Probability Bounds Analysis Applied to the Sandia Verification and Validation Challenge Problem

Aniruddha Choudhary[1]
National Energy Technology Laboratory
3610 Collins Ferry Road, Morgantown, WV 26505
aniruddhac@gmail.com
ASME Member

Ian T. Voyles
Aerospace and Ocean Engineering Department, Virginia Tech
460 Old Turner Street, Blacksburg, VA 24061
itvoyles@vt.edu
ASME Member

Christopher J. Roy
Aerospace and Ocean Engineering Department, Virginia Tech
460 Old Turner Street, Blacksburg, VA 24061
cjroy@vt.edu
ASME Member

William L. Oberkampf
W. L. Oberkampf Consulting
5112 Hidden Springs Trail, Georgetown, TX 78633
wloconsulting@gmail.com
ASME Member

Mayuresh Patil
Aerospace and Ocean Engineering Department, Virginia Tech
460 Old Turner Street, Blacksburg, VA 24061
mpatil@vt.edu

---

[1] Corresponding author

Abstract

Our approach to the Sandia Verification and Validation Challenge Problem is to use probability bounds analysis based on probabilistic representation for aleatory uncertainties and interval representation for (most) epistemic uncertainties. The nondeterministic model predictions thus take the form of p-boxes, or bounding cumulative distribution functions (CDFs) that contain all possible families of CDFs that could exist within the uncertainty bounds. The scarcity of experimental data provides little support for treatment of all uncertain inputs as purely aleatory uncertainties and also precludes significant calibration of the models. We instead seek to estimate the model form uncertainty at conditions where experimental data are available, then extrapolate this uncertainty to conditions were no data exist. The Modified Area Validation Metric (MAVM) is employed to estimate the model form uncertainty which is important because the model involves significant simplifications (both geometric and physical nature) of the true system. The results of verification and validation processes are treated as additional interval-based uncertainties applied to the nondeterministic model predictions based on which the failure prediction is made. Based on the method employed, we estimate the probability of failure to be as large as 0.0034, concluding that the tanks are unsafe.

Keywords: Sandia challenge problem, Verification and validation, Probability bounds analysis, Area Validation Metric, Prediction uncertainty, u-pooling

1   Introduction

As described in the Sandia Verification and Validation (V&V) workshop challenge problem [1][2] (herein referred to as the "challenge problem"), MysteryLiquid Co. is a company that maintains a large number of storage tanks for storing "Mystery Liquid" at various locations around the world. The tanks are cylinders with two hemispherical end caps as shown in Figure 1. At the junction where the cylindrical portion and the end caps meet, the tanks are supported by rings around the circumference. Locations on the tank surface are described by axial distance, x, measured from the central vertical plane, and circumferential angle, ϕ, measured from the vertical down as shown in Figure 1. During standard operation, the liquid level is limited to a certain fraction of the tank's height and the remaining space is filled with pressurized gas. During standard safety testing, one tank's measurements (out of many tanks) exceeded safety specification. This specification has been established from historical data, but is not a regulatory requirement. This out-of-spec tank (Tank 0) never physically failed. The out-of-spec tank and its two neighboring tanks were taken out of service to conduct further testing. Also, four tanks, in four different locations, underwent multiple tests while still in service. Data from these tests is provided to the challenge workshop participants. A computer program is also provided that does inexpensive function evaluations and is assumed to be a proxy for an expensive finite element model. All details regarding the challenge problem can be found in [1][2].

The objective of this analysis is to determine whether the tanks are at risk of failure and should they be replaced or can they remain in service for a few years while the replacements

are ordered. This recommendation by the participants of the challenge problem is to be based upon the calculation of probability of failure. The final decision by the company will be taken based upon the results from experimental testing, study of physical models, and the results of computational simulation of the physical model.

Our view of simulation-informed decision-making is that uncertainty in the simulation results should be clearly conveyed to the decision maker *without* incorporating questionable or debatable assumptions within the analysis. Stated differently, the individuals using the results of a verification, validation, and uncertainty quantification (VVUQ) analysis should clearly see the impact of all important uncertainties on the results of the simulation. Our approach differs dramatically from the traditional or common philosophy in uncertainty quantification (UQ) or risk assessment analyses. In the traditional approach, the analysis team makes seemingly reasonable and common assumptions concerning the conceptualization of the system of interest, numerical solution of the mathematical model, assimilation of experimental data, and then presents the simulation results to the decision maker. The assumptions and approximations may, or may not, be documented in footnotes or an appendix to a written report. Our philosophy is to make assumptions and approximations that are clearly defensible based on (a) information and experimental data on the system of interest and the conditions to which the system is exposed and (b) relevant experience and experimental data for closely related systems exposed to similar surroundings. If information is lacking concerning any aspect of the analysis, then the uncertainty in the information should be characterized such that no *additional* assumptions are incurred in the characterization of the uncertainty itself. The net

result of our philosophy is that the uncertainty in the predicted System Response Quantities (SRQs) will be larger, sometimes much larger, than a traditional approach. We argue that our philosophy is appropriate when uncertainty is primarily caused by lack of knowledge as opposed to random variability, for example in input data describing the system, the physics occurring in the system, and the conditions to which the system is exposed.

Some examples of uncertainties that are dominated by lack of knowledge, i.e., epistemic, uncertainty are: (a) spatial and temporal discretization error that is poorly estimated, (b) little or no experimental data exist for input parameters describing the system of interest, boundary conditions, or excitation of the system, (c) poor agreement between model predictions and experimental data on related systems, and (d) model predictions for physical conditions of interest that are far removed from the conditions where experimental data are available. Epistemic uncertainties such as these are not properly characterized as random variables because they have little relationship to random processes. As a result, mathematical structures such as intervals or Dempster-Shafer structures [3] should be used to forthrightly represent the actual state of knowledge to the decision maker. We present an approach based upon Probability Bounds Analysis (PBA) where aleatory (i.e., random) uncertainties in the input quantities are treated probabilistically as usual, but epistemic (i.e., lack of knowledge) uncertainties are treated as intervals. Our approach accounts for the effects of both types of uncertainties on the SRQs of interest, as well as model form and numerical uncertainty.

This paper is organized as follows. Before presenting our approach to the challenge problem, we make a few qualitative observations in Sec. 2 about the experimental and

computational description provided in the challenge problem statement. In Sec. 3, we provide a discussion of important concepts and terminology relevant to our approach followed by an overview of various steps in our analysis. In Sec. 4, we describe each step of our analysis in detail with results. Here, we present the final estimation of total uncertainty and the decision about tank safety. We end our response to the 2014 Sandia challenge problem in Sec. 5 with some concluding remarks.

2 Qualitative Examination of Problem

2.1 Experimental data

The quantities and types of available data (see Refs. [1][2]) for use in this problem introduce issues that could adversely affect the probability of failure estimate. Data is collected from only six tanks out of a total fleet of 450 tanks – only 1.33% of the population. Additionally, the data from each experiment is limited, especially from the experiments reporting the tank material properties and dimensions. While legacy data is available, no information is provided concerning the methods employed, tolerances allowed, or property distributions identified by the manufacturer. Measurement uncertainty assessments are available for some experiments, but not all. The main quantity of interest, the stress in the tank walls, is never measured and can only be inferred from other quantities.

Yield strength of the material, which is the primary quantity against which stress predictions from simulations are to be compared for reporting failure, is measured at only 10 locations and only using the out-of-spec tank. It is not known how well any data from a given

tank is correlated to the properties for another tank used in the experiment or to another tank within the fleet. No information is provided concerning the operating environments of the tank fleet. The effect that Mystery Liquid may have on a tank (e.g. corrosion or oxidation) along with the environmental effects such as damage due to a dry, sandy or a humid environment, are not provided, nor is it known that the entire fleet experiences the same host of effects. While it is stated that the tanks in the fleet range from 4 to 12 years old, the distribution of ages over the fleet and the age of any particular tank is not known, nor is there any information on tank fatigue due to use. The lack of knowledge concerning the tank environments and life cycle could potentially lead to large uncertainties in the prediction of the probability of failure.

2.2   Modeling and simulation

There are several known modeling and simulation (M&S) issues present that introduce uncertainty into a probability of failure prediction. A tank model has been provided as a Python code which acts a proxy for an expensive finite-element code with 4 available mesh levels. This model uses the series solution evaluation based upon the Timoshenko-Krieger shell theory for cylinders [4]. The finite element model provided considers only the straight, cylindrical portion of the tank and does not model the hemispherical tank end-caps. In addition, the model is incapable of accounting for non-uniformity in the tank dimensions, e.g., variation of tank wall thickness or any tank damage. Uncertainties in SRQs due to the modeling assumptions have not been previously studied and are not known. Additionally, the use of the finite element code introduces numerical uncertainties into the probability of failure prediction. Though the code is reported as having a first-order rate of convergence with consistent grid refinement on

previous problems, the present problem being considered is more complex than prior verification tests. Simulation is confined to only four meshes on which a given SRQ is not necessarily in the asymptotic range. Since our analysis explicitly includes an estimate of numerical uncertainty in the SRQs, the result is that the numerical uncertainty is increased because of lack of knowledge of these issues.

3 Methodology

An overview of our approach is provided in Sec. 3.5. However, first we present various concepts and terminology relevant to our approach in Secs. 3.1-3.4. Here, we also include a brief survey of relevant work from literature to emphasize the reasoning behind the selection of techniques employed in this work.

3.1 Uncertainty Framework

In order to quantify the total uncertainty in M&S predictions, it is important to identify and quantify all of the relevant uncertainty sources. The three main uncertainty sources are input uncertainty (also called parameter uncertainty), uncertainties due to the chosen form of the model (model form uncertainty), and uncertainties due to the numerical solution to the model (numerical uncertainty). Here we employ a broad approach to quantifying total uncertainty in M&S that accounts for all three of these sources [5][6].

