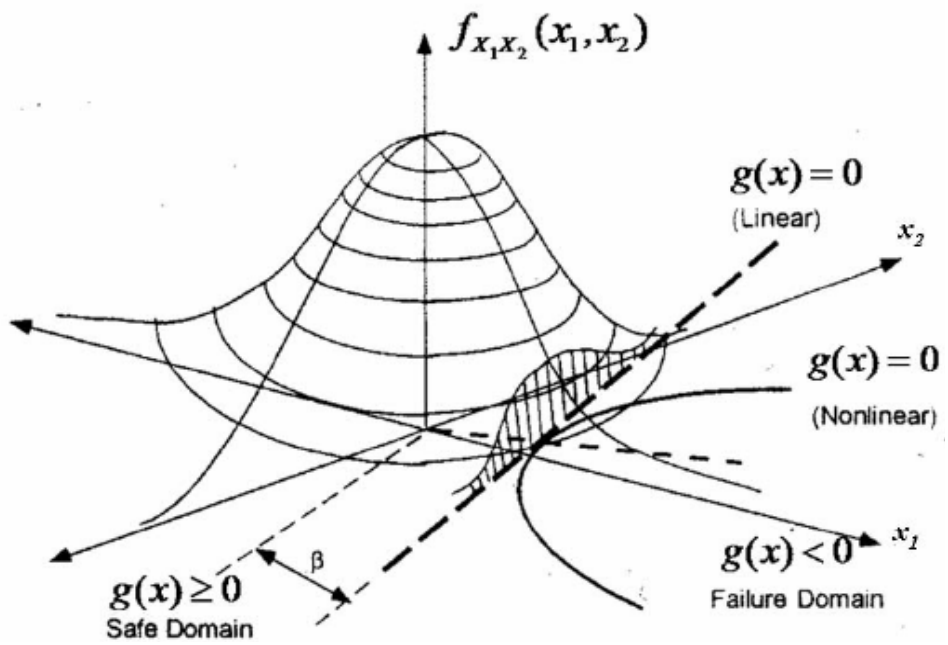

Figure 6. First-order Reliability Method

where:

$\boldsymbol{\eta}$ is the vector of random variables in the space of uncorrelated standard normal variables, and $||\eta||$ denotes the norm of that vector.

Several optimization techniques, such as Newton search (Rackwitz and Fiessler, 1978), and sequential quadratic programming (Schittkowski, 1983) can be used to find the MPP. Second-order reliability methods (SORM) are also available for higher accuracy; these take into account the curvature of the limit state in the failure probability calculation (e.g., Breitung, 1984; Tvedt, 1990). Compared to basic Monte Carlo simulation, FORM and SORM require many fewer iterations to converge to the MPP, and thus drastically reduce the computational expense.

5.2 Multiple Criteria

When a PA is conducted with respect to multiple requirements, the overall system-level probability of meeting the requirements is calculated through unions or intersections of individual failure probabilities.

In the case of unions (i.e., system fails if any one of the individual criteria is not met), the failure probability is

$$P_{F,Series} = P\{\bigcup_k g_k(\mathbf{x}) \leq 0\} \quad (15)$$

This system failure probability may be computed using either Monte Carlo simulation, or by extending the results of the first-order approximation in Section 5.1. Let **B** be the vector of reliability indices for each of the limit states, and the elements of the matrix **R** be the dot products of the corresponding **α** vectors (unit gradient vector of the limit state at the MPP in standard normal space) obtained from the FORM analysis for each limit state. Then the system failure probability in the above equation can be approximated as $1 - \Phi(\mathbf{B}, \mathbf{R})$, where $\Phi(\mathbf{B}, \mathbf{R})$ is the standard normal multivariate CDF with correlation matrix **R**. Closed-form representations of $\Phi(\mathbf{B}, \mathbf{R})$ exist for the

bivariate case (Dunnett and Sobel, 1954). If more than two limit states are considered, then one may elect to use bounding formulae (Ditlevsen, 1979), importance sampling methods (e.g., Mahadevan and Dey, 1998; Ambartzumian et al., 1998), multiple linearizations (Hohenbichler and Rackwitz, 1987), or a moment-based approximation (Pandey, 1998). For nonlinear limit states, the joint failure domain may be identified through an iterative linearization procedure (Mahadevan and Shi, 2001).

