SANDIA REPORT

SAND2005-4777
Unlimited Release
Printed October 2005

The Method of Belief Scales as a Means for Dealing with Uncertainty in Tough Regulatory Decisions

Martin Pilch

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of
Energy under Contract DE-AC04-94AL85000

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2005-4777
Unlimited Release
Printed October 2005

The Method of Belief Scales as a Means for Dealing with Uncertainty in Tough Regulatory Decisions

MARTIN PILCH
Validation and Uncertainty Quantification Department, 1533

Sandia National Laboratories
P. O. Box 5800
Albuquerque, New Mexico 87185-0828

Abstract

Modeling and simulation is playing an increasing role in supporting tough regulatory decisions, which are typically characterized by *variabilities* and *uncertainties* in the scenarios, input conditions, failure criteria, model parameters, and even model form. *Variability* exists when there is a statistically significant database that is fully relevant to the application. *Uncertainty*, on the other hand, is characterized by some degree of ignorance. A simple algebraic problem was used to illustrate how various risk methodologies address variability and uncertainty in a regulatory context. These traditional risk methodologies include probabilistic methods (including frequensic and Bayesian perspectives) and second-order methods where variabilities and uncertainties are treated separately. Representing uncertainties with (subjective) probability distributions and using probabilistic methods to propagate subjective distributions can lead to results that are not logically consistent with available knowledge and that may not be conservative. The Method of Belief Scales (MBS) is developed as a means to logically aggregate uncertain input information and to propagate that information through the model to a set of results that are scrutable, easily interpretable by the nonexpert, and logically consistent with the available input information. The MBS, particularly in conjunction with sensitivity analyses, has the potential to be more computationally efficient than other risk methodologies. The regulatory language must be tailored to the specific risk methodology if ambiguity and conflict are to be avoided.

Intentionally Left Blank

Contents

Figures

Tables

Acronyms and Abbreviations

ASC	Advanced Simulation and Computing
CCDF	complementary cumulative distribution function
CDF	cumulative distribution function
CR	confidence ratio
EPA	Environmental Protection Agency
FEV	Number of evaluations
LHS	Latin Hypercube Sampling
MBS	Method of Belief Scales
PLOAS	Probability of Loss of Assured Safety
QMU	quantifying design margins and uncertainties
QR	quantified reliability
ROAAM	Risk-Oriented Accident Analysis Methodology
Sandia	Sandia National Laboratories
V&V	verification and validation
WIPP	Waste Isolation Pilot Plant

Intentionally Left Blank

1 Introduction

Sandia National Laboratories (Sandia) has the responsibility for providing technical evidence that supports high-consequence decision making on issues concerning the U.S. stockpile of nuclear weapons and other important issues of national interest. Nonlinear multiphysics in complex geometries characterize the systems of interest; consequently, numerous physics-based computer codes are being developed to address the full range of phenomena that arise in the many issues of interest. These codes include fluid dynamics codes (creeping flow to supersonic flow), a thermal response code, a large-deformation structural mechanics code, a structural dynamics code, a shock physics code, radiation transport codes, an electromagnetics code, and a circuit code. Three-dimensional, high-fidelity models are emphasized where appropriate, and some codes can be dynamically coupled with others should the need arise to model coupled multiphysics. These advanced applications codes are run on state-of-the-art computer platforms capable of 3–40 terops with even larger platforms planned for the future. The applications codes and the platforms on which these codes run are being developed as part of the Advanced Simulation and Computing (ASC) program (Kusnezov and Soudah, 2004).

Modeling and simulation is playing an increasing role in supporting weapons qualification and other tough regulatory decisions, which are typically characterized by *variabilities* and *uncertainties* in the scenarios, input conditions, failure criteria, model parameters, and even in the model form. Sandia's approach to weapons certification has historically been test-based, with the design goals based on conservative scenarios, conservative assessments, and conservative requirements. This approach attempts to provide robust designs by enveloping variabilities and uncertainties. In the modern environment, modeling and simulation (enabled by the ASC program) is playing an increasingly important role in providing a more defensible *assertion* of what "conservative" really means, based on a fundamental understanding of the controlling physics. The Method of Belief Scales (MBS) provides a formalism for identifying conservative states while also providing sensitivity information so that research activities can be focused to reduce overall uncertainty.

A distinction can be made between variabilities and uncertainties, but there is no universal agreement that this distinction should be made. Alternate terminologies for *variability* include aleatory uncertainties, irreducible uncertainties, inherent uncertainties, objective uncertainty, and stochastic uncertainty. Variability exists when there is a statistically significant database that is fully relevant to the application.

For example, one might find unit/unit or setup/setup variabilities of nominally identical systems or component-failure thresholds based on large sample testing. Variability can be represented by frequency distributions for both inputs and outputs. When variabilities dominate, the expectation is that input distributions, propagated through a valid model, will yield the same distribution of results that would be realized if the system was tested a large number of times.

Uncertainty, on the other hand, is characterized by some degree of ignorance. Alternate terminologies for uncertainty include epistemic uncertainty, reducible uncertainty, subjective uncertainty, and model-form uncertainty. For example, one might find cases where there are alternate plausible models and cases where there are nonexistent, sparse, or inconsistent experimental data for input distributions. In such situations, it is not possible to ascribe a frequensic interpretation to ensembles of input or output information.