The taxonomy employed for classifying uncertainties is based on their fundamental essence. The input or parameter uncertainty includes the uncertainty in the initial conditions, boundary conditions, material parameters, geometry, and system excitation. This uncertainty

may be purely aleatory (i.e., probabilistic), purely epistemic (arising from lack of knowledge), or a mixture of the two. Aleatory uncertainty is uncertainty due to inherent variation or randomness. Epistemic uncertainty (also called reducible uncertainty or ignorance uncertainty) is uncertainty that arises due to a lack of knowledge on the part of the analyst conducting the M&S. If knowledge is added (through experiments, improved numerical approximations, expert opinion, higher fidelity physics modeling, etc.) then the uncertainty can be reduced. If sufficient knowledge (which costs time and resources) is added, then the epistemic uncertainty can, in principle, be eliminated.

Aleatory uncertainty is typically characterized probabilistically with a precise probability density function or Cumulative Distribution Function (CDF). Epistemic uncertainty (and mixtures of epistemic and aleatory uncertainty) can be characterized in a number of ways including: probabilistically, as second-order probabilities (where the parameters governing the probability distribution are themselves uncertainty), as intervals, as Dempster-Shafer structures, and as fuzzy probabilities [7]. When one has very little knowledge about the value of a parameter, then an interval representation (with no associated probability distribution) is the weakest statement that one can make about the value of that parameter. Stated differently, every value in an interval-valued uncertainty can realize a probability of unity. This is not possible in precise probability theory, but it is allowed in imprecise probability theory. For cases where the parameter is known to be a random variable, but little information is available about its specific distribution, one may choose to characterize the variable using either a uniform distribution or

as a distribution with uncertain descriptive parameters. For this latter case, a mixed aleatory and epistemic characterization is appropriate.

One approach for characterizing mixed aleatory and epistemic uncertainty is a probability box (or p-box), which is similar to a CDF but with a finite width representing the epistemic uncertainty. The shapes of the two outer bounding CDFs reflect the aleatory uncertainty in the variable as seen in Figure 2. The width of the p-box represents the range of parameter values that are possible for a given cumulative probability level, whereas the height of the p-box represents the range of interval-valued cumulative probabilities associated with a given parameter value.

Unless otherwise noted, in this work, the input uncertainties are categorized as purely aleatory, purely epistemic, or mixed based on the available information and, in some cases, external information available in the literature. Purely aleatory uncertainty is characterized by a precise probability distribution. Purely epistemic uncertainty is characterized as an interval. Mixed uncertainty is characterized as imprecise probability distribution, i.e., precise probability distribution with interval-valued mean.

3.2   Uncertainty Propagation

When sufficient information and data are available such that all parametric uncertainties are aleatory, then well-established probabilistic methods for propagating uncertainty can be used (e.g., [10][11]). While a probabilistic treatment of epistemic uncertainty fits nicely within a Bayesian framework [12], it tends to under-predict the true uncertainty [9] and often has a strong (and undesirable) dependence on prior assumptions. For

example, when some information is available for an input parameter, a Bayesian approach would assume some precise probability distribution representing the degree of belief of the analyst (as opposed to actual evidence of frequency of occurrence) [8]. This has two detrimental effects with regard to capturing the actual poor state of knowledge. First, the result of the Bayesian analysis represents the individual belief of the analyst, as opposed to a result that is based on the limited information available. Second, when epistemic uncertainties are characterized as random variables, the final Bayesian result for the analysis is a *single* probability distribution as opposed to a set of distributions captured by a p-box. In fact, Beer et al. [7] point out that the posterior distribution becomes the prior as the available information goes to zero. On the other hand, a more general framework of information theory is described in [18] which can handle imprecise probabilities (i.e., aleatory, epistemic, and mixed uncertainty) and includes approaches such as probability bounds analysis [13] and evidence theory [3].

In order to propagate mixtures of aleatory and epistemic uncertainty, we employ a segregated approach to uncertainty propagation (see [6][16]). In the outer loop, samples from the interval-uncertain model inputs are drawn. For each of these sample values, the probabilistically uncertain model inputs are propagated using Latin Hypercube Sampling (LHS) along with the fixed sample values of the interval-uncertain variables. This propagation of the probabilistic uncertainty forms the inner loop of segregated uncertainty propagation. The result from one iteration of the outer loop is a possible CDF on the SRQ. The total result of the process is an ensemble of possible CDFs on the SRQ. In the limit of infinite samples, this process yields

the correct solution for problems in which there is no assumption of independence between the interval-uncertain inputs; but it is assumed that the interval-uncertain inputs are independent of the probabilistic inputs [9].

3.3   Model Form Uncertainty: Validation and Calibration

There is still a great deal of debate on how to treat model form uncertainty in M&S. One extreme is calibration (e.g., via Bayesian updating) which attempts to remove the model form uncertainty by using experimental data to improve the model [14][15]. When calibration is used to "remove" the model form uncertainty, issues often arise when applying the model outside the range of conditions where data are available (as is generally the case when making predictions). Another extreme is to use all experimental data to quantify the model form uncertainty, a process known as model validation [5][19][20]. There are also numerous approaches between these two extremes. For example, Kennedy and O'Hagan [21] developed a Bayesian calibration approach which includes a model discrepancy term. Roy and Oberkampf [6] allow a partitioning of the available experimental data into a sub-set to be used for calibration and another sub-set to be used for estimating model form uncertainty via comparison of nondeterministic model and experimental outcomes. Both authors provide an approach for extrapolating the model form uncertainty to conditions where experimental data are not available: Kennedy and O'Hagan [21] assume a Gaussian process while Roy and Oberkampf [6] employ a regression fit of the model form uncertainty along with prediction intervals.

3.4 Numerical Uncertainty

Since differential equation-based models rarely admit exact solutions for practical

problems, approximate numerical solutions must be used. The characterization of the

numerical approximation errors associated with a simulation is called verification [5][19][20]. It

includes discretization error, iterative convergence error, round-off error, and also errors due

to computer programming mistakes. The errors due to programming mistakes can usually be

found and eliminated using code verification practices. For cases where numerical

approximation errors can be estimated with a high degree of accuracy, their impact on the M&S

results can, in principle, be eliminated if sufficient computing resources are available. Since this

is often not practical, the uncertainties due to numerical errors should generally be converted

to epistemic uncertainties due to the uncertainties associated with the error estimation process

itself. Since numerical uncertainty is due to a lack of knowledge (i.e., epistemic), we treat it in

the same manner as model form uncertainty, i.e., as an interval about the simulation outcome.

3.5 Overview of the Analysis Process

Our analysis employed the following steps, which are discussed briefly here. The

implementation details of these steps are provided in the following sections.

1. Parametric study – The three control parameters, P, H, and χ, were varied over their

operating envelope to determine their qualitative effects on the maximum von Mises stress.

2. Sensitivity analyses – Global and local sensitivity analyses were performed to

determine which of the uncertain variables had the largest impact on the maximum (von Mises)

stress and maximum displacement. During these sensitivity studies, all uncertain variables were

treated as purely aleatory uncertainties, with epistemic uncertainties approximated as uniform probability distributions over the estimated interval ranges. The sensitivity analyses were used as a screening process to either a) omit certain insignificant variables from the UQ analysis or b) to treat epistemic and/or mixed uncertainties as precisely specified random uncertainties.

3. Numerical uncertainty estimation – Since no information was available regarding the code verification status of the software, we proceeded assuming that the code is free from any coding mistakes. In practice, rigorous code verification must be performed first for any of the downstream verification, validation, and uncertainty quantification activities to have any meaning. For this V&V challenge problem, the only numerical error source that was assumed to be present was discretization error, which was estimated with a combination of Richardson extrapolation and the Grid Convergence Index (GCI).

4. Uncertainty characterization – The input uncertainties were categorized as purely aleatory (A), purely epistemic (E), or mixed (M). Unless otherwise noted, we characterize the input uncertainties as follows: pure aleatory uncertainties are treated probabilistically, pure epistemic uncertainties are treated as intervals, and mixed aleatory and epistemic uncertainties are treated probabilistically with interval-valued means.

5. Uncertainty propagation – We employ segregated propagation of probabilistic and interval-characterized uncertainties using the Probability Bounds Analysis (PBA) framework. Interval uncertainties are sampled in the outer loop using Latin Hypercube Sampling (LHS). Probabilistic uncertainties are sampled in the inner loop using LHS.

6. Model form uncertainty – In this work, we do not employ model calibration. We instead utilize the available experimental data to estimate the model form uncertainty. Since no data are available for the primary SRQ, i.e. von Mises stress, we employ an approach called u-pooling [22][23] to convert experimental realizations for displacements to those for von Mises stress. (Note: u-pooling is discussed in detail with its implementation in Sec. 4.6). We then use the Modified Area Validation Metric (MAVM) to estimate model form uncertainty in von Mises stress [24][25]. Since this uncertainty is epistemic, we treat it as an interval about the simulation outcome. In our case, the simulation outcome is the p-box found by propagating the aleatory, epistemic, and mixed uncertainties through the model.

7. Estimation of total prediction uncertainty – The total prediction uncertainty in von Mises stress is determined by beginning with the p-box found from propagating all input uncertainties (aleatory, epistemic, and mixed) through the model. The estimated model form and numerical uncertainties are then appended as additional, independent, epistemic uncertainties on the bounds of the simulation p-box.

4   Results and Discussion

4.1   Parametric Study

As a first step towards qualitatively assessing the computational model available, a simple parametric study is performed by varying the three operating parameters over the operating range as follows:

- $P = [15,75]$ psig: 11 values at intervals of 6 psig

- $\chi = [0.1,1]$: 10 values at intervals of 0.1

- $H = [0,55]$ inches: 12 values at intervals of 5 in

All other input variables are kept constant at legacy values: $E = 3 \times 10^7$ psi, $v = 0.27$, $L = 60$ inches, $R = 30$ inches, and $T = 0.25$ inches. For this parametric analysis (and later, sensitivity analysis and uncertainty propagation study), the Dakota software toolkit [26] is used mainly due to its capability to perform large number of parallel runs of the model over multiple processors. Initially, when information about the computational cost was not known, the finest mesh (i.e., $m = 4$) was used for all $1320 = 11 \times 10 \times 12$ cases using the non-uniform mesh (option: resultStyle $= 2$) in the provided code. After information about the cost of computation on each mesh level was made available (as shown in Table 1), we recommend the use of the coarsest mesh ($m = 1$) for parametric analysis.