Similar concepts can be applied when the system failure is defined through intersections of individual failures (i.e., system fails only if all the individual criteria are not met). In that case, the failure probability is

$$P_{F,Parallel} = P\{\bigcap_k g_k(\mathbf{x}) \le 0\} \tag{16}$$

Again, the failure probability of the parallel system can be calculated either by Monte Carlo simulation, or from the results of the FORM analysis of its components as Φ (**-B**, **R**). In case FORM-based estimation is too approximate, Monte Carlo simulation can be used for higher accuracy, but with a large number of simulations. Efficient sampling techniques such as importance sampling (Mahadevan and Dey, 1998) may be used to reduce the computational expense.

In some cases, overall system failure definition may not be a simple union or intersection of individual failures, but may need to be represented as combinations of unions and intersections. In most cases, the system will not necessarily be in one of the two states (failed or safe), but in one of several levels of performance or degradation. Accounting for evolution of system states through time considerably increases the computational effort. The effort increases further when iterative multi-physics analysis is necessary, as in the case of several simultaneously active degradation processes. One option is to use first-order, second moment approximations to **B** and **R** (Mahadevan and Smith, 2006), to reduce the computational expense, but at the cost of accuracy. A trade-off between accuracy and computational expense may be necessary.

An important observation to note is that the probability calculations described in Sections 5.1 and 5.2 are only with respect to physical variability, represented by the random variables ***X***. The presence of data uncertainty and model errors makes the probability estimates themselves uncertain. Thus one can construct confidence bounds on the CDF of the output, based on a nested two-loop analysis. In the outer loop, realizations of the variables representing information uncertainty (such as distribution parameters of the probability distributions) and model errors are generated, and for each such realization, the output CDF is constructed in the inner loop. The collection of the resulting multiple CDFs is then used to construct the confidence bounds on the CDF. This nested implementation can become computationally demanding; in that case, a single loop implementation that simultaneously performs both outer loop and inner loop analyses may be pursued (McDonald et al., 2009).

6.0 CONCLUSION

Uncertainty quantification in performance assessment involves consideration of three sources of uncertainty – inherent variability, information uncertainty, and model errors. This report described available methods to quantify the uncertainty in model-based prediction due to each of these sources, and addressed them in four stages – input characterization based on data; propagation of uncertainties and errors through the system model; model calibration, validation and extrapolation; and performance assessment. Flexible distribution families were discussed to handle sparse data and interval data. Autoregressive models were discussed to handle time dependence. Methods to quantify model errors resulting from both model form selection and solution approximation were discussed. Bayesian methods were discussed for model calibration, validation and extrapolation. An important issue is computational expense, when iterative analysis

between multiple codes is necessary. Uncertainty quantification multiplies the computational effort of deterministic analysis by an order of magnitude. Therefore the use of surrogate models, and first-order approximations of overall output uncertainty, were described to reduce the computational expense.

Many of the methods described in the report have been applied to mechanical systems that are small in size, or time-independent, and the uncertainties considered were not very large. None of these simplifications is available in the case of long-term performance assessment of engineered barriers for radioactive waste containment, and real-world data to validate long-term model predictions is not available. Thus the extrapolations are based on laboratory data or limited term observations, and come with large uncertainty. Therefore the benefit of uncertainty quantification is not so much in predicting failure probability or similar measures, but in facilitating engineering decision making, such as comparing different design and analysis options, and allocating resources for uncertainty reduction through further data collection and/or model refinement.

7.0 REFERENCES

Ambartzumian, R, Der Kiureghian, A, Ohanian, V & Sukiasian, H 1997, 'Multinormal probability by sequential conditioned importance sampling.' In Advances in Safety and Reliability, Proceedings of ESREL '97, Vol. 2, 17–20 June, Lisbon, pp. 1261–1268.

Bae, H-R, Grandhi, RV & Canfield, RA 2003, 'Uncertainty quantification of structural response using evidence theory.' AIAA Journal, 41(10), 2062–2068.