The output quantities from each of the 1320 runs include the maximum stress value and its location on the cylinder (i.e., axial, and angular locations and whether the maximum stress occurs on the internal or external surface). For brevity, only the key observations are discussed as follows:

- As expected, the maximum stress for any liquid composition always appears for the largest pressure and liquid height.

- Maximum stress varies most strongly with the liquid height in the container, less strongly with pressure and with liquid composition. A sample set of results are

shown in Figure 3 where the effect of varying pressure and liquid composition in the parametric space is shown for different values of liquid heights.

- Variation of maximum stress with liquid composition is not monotonic, i.e., it is large at $\chi = 0.1$, then decreases as χ increases until $\chi \approx 0.3$, and then increases from $\chi \approx 0.4$ to $\chi = 1.0$. The variation of maximum stress with pressure is almost always linear as expected.

- The tank has a circular cross section which results in slight nonlinearity in the variation of maximum stress with liquid height near 30-50 inches (i.e., when the tank is just over half-full). The cylindrical geometry of the tank also effects the angular location of maximum stress which moves up to approximately $80°$ (i.e., near the middle of the tank) when liquid height is approximately 35 inches and then falls down as the liquid height is further increased. This suggests that the middle region of the container is an important region for assessing maximum stress. A similar analysis of variation in axial location of maximum stress suggests that the regions near the support are important for failure analysis. The information about angular and axial locations of maximum stress suggests that the input option, $resultStyle = 2$, which reports results on a non-uniform grid (with refinement at the centerline and the supports of the tank) is suitable for this V&V study.

4.2   Sensitivity Analysis (SA)

Two sensitivity analyses are presented here at the maximum loading conditions (i.e.,

$H = 55$ inches, $P = 75$ psig, $\chi = 1$) with legacy values of the material and geometry

parameters. The first involves performing a global sensitivity analysis using the Variance-Based

Decomposition (VBD) method in Dakota [26]. The second involves a local sensitivity analysis

using simple finite-difference calculations. If resources are a constraint, we recommend the

local SA around the desired conditions. However, the global SA approach provides measures of

sensitivity in terms of main and total Sobol' indices giving an insight into uncertainty in the

output not only due to the uncertainty in the input variable but also due to the interactions

between the variables.

4.2.1   Global Sensitivity Analysis

VBD is used along with a sampling method (LHS). Since this SA was performed prior to

the characterization of input uncertainties, all input variables are treated as normal probability

distribution functions during this analysis. We used N=100 samples for each of the M=8 input

parameters as described in Table 2 for a total of $N \times (M + 2) = 1000$ runs. While sampling for

input parameters, it was found that the provided code has a calculation of arccosine function

using liquid height, H, and tank radius, R, as arccos$((H - R_1)/R_1)$, where R_1 has been created

by creating a bias/error in the input value of R as

$$R_1 = (1.5R - 17)0.992^{6/m}, \tag{1}$$

where, R is the tank radius, and $m = [1,2,3,4]$ is the meshID. This transformation shifts an R of

$28 - 32$ inches to an R_1 of $25 - 31$ inches. Since the arccosine function here cannot accept

18

$H > 2R_1$, this results in a rather strict (and complicated) constraint on the upper bound of H

and the lower bound of R. We use a height ratio to work around this issues during the sampling

process (discussed later). Another consideration that must be made while selecting the input

parameters is that the liquid composition must strictly follow, $0 \leq \chi \leq 1$. To deal with these

requirements, we have used truncated normal distribution functions which affect the mean and

SD (standard deviation) for H and R as seen in Table 2. The coefficient of variation (COV) and

mean values are determined based upon the limited experimental data available.

The VBD approach provides two primary measures to study how the uncertainty in

model output can be apportioned to uncertainty in input variables: the main effect and the

total effect sensitivity indices. The main effect sensitivity index, S_i, corresponds to the fraction

of the uncertainty in the output (e.g., Y) that can be attributed to an input (e.g., x_i) alone. The

total effects index corresponds to the fraction of the uncertainty in the output, Y, that can be

attributed to input, x_i, and its interactions with other input variables. Large values of S_i indicate

that the uncertainty in the input variable, x_i, has a large effect on the variance of the output.

Sum of the main effect indices from different input variable should be approximately equal to 1.

However, if the sum is significantly less than one, then there could be significant interactions

between input variables. Further details on interpretation of the Sobol' indices can be found in

[27].

Two different conditions (at $\chi = 1$ and $\chi = 0.1$) were explored, though the results of

only the more relevant case ($\chi = 1$) are presented here in Table 3 for brevity, where large

values of sensitivity indices are highlighted. Note that apart from the suggested SRQ of

maximum stress (σ_{max}), we also looked at the deflection at the maximum stress location ($w@\sigma_{max}$), and the maximum deflection (or normal displacement, w_{max}) in the simulation. Importance is given to both stress and deflection related SRQs in the current V&V study since the main quantity of interest for final failure prediction (i.e., σ_{max}) is never measured directly during the experiments and must be inferred from other properties such as deflection. Based on the results shown in Table 3, it is evident that maximum stress is most sensitive to uncertainties in tank thickness and radius values. While the deflection SRQs are sensitive to tank thickness and radius, they are most sensitive to uncertainties in Young's modulus of the material. The results from all the samples are presented pictorially in Figure 4 and Figure 5 where maximum stress and maximum deflection, respectively, are plotted against different input values. It can be seen from Figure 4(a)(b) and Figure 5(a)(b) that both SRQs are strongly correlated with tank radius and thickness. Figure 5(c) shows that the deflection SRQ is most strongly correlated with the Young's modulus of the material whereas the maximum stress has zero correlation with the Young's modulus as seen from Figure 4(c).

4.2.2   Local Sensitivity Analysis

A simple local sensitivity analysis is performed using one-sided finite differences. Here, the 8 input variables are perturbed one-at-a-time by 0.01% of the initial value at maximum loading conditions. Thus this analysis requires only a total of 9 runs which can be performed using results from the fine mesh (i.e., $m = 3$). As the inputs have a wide range of magnitudes, gradients are formed in a dimensionless manner. For an input x and the resulting SRQ, $Y(x)$, the gradient is calculated about the desired initial condition, x_0 as simply:

$$\frac{dY}{dx}\Big|_{x_0} \approx \frac{(Y(x_0 + \delta x) - Y(x_0))/Y(x_0)}{\delta x/x_0} \tag{2}$$

The results from this local sensitivity analysis are shown in Table 4 and largely support the conclusions derived from the earlier Sobol' indices computations, i.e., the examined SRQs are most sensitive to the wall thickness and tank radius, with the deflection SRQs also being sensitive to the Young's modulus. Based only on the finite-difference analysis, arguments could be made that variations in the liquid height and tank length are also moderately important, especially for the deflection SRQs.

4.3 Numerical Uncertainty: Grid Effects

To estimate numerical uncertainty, we examined three cases: two cases at $H = H_{max} = 55$ inches with (a) $\chi = 1$ and (b) $\chi = 0.1$, and one pressure loading only case (i.e. $H = 0$ inch). For each of the three cases, simulations were run on all four grid sizes (i.e., $m = 1,2,3,4$). Since the exact solution is unknown, the triplets of grids (1,2,3) and grids (2,3,4) (where 1 is the coarsest and 4 is the finest grid) were used to obtain observed orders of accuracy for each of the two SRQs (maximum stress and maximum displacement). The observed order with the finer grid triplet (2,3,4) always produced negative orders. For the coarser grid triplet (1,2,3), the observed order of accuracy of the maximum displacement was approximately 1.5 while the observed order for the maximum stress was near zero. For a real FEM solution, this erratic behavior for observed orders could be attributed to issues with systematic refinement of the finest grid or iterative non-convergence. For the current proxy FEM model, we believe that this

erratic behavior is the result of the tank radius transformation (discussed in Eq. 1) which may

not be well-posed with systematic refinement.

Since we were limited to the four meshes identified in the problem statement, we chose

to simply estimate the numerical error using a conservative factor of safety of three.

Furthermore, we decided to ignore the finest grid results and estimated the exact solution ($\bar{S}$)

using the fine ($m = 3$) and the medium ($m = 2$) grids as follows

$$\bar{S} = S_3 + \frac{S_3 - S_2}{r^p - 1} \tag{3}$$

where, S represents the simulation solution, r is the refinement factor between grids ($r = 2$ in

this study), and p is the observed order accuracy limited between, $0.5 \leq p \leq 1$ as

recommended in [5] (note that $p = 1$ is the stated formal order of accuracy of the FEM code for

this problem). The numerical uncertainty can then be then estimated for solution on a given

grid level, S_m, using a factor of safety of, $F_s = 3$, as follows:

$$U_{num}(\%) = F_s \left| \frac{\bar{S} - S_m}{\bar{S}} \right| \times 100 \tag{4}$$

Note that this procedure is similar to the Grid Convergence Index of [28], but slightly modified

to explicitly include the estimated exact solution which allows for the estimation of numerical

uncertainty in each of the coarser mesh levels. The final numerical uncertainties for $H =$

55 inches and $\chi = 1$ are given in Table 1. A crucial observation was that the maximum stress

always decreased monotonically with mesh refinement while the maximum displacement

always increased monotonically with grid refinement (numerical values not shown here for

brevity). Thus, one-sided numerical uncertainty intervals can be justifiably used about the

numerical solution as $[S - U_{num}, S]$ for maximum stress and $[S, S + U_{num}]$ for maximum

displacement. Here S is the simulation result (deterministic value, CDF, or p-box) and U_{num} is

the estimated numerical uncertainty. Based upon the uncertainty levels found in Table 1, we

propose using the medium mesh ($m = 2$) for the uncertainty propagation at nominal

conditions. This will require using the estimated U_{num} of $(-)24.1\%$ for maximum stress values

and $(+)5.5\%$ when maximum displacement SRQ is considered.

4.4   Uncertainty Characterization

4.4.1   Description and Characterization of Input Data

All input parameters can be divided into the categories of model input, system

excitation input (also called operating conditions), and numerical input parameters, as shown in

Table 5. Note that the liquid specific weight, γ, can be determined directly from a correlation

function dependent on the liquid composition, χ (with a relatively small error of $\pm 2\%$). Thus,

there is no need to discuss both χ and γ during this analysis.

Proper characterization of the input uncertainties is a crucial step in the current

probability bounds analysis framework. The three types of characterization (aleatory, epistemic,

and mixed) are suggested here based upon the nature of the property, information provided in

the problem statement, the parametric analysis of the input data, and the sensitivity analysis

conducted. An attempt was made to limit the number of quantities with interval-valued

23

characterization because epistemic and mixed uncertainties rapidly increase the number of

samples required during the propagation of uncertainties through the model. Thus, if the

sensitivity analysis showed that an input quantity has very little effect on the maximum stress

or the wall displacement, then that quantity is characterized as a precise probability

distribution using a normal distribution function. Although the results of SA matched well with

our physical understanding of the problem, it must be qualified here that these observations

were based upon SA performed before any model validation which can sometimes be

misleading.

The input uncertainty characterization is summarized in Table 6. The uncertainties are

characterized about the nominal condition which is defined as the condition at which the

ultimate failure prediction is to be made. The nominal condition for the control parameters is

described in the problem statement as: $P = 73.5$ psig, $H = 50$ inches, and $\chi = 1$. For the

material and geometry model input parameters, the legacy values are used as the nominal

values. The liquid composition, χ, and the liquid height, H, are characterized as interval-valued

uncertain parameters (i.e., epistemic) with precise lower and upper bounds. Note that the

upper bound of $\chi = 1$ is ensured here. Also, the liquid height is input into the code using a

user-defined normalized height parameter, $\psi = \frac{H}{2R_1}$, where R_1 is described in Eq. 1. The use of

ψ to define liquid height ensures the strict mathematical constraint of $H_{max} \leq 2R_1$ during the

sampling process and avoids the otherwise fatal segmentation fault within the

$\arccos((H - R_1)/R_1)$ calculation in the code. The lower and upper bounds of H are set based

upon tank levelling concerns as $\Delta H = \pm 1$ inch, while those for χ were set as $\Delta \chi = \pm 0.05$ based upon the information in the problem statement.

The gauge pressure, P, the Poisson's ratio, ν, and the tank length, L have been characterized as aleatory uncertainties using normal (or Gaussian) probability distribution functions with known mean (μ) and standard deviation (SD). The mean and standard deviation (or equivalently, coefficient of variation defined as $COV = SD/\mu$) have been selected based upon the available experimental data. From Table 6 it can be seen that the COV for P is selected as 2.5%, for ν as 4%, and for L as 1%. Given the small quantity of experimental data available for these input properties, we selected the mean as the legacy value (rather than the mean of the experimental data) and COV as 3.5 times the coefficient of variation from the experimental data to encompass ~95% of the available data points.

The mixed uncertainty characterization is employed for the following three model input parameters: Young's modulus, E, tank surface thickness, T, and tank radius, R. Here the mean is represented as an uncertain interval-valued parameter with known upper and lower bounds, however for each mean value a normal probability distribution is assumed with a selected standard deviation. The upper and lower bounds for the mean and the standard deviation for the probabilistic distribution are selected after observing the distribution of the (limited) experimental data provided.

4.4.2   Description of System Response Quantities

For a sampled point in the 8-dimensional input space, we can obtain a solution of the mathematical model. While the failure prediction is to be made based upon the von Mises stress response from the computational model, the only experimentally measured SRQ is the normal (radial) displacement of the tank wall under various loading conditions. Data set #5 (see [1][2] for details on various Data sets) provides measured displacements for Tanks 1 and 2, for three pressure-only loading conditions at 4 locations on each tank. Measurements on each tank were repeated in a second experiment, so we have two independent sets of measurements. Data set #6 contains measured displacements for Tanks 3 – 6 for various combinations of pressure, liquid height, and Chi, at 20 locations on each tank. Measurements on each tank were repeated three times, so three independent sets of measurements are available on each tank. One could argue that these data could also be used for computing a validation metric, or for calibration of input parameters. However, we decided to not use these data for both these purposes since failure prediction is to be based upon the stress SRQ and not displacements. Instead we choose to use the experimental realizations of the displacements at the provided locations to estimate the von Mises stress using u-pooling in order to estimate the model form uncertainty.

The given mathematical model provides a local value of the von Mises stress, σ, over the entire tank surface which can be used to find the maximum value of stress, σ_{max}, under a given loading condition. The failure criterion in this problem is the requirement that σ_{max} does not exceed the yield stress, σ_Y, by a probability of 10^{-3}. This criterion can be stated as

$$P(\sigma_{max} > \sigma_Y) < 10^{-3} \tag{5}$$

A key issue in interpreting the failure criterion is the characterization of yield stress. Yield stress

varies from unit-to-unit during manufacturing, depends upon operational history,

environmental conditions, and also varies locally throughout the volume of a solid. However,

experimental measurements for σ_Y are available at only 10 locations and only on Tank 0. This

lack of knowledge about yield stress, which is a key property for determining failure, is a major

challenge and its proper treatment is addressed during the final prediction. Note that the

uncertainty in the von Mises yield criterion itself should not be significant and gets addressed

through the uncertainty in yield stress.

4.5   Uncertainty Propagation

Input uncertainties are propagated through the model using segregated propagation of

uncertainties as implemented in the Dakota toolkit [26]. Here, the epistemic uncertainties are

sampled in the outer loop (number of samples=M) and aleatory uncertainties are sampled

(number of samples=N) within the inner loop. This results in a total of $M \times N$ total samples

(and thus total number of simulations) within the parametric space of 8 input uncertainties. The

standard Latin Hypercube Sampling (LHS) is used at all steps with the characterization of input

uncertainties as described earlier and in Table 6. For the final uncertainty propagation at

nominal conditions and for the u-pooling study (discussed in next section), the number of

samples were selected as $M = 50$ for the epistemic variables and $N = 200$ for the aleatory

variables. The results of uncertainty propagation performed at nominal conditions are shown in

Figure 6, where the left figure shows all 50 CDFs for the maximum stress. The outer bounds of the ensemble of CDFs result in a p-box which is shown in the right figure. Note that no uncertainty structure is assumed within this p-box.

4.6 Model Form Uncertainty

4.6.1 Obtaining Stress Data

Tanks 3, 4, 5, and 6 were field tested and 3 sets of experimental results are available for each tank at different loading conditions. For each of the resulting 12 experiments, displacement measurements are available at the 20 location as identified in Figure 7. As mentioned earlier, the experimental data is only available for displacements while a prediction is to be made for von Mises stress. To address this concern, we use a surrogate model developed from the simulations in order to correlate the measured displacements with stress components at the locations provided. Given the complexity of the problem, there is no direct correlation between stress and displacement. However, assuming that the system response can be represented by a single mode-shape (i.e., the functional form of the solution is known but the amplitude is unknown), then we can find exact correlation between the two quantities. Furthermore, we find that fitting the displacements from the simulations to the stress components resulted in nearly exact fit between the displacements and the von Mises stress. The axial stress (σ_x), the tangential stress (σ_ϕ), and the cross stress ($\sigma_{x\phi}$) terms can be combined to obtain the "surrogate" von Mises stress (σ) at each of these locations within the domain. This approach is somewhat similar to a Proper Orthogonal Decomposition analysis of the most important "mode shapes".

In order to create the surrogate model that could convert displacements into stress, we find the coefficients that correlate each of stress components at a few locations (numbers 1, 5, 8, 16 and 20) on the tank surface with their respective displacements obtained from a parametric run over the operating condition space (P, χ, H). A total of $125 = 5 \times 5 \times 5$ runs were made on the fine mesh ($m = 3$) to obtain the coefficients of the displacement-stress surrogate model.

4.6.2 u-pooling

The goal here is to employ the MAVM to obtain the model form uncertainty in the predicted stress which requires multiple experimental data points for each set of conditions. However, for each set of experimental conditions only 1 data point for stress is available at several locations on the tank. The method of u-pooling [23] is used here to transfer the 5 stress values from the 5 locations (numbers 16, 17, 18, 19 and 20) to one location (number 16) so that MAVM can be employed using this u-pooled "experimental" stress distribution. The key assumption for the use of u-pooling is that the effects of input uncertainties on these 5 different stress locations are correlated. Essentially, the u-pooling process is used here to transform the surrogate stress data to probability space by performing uncertainty propagation (while treating the interval uncertainties as uniform probabilities) at the experimental operating conditions. Next, the probability values (u-values) obtained from different experiments are all pooled to one location at nominal conditions. For example, we used case52 and case53 (see Table 7) to pool the data to $H = 51$ inches condition as follows:

- For conditions of case52 ($P = 63.2$ psig, $\chi = 0.7, H = 51$ inches), 10000 samples (M=50, N=200) were selected over the input parameter space and nondeterministic simulations were performed. The CDFs for the 5 von Mises stress values, called s_1, s_2, s_3, s_4, s_5, at the 5 extreme locations, numbered 16 to 20, respectively, are probed to find the corresponding probabilities (u-values) for the 5 stress values obtained from the surrogate model. Note that the "surrogate" stress data was obtained earlier using the displacement data from this specific experiment.

- Similarly, 10000 samples were run for case53 ($P = 64.6$ psig, $\chi = 0.4, H = 54$ inches) and after probing the CDFs, 5 u-values are obtained for the 5 surrogate stress data for this specific experiment.

- Another non-deterministic run of 10000 samples is performed for the operating conditions where we want to pool the data (i.e., $P = 73.5$ psig, $\chi = 1.0, H = 51$ inches). The 10 u-values (5 from each of the two experiments) are then used to probe the CDF for the stress at location 1 to obtain 10 stress values which we refer to as "experimental" stress data.

The experimental stress data can now be used to estimate the area validation metrics at $H = 51$ inches. For this case, the 10 surrogate stress data points, the corresponding u-values, and the experimental stress data are tabulated in Table 8. Figure 8 shows the process pictorially, where in (a) to (e), the probability levels are collected from each of the 5 stress values at the 5

locations (shown for case52 only), while in (f), the 10 cumulative probabilities from both cases are used to probe for the 10 experimental stress values.