Bichon, BJ, Eldred, MS, Swiler, LP, Mahadevan, S & McFarland, JM 2007. 'Multimodal Reliability Assessment for Complex Engineering Applications using Efficient Global Optimization,' Proceedings, 9th AIAA Non-Deterministic Approaches Conference, Waikiki, HI.

Barford, NC 1985, Experimental measurements: precision, error, and truth. Wiley, NewYork, 1985.

Blischke, WR & Murthy, DNP 2000, Reliability: Modeling, Prediction, and Optimization. Wiley, New York.

Breitung, K 1984, 'Asymptotic approximations for multinormal integrals.' Journal of Engineering Mechanics, 110(3), 357-366.

Box, GEP, Hunter, WG & Hunter JS 1978, Statistics for Experimenters, An Introduction to Design, Data Analysis, and Model Building. Wiley Interscience.

Box, GEP, Jenkins, GM & Reinsel GC 1994, Time series analysis forecasting and control (3rd edition). Prentice Hall, Englewood Cliffs, New Jersey,.

Campolongo, F, Saltelli, A, Sorensen, T & Tarantola, S 2000, 'Hitchhiker's guide to sensitivity analysis.' Sensitivity Analysis, A Saltelli, K Chan, and EM Scott. eds., John Wiley & Sons, pp. 15–47.

Chen, D 2006, Computational framework for durability assessment of reinforced concrete structures under coupled deterioration processes. Ph.D. Dissertation, Department of Civil and Environmental Engineering, Vanderbilt University.

Chen, W, Baghdasaryan, L, Buranathiti, T & Cao, J 2004, 'Model validation via uncertainty propagation and data transformation.' AIAA Journal, 42(7), 1406–1415.

Devolder, B, Glimm, J, Grove, JW, Kang, Y, Lee, Y, Pao, K, Sharp, DH & Ye, K 2002, 'Uncertainty quantification for multiscale simulations.' ASME Journal of Fluids Engineering, 124(1), 129 – 141.

Ditlevsen, O 1979, 'Narrow reliability bounds for structural systems.' Journal of Structural Mechanics, 7(4), 453–472.

Dowding, KJ, Hills, RG, Leslie, I, Pilch, M, Rutherford, BM & Hobbs, ML 2004, Case study for model validation: Assessing a model for thermal decomposition of polyurethane foam, Sandia National Laboratories Tec. Rep. Sand. No 2004-3632, Albuquerque, New Mexico, 2004.

Dunnett CW & Sobel, M 1954, 'A Bivariate generalization of student's t-distribution, with tables for certain special cases.' Biometrika, 41(1/2), 153–169.

Efron, B & Tibshirani, RJ 1994, An Introduction to the Bootstrap, Chapman & Hall/CRC.

Farrar, CR, Sohn, H, Hemez, FM, Anderson, MC, Bement, MT, Cornwell, PJ, Doebling, SW, Schultze, JF, Lieven, N & Robertson, AN 2003, Damage prognosis: Current status and future needs. Tec. Rep. LA–14051–MS, Los Alamos National Laboratory, Los Alamos, New Mexico.

Ferraris, CF, Clifton, JR, Stutzman, PE & Garboczi, EJ 1997, 'Mechanisms of degradation of Portland cement-based systems by sulfate attack.' Mechanisms of Chemical Degradation of Cement-Based Systems, London, 185-192, 1997.

Ferson, S, Kreinovich, V, Hajagos, J, Oberkampf W & Ginzburg, L 2007, Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty. Sandia National Laboratories Tec. Rep. Sand. No 2003-0939, Albuquerque, New Mexico.

Ghanem, R & Spanos, P 2003, Stochastic Finite Elements: A Spectral Approach, Springer-Verlag, New York.

Gilks, WR, Richardson, S & Spiegelhalter, DJ 1996, Markov Chain Monte Carlo in Practice, Interdisciplinary Statistics Series, Chapman and Hall, Boca Raton, Florida.

Glimm, J & Sharp, DH 1999, 'Prediction and the quantification of uncertainty.' Physica D, 133(1-4), 152–170.