4.6.3 Modified Area Validation Metrics

Area validation metrics employ the area between CDFs as the basis for the formation of a metric quantifying the disagreement between simulation and experimental outcomes [5][23]. The metric d is defined for the area validation metric as

$$d(F, S_n) = \int_{-\infty}^{\infty} |F(Y) - S_n(Y)| dY \tag{6}$$

where, $F(Y)$ is the p-box from the simulations, $S_n(Y)$ is the empirical CDF from the experiments, and Y is the SRQ of interest. Within the MAVM employed here [25] which accounts for regions in the cumulative probability space where the experimental values are larger than the simulation (d^+) and where the experimental values are smaller than the simulation values (d^-) (Please refer to [24][25] for a complete discussion on the process of MAVM evaluation). Once these metrics are computed, the model form uncertainty is constructed as the following interval around the simulation p-box

$$[F(Y) + \left(\frac{1 - F_s}{2}\right) d^+ - \left(\frac{1 + F_s}{2}\right) d^-, F(Y) + \left(\frac{1 + F_s}{2}\right) d^+ - \left(\frac{1 - F_s}{2}\right) d^-] \tag{7}$$

where F_s is a factor of safety based on the number of experimental samples available. F_s is determined here as

$$F_s(k) = F_1 + 1.2 \frac{F_0 - F_1}{k^{1/3}} \tag{8}$$

where, k is the number of available experimental samples, $F_1 = 1.25$, and $F_0 = 4.0$. For

additional conservativeness, the "average area" definition given in Voyles and Roy [24][25] is

used. Note that another formulation for F_s can be found in [25] based on confidence intervals.

After employing u-pooling, the new "experimental" CDF is compared with the

simulation p-box (from Figure 6) as shown in Figure 9. Using the u-pooled case at $H = 51$ in,

the MAVM metrics d^+ and d^- employing the average area are found as

$$d^+ = 2210 \text{ psi}; \text{ and } d^- = 20.7 \text{ psi} \tag{9}$$

Here, the factor of safety for 10 available experimental samples was found to be $F_s = 2.72$

from Eq. 8. Thus, from Eq. 7 the final model form uncertainty is estimated to be

$$[F(s_1) - 1878, F(s_1) + 4067] \text{ psi} \tag{10}$$

where $F(s_1)$ represents the p-box in this analysis. A similar method can be employed to obtain

the MAVMs at different liquid heights which can then be interpolated to obtain model form

uncertainty at any liquid height. However, we use the MAVM from Eq. 10 (i.e., for $H = 51$

inches) as the final model form uncertainty for total uncertainty estimation as an

approximation to that at the nominal condition ($H = 50$ inches).

4.7 Total Prediction Uncertainty

The total prediction uncertainty is determined by appending the numerical uncertainty which was estimated in Sec. 4.3 and the estimated model form uncertainty from Sec. 4.6 to the p-box obtained from the segregated propagation performed at nominal conditions in Sec. 4.5. The resulting total prediction p-box is shown in Figure 10. Note that the numerical uncertainty is appended only to the left side of the p-box since maximum stress was found to be monotonically decreasing with mesh refinement. Numerical uncertainty remains the largest source of uncertainty since the segregated propagation was performed with the medium grid ($m = 2$). The model form uncertainty is also a large contribution. Although the increase in predictive uncertainty due to numerical and model form uncertainties is large compared to the contribution of the input uncertainties, it is not claimed that the total uncertainty after appending these contributing uncertainties is a certain bound.

To evaluate the maximum probability of failure at the nominal conditions, the rightmost CDF of the total prediction p-box is fitted with a probability distribution. After comparing several distributions, it was found that the 3-point log-normal distribution with the parameters, $\mu = 9.8296$, $SD = 0.13876$, and threshold parameter, $\lambda = 12404$ provides the best fit based upon the Kolmogorov-Smirnov goodness of fit statistic. Similarly, another 3-point log-normal distribution is fitted to the leftmost CDF of the total prediction p-box using the parameters, $\mu = 10.631$, $SD = 0.03996$, and $\lambda = -26315$ for determining the minimum probability of failure. The fitted distributions for prediction p-box are shown in Figure 11 using solid curves.

The treatment of yield stress is more challenging because the small sample size results in a poor fit for most probability distributions. This is a case of lack of enough knowledge about a variable that is known to be random and is best characterized as a mixed uncertainty. A normal distribution is assumed for the yield stress with the standard deviation calculated from the provided sample as, $SD = 1755.857$, while the mean value is an uncertain interval. The lower and upper bounds for the mean of the normal distribution are selected as the sample mean and the legacy value for yield stress, respectively, resulting in the interval, $\mu = [44203.8, 45000]$. The resulting p-box for yield stress is plotted as the probability of exceedance (also called Complementary CDF (CCDF) or survival function) in Figure 11 using dotted curves. CCDF is simply defined as: $CCDF(\sigma) = 1 - CDF(\sigma)$.

In this case, the maximum probability of failure is determined by the complement of cumulative probability level where the rightmost prediction CDF exceeds the leftmost CCDF. This point of intersection is shown in Figure 11 using a solid circle (right). Similarly, minimum probability of failure is determined by the complement of cumulative probability level where the leftmost prediction CDF exceeds the rightmost CCDF. This value is found to be zero (within machine precision). Our final conclusions regarding failure prediction are as follows:

- The probability of failure is found as: $P_{fail} = [0, 0.0034]$.

- Based upon the available data on yield stress and defined failure criterion, we conclude that the tanks are not safe.

The final conclusion strongly depends upon the proper characterization of yield stress. However, detailed knowledge of yield stress from experimental measurements has not been provided in the problem statement. We recommend that accurate interval bounds as well as probabilistic distribution be established for σ_Y so it can be accurately characterized as a mixed uncertainty for comparison with maximum stress prediction.

4.8   Computational Cost

The proxy code for FEM simulation provided with the problem statement simply performs function evaluations based on Timoshenko-Krieger series solution for cylindrical shells [4]. Though these function evaluations are inexpensive (almost instantaneous), the corresponding FEM simulation needs several CPU hours depending on the grid level as per the data provided in the problem statement. The number of runs performed and the total cost of computation in terms of CPU hours are shown in Table 9.

The largest number of runs (30,000) was performed during the u-pooling step, where 3 different cases for u-pooling to $H = 51$ inches each were run with 10,000 simulations. Thus, the u-pooling study was performed using the coarsest mesh ($m = 1$). The MAVM calculations used the u-pooling runs and hence no additional runs were required for this part. A bias error arises due to the use of different grid sizes for u-pooling and for formation of the simulation p-box. However, we estimate that the bias error in the formation of MAVM due to the use of coarse mesh is small and would lead to further reduction in total uncertainty.

1,000 runs were made each at 2 different operating conditions ($\chi = 1$ and $\chi = 0.1$) on the coarse mesh for the global SA to determine the Sobol' indices. For local SA using finite

difference method only 9 runs were required for each of the 2 conditions, and thus medium mesh ($m = 2$) was used for this purpose. The parametric study mentioned in Sec. 4.1 was performed with 1320 runs on the finest mesh ($m = 4$) before the cost of computation was known. However, after the knowledge of the computational cost we recommend replacing this study with the parametric runs over the operating conditions space that was done for the surrogate model construction in Sec. 4.6, for a total of 125 runs on the fine mesh ($m = 3$).

The solution verification to determine numerical uncertainty was performed using 3 runs (for the three conditions) on each of the four mesh levels. Finally, the uncertainty propagation at the nominal conditions for failure prediction was performed with 10,000 runs using the medium mesh ($m = 2$). The total cost of computation was found to be approximately 1.6 million CPU hours which could be expensive for a practical application. Note that the largest cost in our analysis arises from the uncertainty propagation step (i.e., 1.05 million CPU hours for 10,000 simulations). Given the large variation in computational cost with mesh sizes, we recommend that more mesh levels be examined (such as intermediate mesh sizes between the coarse and medium meshes) to determine whether one could obtain satisfactory results for lower computational cost. Also, more information from experimental measurements could significantly reduce the interval uncertainty in input variables leading to requiring fewer computational simulations for an accurate estimation of total prediction uncertainty.

5   Conclusions

In this work, we employed the segregated propagation of probabilistic and interval-characterized uncertainties using the Probability Bounds Analysis (PBA) framework to address

the Sandia challenge problem of tank failure prediction. Systematic analysis was performed involving a parametric study to obtain a feel for the problem, global and local sensitivity analyses, numerical uncertainty estimation using Richardson extrapolation, and model form uncertainty estimation using MAVMs. Input uncertainties were characterized into aleatory, epistemic, and mixed uncertainties after careful observation of the provided data, qualitative observation from the parametric study, and sensitivity analyses. In order to convert the experimental data for displacements to stress data so MAVM could be generated, we used a surrogate model to convert experimental displacement data to "surrogate" stress data. We then used u-pooling to convert stress data from various conditions to "experimental" stress data points. The total uncertainty prediction was formed by appending the numerical uncertainty and the model form uncertainty to the p-box obtained from segregated propagation of input uncertainties.

The final p-box for maximum stress SRQ was used to determine that the probability of failure at nominal conditions is $P_{fail} = [0, 0.0034]$. The tanks are determined to be not safe. However, this conclusion is strongly dependent on the uncertainty in yield stress. Specifically, not enough data has been provided to characterize a key element of this failure analysis, i.e. yield stress.

The lack of experimental data in quantity (e.g., number of data points), quality (e.g., no tolerances or distributions for legacy data), and general information (e.g., age/environmental conditions of the tank population) were a major challenge in estimating the model form uncertainty. Also, the nature of bias created in the proxy code for radius of the tank (possibly

used to mimic the true solution) created problems at various stages of this process. Another challenge was that experimental data, especially geometry data, was all from the same tank and thus there were no estimates of tank-to-tank uncertainties.

The uncertainties just mentioned contributed large epistemic and aleatory uncertainty to our prediction of tank safety. Because of the large interval value of P_{fail}, one could conclude that the tanks *may* be safe; but that is not the issue that a decision maker must deal with. Faced with large uncertainty for a system that does not meet safety or performance requirements, a decision maker will then ask: "What are the major contributors to the predicted uncertainty?" While the present approach does segregate the effects of epistemic model input uncertainty, aleatory model input uncertainty, model form uncertainty, and numerical uncertainty, further information on which aleatory or epistemic uncertainties are important requires a sensitivity analysis. When interval-valued uncertainties and random variables exist together, as in the present analysis, special sensitivity analysis techniques must be used.