Goktepe, AB, Inan, G, Ramyar, K & Sezer, A 2006. 'Estimation of sulfate expansion level of pc mortar using statistical and neural approaches', Construction and Building Materials, 20, 441–449.

Gulikers, J 2006, 'Considerations on the reliability of service life predictions using probabilistic approach. Journal de Physique IV, 136, 233-241, 2006.

Gurley, KR 1997, 'Modeling and simulation of non-Gaussian processes', PhD thesis, University of Notre Dame, April 1997.

Haldar, A & Mahadevan, S 2000, Probability, Reliability and Statistical Methods in Engineering Design, J. Wiley & Sons, New York.

Hanson, KM 1999, 'A framework for assessing uncertainties in simulation predictions.' Physica D, 133(1-4), 179–188.

Hanson, KM & Hemez, FM 2003, 'Uncertainty quantification of simulation codes based on experimental data.' Proceedings of 41st AIAA Aerospace Sciences, Jan. 6-9, 2003, Reno, Nevada.

Hasofer, AM & Lind, NC 1974, 'Exact and invariant second moment code format.' Journal of the Engineering Mechanics Division, ASCE, 100(EM1), 111-121.

Helton, JC & Sallabery, CJ 2009a, 'Conceptual basis for the definition and calculation of expected dose in performance assessments for the proposed high-level radioactive waste repository at Yucca Mountain, Nevada', Reliability Engineering and System Safety 94, 677- 698.

Helton, JC & Sallabery, CJ 2009b, 'Computational implementation of sampling-based approaches to the calculation of expected dose in performance assessments for the proposed high-level radioactive waste repository at Yucca Mountain, Nevada', Reliability Engineering and System Safety 94, 699-721.

Hills, RG & Leslie, I 2003, Statistical validation of engineering and scientific models: Validation experiments to application. Sandia National Laboratories Tec. Rep. Sand. No 2003-0706, Albuquerque, New Mexico.

Hills, RG & Trucano, TG 2002, Statistical validation of engineering and scientific models: A maximum likelihood based metric, Sandia National Laboratories Tec. Rep. Sand. No 2001-1783, Albuquerque, New Mexico.

Hohenbichler M & Rackwitz, R 1987, 'First-order concepts in systems reliability.' Structural Safety; 1(3), 177–188.

Holdren, KJ, Anderson, DL, Becker, BH, Hampton, NL, Koeppen, LD, Magnussen, SO & Sondrup, AJ 2006, Remedial investigation and baseline risk assessment of operable unit 7 13/14, U. S. Department of Energy, Idaho Operations Office.

Hong, HP 2000, Assessment of reliability of ageing reinforced concrete structures. Journal of Structural Engineering, 1260(12), 1458–1465.

Huang, S, Mahadevan, S & Rebba, R 2007, 'Collocation-based stochastic finite element analysis for random field problems', Probabilistic Engineering Mechanics, 22, 194–205, 2007.

Isukapalli, SS, Roy, A & Georgopoulos, PG 1998, 'Stochastic response surface methods (SRSMs) for uncertainty propagation: Application to environmental and biological systems.' Risk analysis, 18(3), 351–363.

Jeffreys, H 1961, Theory of Probability, 3rd. ed., Oxford University Press, London.

Jensen, FV & Jensen, FB 2001, Bayesian Networks and Decision Graphs, Springer-Verlag, New York.

Jiang, X & Mahadevan, S 2007, 'Bayesian risk-based decision method for model validation under uncertainty.' Reliability Engineering and System Safety, 92(6), 707–718.

Jiang, X and Mahadevan, S 2008, 'Bayesian validation assessment of multivariate computational models', Journal of Applied Statistics, 35(1), 49-65.

Kennedy, MC & O'Hagan, A 2001, Bayesian calibration of computer models (with discussion). Journal of the Royal Statistical Society, Series B, 63(3), 425–464.

Kong, JS, Ababneh, AN, Frangopol, DM & Xi, Y 2002, Reliability analysis of chloride penetration in saturated concrete. Probabilistic Engineering Mechanics, 17, 305–315.

Krajcinovic, D, Basista, M, Mallick, K & Sumarac, D 1992, Chemo-micromechanics of brittle solids. Journal of the Mechanics and Physics of solids, 40(5), 965-990.

Langley, RS 2000, 'A unified approach to the probabilistic and possibilistic analysis of uncertain systems', ASCE Journal of Engineering Mechanics, 126, 1163-1172.

Mahadevan, S & Dey, A 1998, 'Ductile system reliability analysis using adaptive importance sampling.' Structural Safety, 20(2), 137–154.

Mahadevan, S & Rebba, R 2005, 'Validation of reliability computational models using Bayes networks.' Reliability Engineering and System Safety, 87(2), 223–232.

Mahadevan, S & Raghothamachar, P 2000, 'Adaptive simulation for system reliability analysis of large structures.' Computers and Structures, 77(6), 725–734.

Mahadevan, S & Shi, P 2001, 'Multiple linearization method for nonlinear reliability analysis.' ASCE Journal of Engineering Mechanics, 127(11), 1165–1173.

Mahadevan, S & Smith, N 2006, 'Efficient first-order reliability analysis of multidisciplinary systems.' International Journal of Reliability and Safety, 1(1/2), 137–154.

Marhadi, KS, Venkataraman, S and Pai, S 2008, 'Quantifying uncertainty in statistical distribution of small sample data using Bayesian inference of unbounded Johnson distribution', Proceedings, 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Schaumburg, Illinois.

Mathelin, L, Hussaini, MY, Zhang, TA & Bataille, F 2004, 'Uncertainty propagation for a turbulent, compressible nozzle flow using stochastic methods.' AIAA Journal, 42(8), 1169–1176.

Mathelin, L, Hussaini, MY & Zang, TA 2005, 'Stochastic approaches to uncertainty quantification in CFD simulations', Numerical Algorithms, 38, 209–236.

McDonald, M, Zaman, K & Mahadevan, S 2008, Uncertainty quantification and propagation for multidisciplinary system analysis, 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Paper No. 134794, Victoria, British Columbia, Canada.

McDonald, M, Zaman, K & Mahadevan, S 2009, 'Representation and first-order approximations for propagation of aleatory and distribution parameter uncertainty.' AIAA-2009-2250, 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural dynamics, and Materials Conference, 4-7 May 2009, Palm Springs, California.

McFarland, JM 2008, Uncertainty analysis for computer simulations through validation and calibration. PhD dissertation, Vanderbilt University, Nashville, TN.

McFarland, J, Mahadevan, S, Swiler, L & Giunta, A 2007, 'Bayesian calibration of the QASPR simulation.' In Proceedings of the 9th AIAA Non-Deterministic Approaches Conference, Honolulu, Hawaii.

Mckay, MD, Conover, WJ & Beckman, RJ 1979. 'A comparison of three methods for selecting values of input variables in the analysis of output from a computer code.' Technometrics, 21, 239–245.

Millman, DR, King, PI, Maple, RC, Beranx, PS & Chiltonk, LK 2006, 'Uncertainty quantification with a B-spline stochastic projection.' AIAA journal, 44(8), 1845–1853.

Nigam, NC 1983, Introduction to random vibrations. MIT Press.

Oberkampf, WL & Barone, MF 2006, 'Measures of agreement between computation and experiment: Validation metrics.' Journal of Computational Physics 217(1), 5–36.

Oberkampf, WL, Trucano, TG & Hirsch, Ch 2003, Verification, Validation, and Predictive Capabilities in Computational Engineering and Physics, Sandia National Laboratories Tec. Rep. Sand. No 2003-3769, Albuquerque, New Mexico.

Paez, TL & Urbina, A 2002, Validation of mathematical models of complex structural dynamic systems, Proceedings of the Ninth International Congress on Sound and Vibration, Orlando, Florida.

Pandey, MD 1998, 'An effective approximation to evaluate multinormal integrals.' Structural Safety, 20(1), 51–67.

Rackwitz, R & Fiessler B 1978, 'Structural reliability under combined load sequences.' Computers and Structures, 9(5), 489–494.