Acknowledgments

Nomenclature

|	
σ (psi)	von Mises stress
σ_Y (psi)	Yield stress
d or w (in)	Tank wall displacements, normal to the surface
x (in)	Axial location
ϕ (rad)	Circumferential angle or angular location
P (psig)	Gauge pressure
γ (lbs/in^3)	Liquid specific weight
H (in)	Liquid composition or mass fraction
E (psi)	Young's modulus
F_s	Factor of safety for numerical uncertainty
$F_s(k)$	Factor of safety for modified area validation metric
F_0, F_1	Constants to determine $F_s(k)$
$F(Y)$	Simulation outcome (value, CDF, or p-box)
ν	Poisson's ratio
L (in)	Length
R (in)	Radius
T (in)	Wall thickness
m	Mesh ID
P_{fail}	Probability of failure

Figures
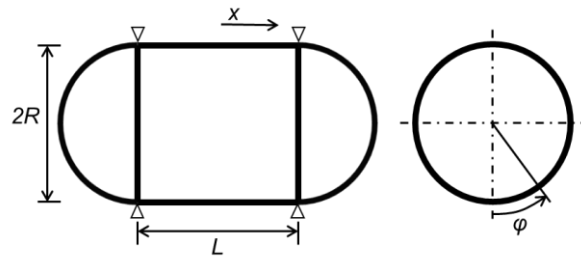
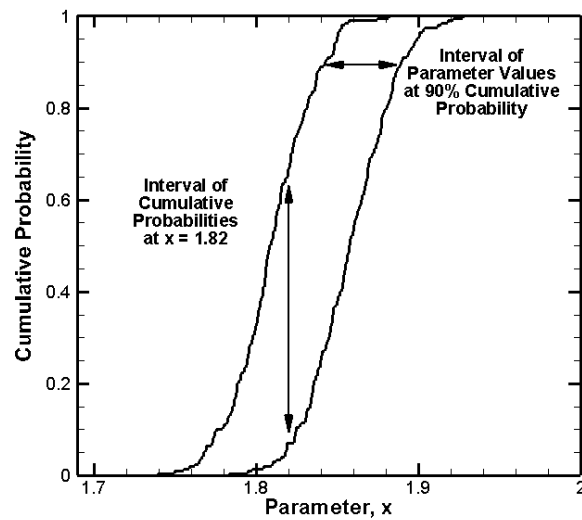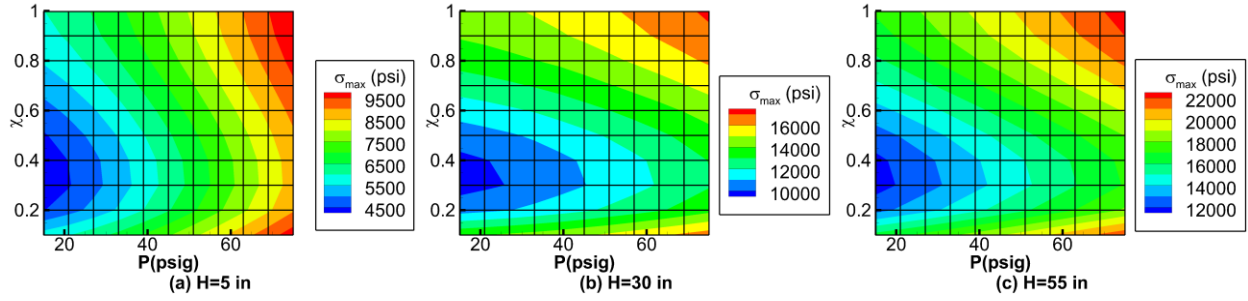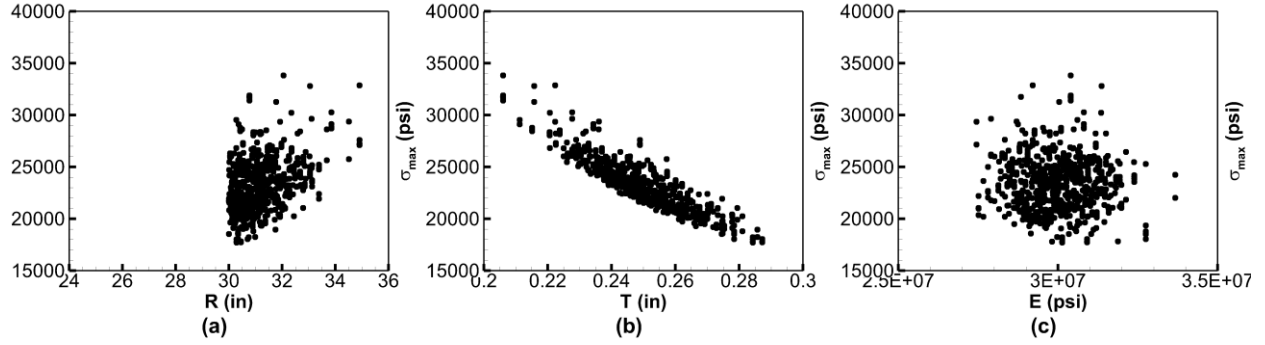**Figure 1 Side view (left) and axial view (right) of the tanks.**

Figure 2 An example of probability box (p-box) for a parameter (x) that is a mixture of both aleatory (random) and epistemic (lack of knowledge) uncertainty
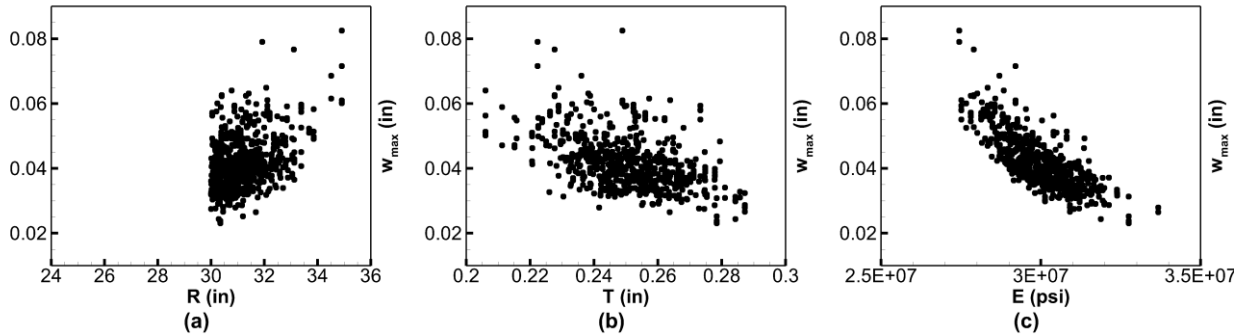
Figure 3 Maximum stress values for the parametric space of pressure and liquid composition for different liquid heights: (a) $H = 5$ **in, (b)** $H = 30$ **in, and (c)** $H = 55$ **in.**

Figure 4 Effect of perturbing various input variables on maximum stress: (a) tank radiusn (b) tank surface thickness, and (c) Young's modulus. Note that input distribution for radius is truncated at $R = 30$ **in.**

Figure 5 Effect of perturbing various input variables on maximum normal displacement: (a) tank radius, (b) tank surface thickness, and (c) Young's modulus. Note that input distribution for radius is truncated at $R = 30$ **in.**
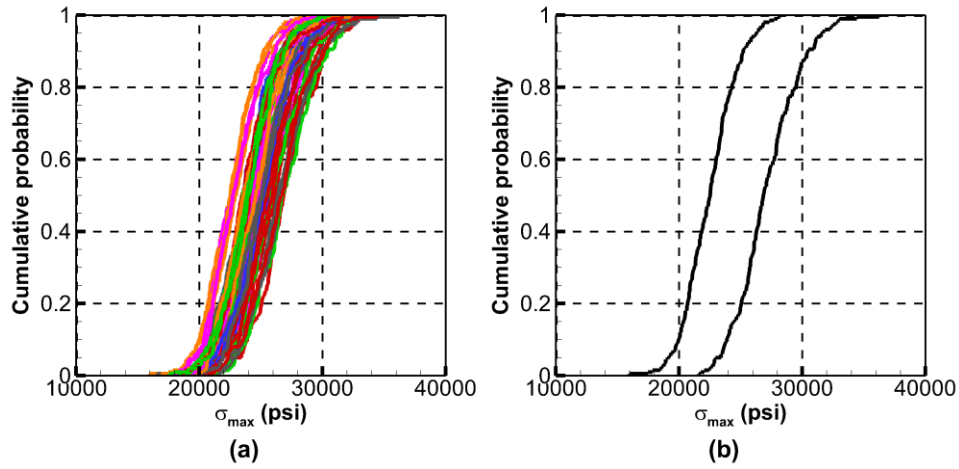
Figure 6 Uncertainty propagation at nominal conditions. $M = 50$ epistemic samples and $N = 200$ aleatory samples were used (for a total of 10000 simulations) on medium grid ($m = 2$): (a) all CDFs, (b) p-box.
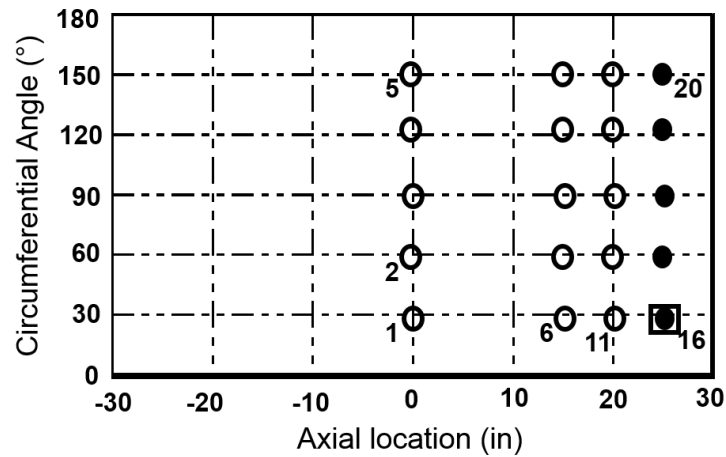

Figure 7 Locations on the tank surface where displacement data is measured during field tests. 20 locations are marked with circles. Filled circles are locations where we convert the displacement data to stress data. Location#16 is where all experimental data is pooled for MAVM calculation.
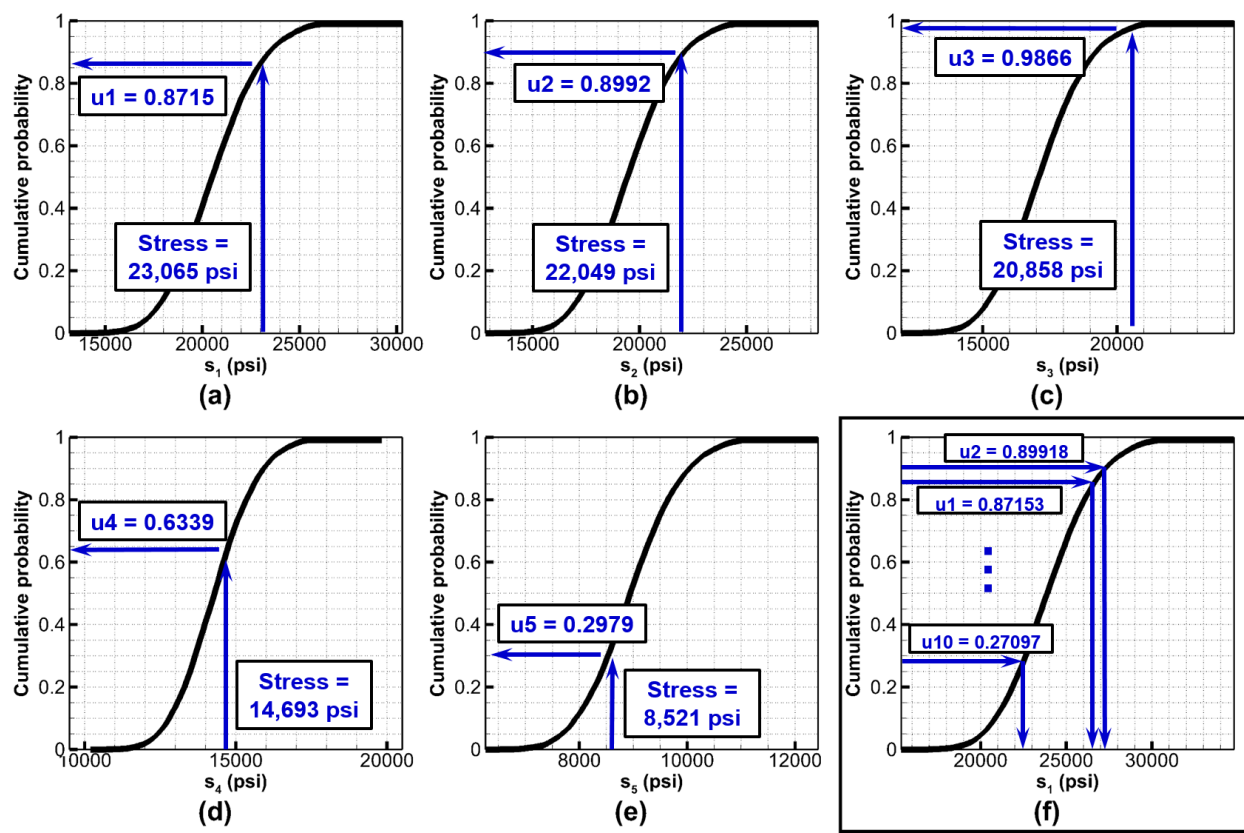
Figure 8 The u-pooling process explained for pooling data to one location at $H = 51$ inches liquid height: (a)-(e) probing for probability levels (u-values) based upon the surrogate stress data, and (f) collecting stress data for the experimental CDF.
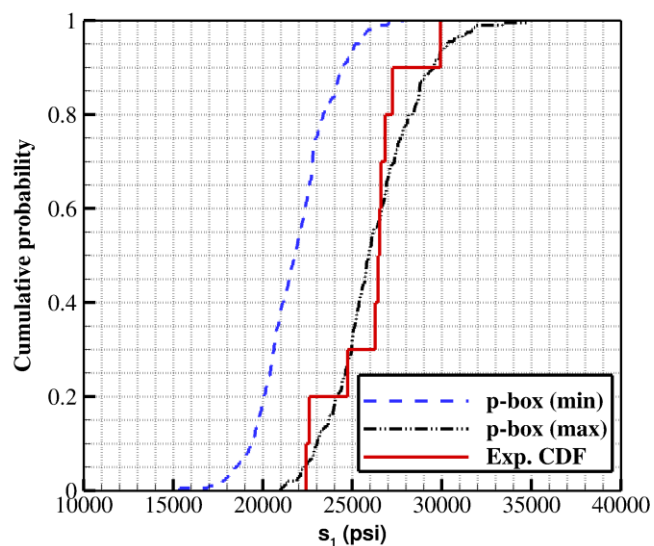
Figure 9 MAVM calculation using simulation p-box and experimental (u-pooled) discrete CDF at $H =$ 51 inches operating condition.
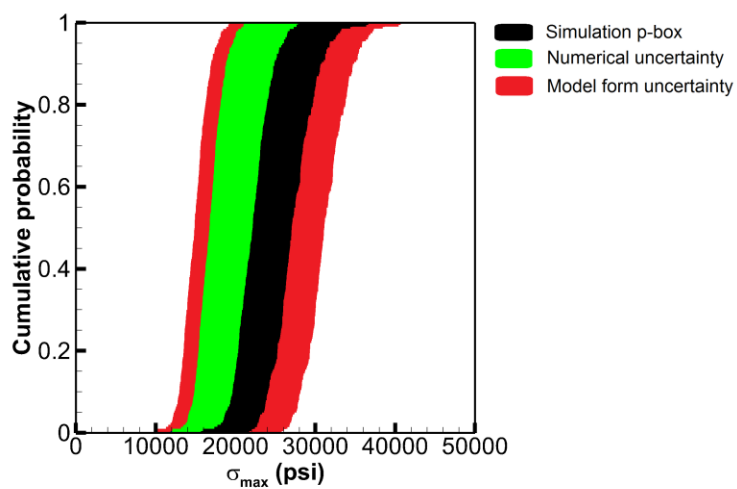
Figure 10 Total prediction uncertainty.
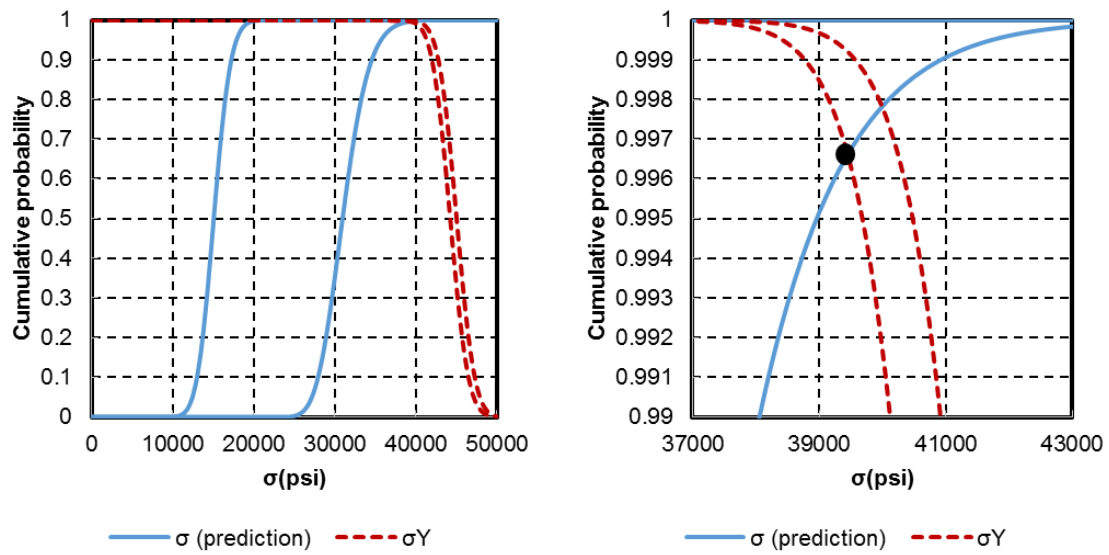
Figure 11Prediction p-box for maximum stress (solid curves) and p-box for yield stress (dotted curves): (left) CDFs for maximum stress and CCDFs for yield stress, (right) enlarged view. The relevant point of intersection is shown with a solid circle which results in maximum probability of failure as: $1 - 0.9966 = 0.0034.$

Tables

Table 1 Maximum numerical uncertainty (U_{num}) for different grid sizes.

Grid Level	Cost/Run (CPU Hours)	U_{num} (Max stress)	U_{num} (Max Displacement)
Mesh 4 (finest)	10,200	11.3%	2.2%
Mesh 3 (fine)	1,100	17.0%	3.3%
Mesh 2 (medium)	105	**24.1%**	**5.0%**
Mesh 1 (coarse)	12	31.2%	7.6%

Table 2 User-defined and Dakota-sampled values of input parameter distribution for global sensitivity analysis.