Rafiq, MI, Chryssanthopoulos, MK & Onoufriou, T 2004, 'Performance updating of concrete bridges using proactive health monitoring methods', Reliability Engineering and System Safety, 86(3), 247-256.

Rebba, R 2005, Model Validation and Design under Uncertainty. PhD dissertation, Vanderbilt University, Nashville, TN, USA.

Rebba, R & Mahadevan, S 2006, 'Model predictive capability assessment under uncertainty', AIAA Journal, 44(10), 2376–2384.

Rebba, R & Mahadevan, S 2008, 'Computational methods for model reliability assessment', Reliability Engineering and System Safety, 93, 1197–1207.

Rebba, R, Mahadevan, S & Huang, S 2006, 'Validation and error estimation of computational models.' Reliability Engineering and Safety System, 91(10-11), 1390–1397.

Red-Horse, JR & Benjamin, AS 2004, 'A probabilistic approach to uncertainty quantification with limited information', Reliability Engineering and System Safety, 85, 183–190, 2004.

Robert, CP & Casella, G 2004, Monte Carlo Statistical Methods. 2nd ed. Springer-Verlag, New York.

Ross, TJ, Booker, JM & Parkinson, WJ 2002, Fuzzy Logic and Probability Applications: Bridging the Gap. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Richards, SA 1997, 'Completed Richardson extrapolation in space and time.' Communications in Numerical Methods in Engineering, 13(7), 558–573.

Rubinstein, RY 1981, Simulation and the Monte Carlo method. New York, Wiley.

Rutherford, BM, Dowding, K 2003, An approach to model validation and model-based prediction—polyurethane foam case study, Sandia National Laboratories, Tec. Rep. Sand. No 2003–2336, Albuquerque, New Mexico.

Saltelli, A, Chan, K & Scott, EM 2000, Sensitivity Analysis. John Wiley & Sons.

Schittkowski, K 1983, 'On the convergence of a sequential quadratic programming method with an augmented Lagrangian search direction.' Mathematische Operationsforschung und Statistik, Series Optimization, 14, 197–216.

Stewart, MG & Mullard, JA 2007, Spatial time-dependent reliability analysis of corrosion damage and the timing of first repair for RC structures. Engineering Structures, 29, 1457-1464.

Tatang, MA, Pan, W, Prinn, RG & McRae, GJ 1997, 'An efficient method for parametric uncertainty analysis of numerical geophysical models.' Journal of Geophysical Research, 102(D18), 21925–21932.

Tixier, R & Mobasher, B 2003, Modeling of damage in cement-based materials subjected to external sulfate attack. II: comparison with experiments. Journal of Materials in Civil Engineering, 15(4), 314-322.

Trucano, TG, Easterling, RG, Dowding, KJ, Paez, TL, Urbina, A, Romero, VJ, Rutherford, BM & Hills, RG 2001, Description of the Sandia validation metrics project, Sandia National Laboratories Tec. Rep. Sand. No 2001-1339, Albuquerque, New Mexico.

Tvedt, L 1990, 'Distribution of quadratic forms in normal space – application to structural reliability.' Journal of Engineering Mechanics, ASCE, 116(6), 1183-1197.

Urbina, A & Mahadevan, S 2009, Uncertainty quantification in hierarchical computational model development, Proceedings, 12th AIAA Non-Deterministic Approaches Conferences, Palm Springs, California.

Witteveen, J & Bijl, H 2006, 'Using polynomial chaos for uncertainty quantification in problems with Nonlinearities.' Paper No. AIAA-2006-2066, Proceedings, 47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, 14th AIAA/ASME/AHS Adaptive Structures Conference, Newport, Rhode Island.

Xiu, D & Karniadakis, GE 2003, 'Modeling uncertainty in flow simulations via generalized polynomial chaos.' Journal of Computational Physics, 187(1), 137-167.

Zhang, R & Mahadevan, S 2003, 'Bayesian methodology for reliability model acceptance.' Reliability Engineering and System Safety, 80(1), 95–103.

Zou, T, Mahadevan, S & Mourelatos, Z 2003, 'Reliability-based evaluation of automotive wind noise quality.' Reliability Engineering and System Safety, 82(2), 217–224.