	User-defined				Dakota-sampled			
	Dist.	Mean	SD	COV (%)	Mean	Var	SD	COV (%)
P (psig)	Normal	75	1.875	2.5	74.99	3.47	1.86	2.5
χ	Normal	1	0.025	2.5	0.98	2.2E-04	0.0149	1.5
H (in)	Normal	55	0.275	0.5	54.92	0.05	0.22	0.4
E (psi)	Normal	3.0E+07	1.1E+06	3.5	3.0E+07	1.1E+12	1.0E+06	3.5
ν	Normal	0.27	0.0108	4.0	0.27	1.2E-04	0.0109	4.0
L (in)	Normal	60	1.2	2.0	60	1.41	1.19	2.0
R (in)	Normal	30	1.5	5.0	31.20	0.86	0.93	3.0
T (in)	Normal	0.25	0.015	6.0	0.25	0.00	0.02	6.0

Table 3 Main and total Sobol' indices for global sensitivity analysis. (Bold numbers indicate significant correlations)

	σ_{max}		$w@\sigma_{max}$		w_{max}	
	Main	Total	Main	Total	Main	Total
P	0.00	0.00	0.01	0.00	0.00	0.00
χ	0.00	0.00	0.00	0.00	0.00	0.00
H	0.00	0.00	0.00	0.00	0.00	0.00
E	0.00	0.00	**0.64**	**0.63**	**0.71**	**0.69**
ν	0.00	0.00	0.00	0.00	0.00	0.00
L	0.00	0.00	0.01	0.01	0.01	0.01
R	**0.14**	**0.18**	**0.16**	**0.25**	**0.08**	**0.16**
T	**0.84**	**0.85**	**0.16**	**0.24**	**0.18**	**0.24**

Table 4 Finite-difference based sensitivity. (Bold numbers represent significant correlations while shaded cells represent moderately significant correlations)

	σ_{max}	$w@\sigma_{max}$	w_{max}
d/dP	-0.35	-0.31	-0.23
$d/d\chi$	-0.13	-0.14	-0.15
d/dH	-0.69	-0.62	-0.46
d/dE	0.00	**4.09**	**4.09**
d/dv	0.04	0.00	0.00
d/dL	0.08	-0.42	-0.83
d/dR	**-1.65**	**-2.75**	**-2.35**
d/dT	**1.55**	**1.67**	**1.57**

Table 5 Categorization of input parameters.

Model input parameters (material)	Model input parameters (geometry)
Young's modulus, E Poisson's ratio, v	Thickness, T Radius, R Length, L
System excitation (or control) parameters	Numerical input parameter
Gauge pressure, P Liquid composition, χ Liquid height, H	Mesh ID, m

Table 6 Characterization of input uncertainties.

Epistemic	Nominal	Lower bnd.	Upper bnd.	Type	
χ	1.0	0.95	1.0	Interval	
H (in)	50.0	49.0	51.0	Interval	
$\psi = H/2R_1$	--	0.9182	0.9557	Interval	(for $m = 1$)
$\psi = H/2R_1$		0.8963	0.9329	Interval	(for $m = 2$)

Aleatory variable	Nominal				COV (%)	SD
P (psig)	73.5			Normal	2.5	1.8375
L (in)	60.0			Normal	1	0.6
v	0.27			Normal	4	0.0108

Mixed	Nominal	Lower bnd.	Upper bnd.		COV (%)	SD
E (psi)	3E7	2.8141E7	3E7	Interval & Normal	2	0.06E7
R (in)	30.0	30.0	31.1527	Interval & Normal	3	0.9
T (in)	0.25	0.23132	0.25	Interval & Normal	3	0.0075

Table 7 Operating conditions for which displacement data is available from Data set #6. (Only conditions relevant to u-pooling are shown here)

Case#	P (pisg)	χ	γ	H (in)
Case52	63.167	0.7	3.0183	51
Case53	64.636	0.4	2.7334	54

Table 8 u-pooling to get experimental stress data points at H = 51 inches.

Case	Location	surrogate stress	u	exp. stress at location#16 (H=51 in)
case52 (H=51 in)	16	23065	0.8715	26833.7
	17	22049	0.8992	27250.2
	18	20858	0.9866	29925.8
	19	14693	0.6339	24748.0
	20	8521	0.2979	22601.7
case53 (H=54 in)	16	20086	0.8380	26446.5
	17	19041	0.8527	26599.7
	18	16798	0.8459	26527.8
	19	13651	0.8230	26286.2
	20	9049	0.2710	22411.9

Table 9 Cost of computation.

Grid Level	Cost (CPU Hours)	# of Runs	Cost (CPU Hours)
4 (finest)	10,200	3	30,600
3 (fine)	1,100	128	140,800
2 (medium)	105	10,021	1,052,205
1 (coarse)	12	32,003	384,036
TOTAL	-----	**82,149**	**1,607,641**

References

[1] Hu, Kenneth T., and George E. Orient. 2015. The 2014 Sandia V&V Challenge Problem: A case study in simulation, analysis, and decision support. *ASME Journal of Verification, Validation and Uncertainty Quantification*, in same volume.

[2] Hu, Kenneth. 2014. 2014 V&V Challenge: Problem Statement, SAND Report 2013-10486P. Sandia National Laboratories.

[3] Bernardini, A. and Tonon, F. 2010. *Bounding Uncertainty in Civil Engineering: Theoretical Background*, Springer-Verlag Berlin Heidelberg. doi: http://dx.doi.org/ 10.1007/978-3-642-11190-7.

[4] Timoshenko, S., and S. Woinowsky-Krieger. 1987. *Theory of Plates and Shells*. New York: McGraw-Hill.

[5] Oberkampf, William L., and Christopher J. Roy. 2010. *Verification and Validation in Scientific Computing*: Cambridge University Press, Cambridge.

[6] Roy, Christopher J., and William L. Oberkampf. 2011. "A Comprehensive Framework for Verification, Validation, and Uncertainty Quantification in Scientific Computing." *Computer Methods in Applied Mechanics and Engineering* 200 (25):2131-2144.

[7] Beer, Michael, Scott Ferson, and Vladik Kreinovich. 2013. "Imprecise probabilities in engineering analyses." *Mechanical Systems and Signal Processing* 37 (1-2):4-29. doi: 10.1016/j.ymssp.2013.01.024.

[8] Ghosh, Jayanta K., Mohan Delampady, and Tapas Samanta. 2006. *An Introduction to Bayesian Analysis: Theory and Methods*. Berlin: Springer.

[9] Roy, Christopher J., and Michael S. Balch. 2012. "A Holistic Approach to Uncertainty Quantification with Application to Supersonic Nozzle Thrust." 2 (4):363-381. doi: http://dx.doi.org/10.1615/Int.J.UncertaintyQuantification.2012003562.

[10] Deodatis, G., and P. D. Spanos. 2011. "Computational Stochastic Mechanics." Proceedings of the Sixth International Conference on Computational Stochastic Mechanics.

[11] Cullen, A. C., Frey, H. C. 1999. *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. New York: Plenum Press.

[12] Veneziano, Daniele, Anna Agarwal, and Erdem Karaca. 2009. "Decision making with epistemic uncertainty under safety constraints: An application to seismic design." *Probabilistic Engineering Mechanics* 24 (3):426-437. doi: http://dx.doi.org/10.1016/j.probengmech.2008.12.004.

[13] Ferson, Scott, and Janos G. Hajagos. 2004. "Arithmetic with uncertain numbers: rigorous and (often) best possible answers." *Reliability Engineering & System Safety* 85 (1–3):135-152. doi: http://dx.doi.org/10.1016/j.ress.2004.03.008.

[14] Leonard, T. and Hsu, J.S.J. 1999. Bayesian Methods: *An Analysis for Statisticians and Interdisciplinary Researchers*: Cambridge University Press, New York.

[15] van den Bos, A. 2007. Parameter Estimation for Scientists and Engineers, Wiley-Interscience, Hoboken, NJ.

[16] Ferson, Scott, and Lev R. Ginzburg. 1996. "Different methods are needed to propagate ignorance and variability." *Reliability Engineering & System Safety* 54 (2–3):133-144. doi: http://dx.doi.org/10.1016/S0951-8320(96)00071-3.

[17] Roy, Christopher J. 2005. "Review of code and solution verification procedures for computational simulation." *J. Comput. Phys.* 205 (1):131-156. doi: http://dx.doi.org/10.1016/j.jcp.2004.10.036.

[18] Klir, George J. 2006. *Uncertainty and Information: Foundations of Generalized Information Theory*. Hoboken, N.J.: Wiley-Interscience.

[19] AIAA. 1998. AIAA Guide for the Verification and Validation of Computational Fluid Dynamics Simulations (G-077-1998e). Reston, VA.

[20] ASME. 2006. Guide for Verification and Validation in Computational Solid Mechanics, American Society of Mechanical Engineers, ASME Standard V&V 10-2006. Newyork, NY.

[21] Kennedy, Marc C., and Anthony O'Hagan. 2001. "Bayesian calibration of computer models." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63 (3):425-464. doi: 10.1111/1467-9868.00294.

[22] Oberkampf, W. L. and Scott Ferson. 2007. "Model validation under both aleatory and epistemic uncertainty." NATO/RTO Symposium on Computational Uncertainty in Military Vehicle Design, AVT-147/RSY-022. Athens, Greece, NATO.

[23] Ferson, Scott, William L. Oberkampf, and Lev Ginzburg. 2008. "Model validation and predictive capability for the thermal challenge problem." *Computer Methods in Applied Mechanics and Engineering* 197 (29–32):2408-2430. doi: http://dx.doi.org/10.1016/j.cma.2007.07.030.

[24] Voyles, Ian T., and Christopher J. Roy. 2014. "Evaluation of Model Validation Techniques in the Presence of Uncertainty." In *16th AIAA Non-Deterministic Approaches Conference*. American Institute of Aeronautics and Astronautics. doi: http://dx.doi.org/10.2514/6.2014-0120.

[25] Voyles, Ian T., and Christopher J. Roy. 2015. "Evaluation of Model Validation Techniques in the Presence of Aleatory and Epistemic Input Uncertainties." In *17th AIAA Non-Deterministic Approaches Conference*. American Institute of Aeronautics and Astronautics. http://dx.doi.org/10.2514/6.2015-1374.

[26] Adams, B M, L E Bauman, W J Bohnhoff, K R Dalbey, M S Ebeida, J P Eddy, M S Eldred, P D Hough, K T Hu, J D Jakeman, L P Swiler, and D M Vigil. December 2009. Updated November 7, 2013. DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 5.4 User's Manual, Sandia Technical Report SAND2010-2183. Sandia National Laboratories.

[27] Saltelli, Andrea, Stefano Tarantola, Francesca Campolongo, and Marco Ratto. 2002. *Sensitivity Analysis in Practice*: John Wiley & Sons, Ltd.

[28] Roache, Patrick J.. 2009. *Fundamentals of Verification and Validation*. Socorro, NM: Hermosa Publishers.