The goal of any particular risk analysis is to make better decisions by acknowledging relevant variabilities and uncertainties and assessing the impact they have on regulatory metrics of interest. Assessments of the type envisioned here are often termed risk analyses, performance assessments, or *risk-informed* decision analyses. We distinguish *risk-informed* decisions from *risk-based* decisions. The latter implies that any decision is solely dependent on absolute results of the study, whereas most high-consequence decisions are ultimately assertion-based (Helton and Breeding, 1993; Helton et. al., 2000). Assertion-based decisions can be risk-informed while allowing the decision maker to subjectively weigh technical and programmatic "intangibles." More specifically, risk-informed decisions are an exercise in *due diligence* by taking a systems approach to identifying failure modes and by providing an infrastructure that favors transparency for peer review into the relative contributors to uncertainty (Apostolakis, 2004)

The credibility of models, through which variabilities and models are propagated, is a critical prerequisite to risk-informed decisions. Ideally, scrutable evidence

(1) that computer codes are developed to institutional software quality requirements,

(2) that codes and analyses are verified to ensure convergence to the right answer for the intended application, and

(3) that the models are validated to ensure adequate representation of the right physics

should be available for high-consequence issues. Validation activities play a central role in quantifying variabilities and uncertainties for the intended application. Increased risk from unknown/unknowns is greater when due diligence is not exercised in the use of modeling and simulation. The focus of this report is on the representation and propagation of variabilities and uncertainties; however, for any real application, the credibility of the models and their inputs cannot be ignored.

Risk analyses can be computationally expensive. Approximately 50 to 100 function evaluations is a challenging, but achievable, target for many of the weapons applications targeted by the ASC program. Variabilities that might be treated could number in the hundreds, one for every material property or geometric dimension. Typically, only a couple of these variabilities are shown to be important relative to the others; and even then, variabilities are often dominated by uncertainties, which could number in the dozens. Once again, it is typical (subject to actual verification) that only a few of the many uncertainties actually dominate for any given application. Expert judgment is unavoidable in determining which variabilities and uncertainties should be treated in the risk analysis. However, it is desirable to minimize the reliance on experts because such judgments can be unreliable when nonlinear multiphysics in complex geometries are involved. The grand challenges, then, are to address as many variabilities and uncertainties as is practical; to perform sensitivity analyses to quantitatively sort out the dominant contributors to uncertainty; to assess whether regulatory requirements are met and to quantify confidence in those assessments; and, finally, to do all this with a very limited number of evaluations of the full physics models.

In August 2002, Sandia conducted an international workshop on epistemic uncertainties (Oberkampf et al. 2004). The workshop used a series of challenge problems to facilitate comparison of different approaches to representing variabilities and uncertainties within the context of regulatory decisions. A simple mathematical model was employed to focus on aggregating the information, propagating the inputs through to the outputs, and representing the outputs.

Keeping in mind that the challenge problems are computationally simple surrogates for the issues and models of interest, requirements and desirable features of the solution approach are as follows:

1. The method for aggregating evidence and constructing input distributions should favor repeatability by independent experts looking at the same evidence.

2. The method of aggregating evidence and constructing input distributions should be logically consistent with the available evidence.

3. The mechanics of propagating information through the model should preferably be nonintrusive,[1] i.e., black-box treatment of the model.

4. The interpretation of the output distributions must be clear, explicit, and consistent with the information embodied in the input distributions, i.e., propagation should be *information preserving*, implying no more information than the evidence justifies.

5. The methodology should be applicable to scenario uncertainties, parameter uncertainties, and model-form uncertainties.

6. The sensitivity of key outputs should be quantifiable in terms of uncertainties in inputs and model form.

7. The results should be numerically *converged* in some defined sense, and any "representational errors" should be quantified.[2]

---

[1] Intrusive methodologies add significantly to the cost of developing and maintaining ASC codes because such methodologies are invariably code-specific, requiring repetition of development costs from code to code. In addition, intrusive techniques are often founded on certain assumptions, approximations, or linearizations that limit their utility in certain real-world applications.
[2] The assumption of a distribution form or the use of a surrogate model to replace full physics models introduces "representational errors."

8. The computational burden should be quantifiable and not prohibitive.

It is not the position of Sandia's ASC Verification and Validation (V&V) program to prescribe how these requirements and desirable features are met. However, the MBS will be shown to have favorable performance characteristics when judged according to these concepts and in relation to other traditional approaches that might be employed. It should be noted that the Sandia ASC V&V program is funding two other activities for the simultaneous representation and propagation of variabilities and uncertainties: one activity is based on Dempster-Schafer theory (Helton et al. 2004), and the other activity is based on polynomial chaos concepts (Red-Horse and Benjamin 2004). Other methods can also be found in the literature, some of which will be cited throughout this report. Ultimately, experience with relevant applications will dictate which methods have adequate rigor and adequate practicality to warrant adoption by the broader community.

Section 2 of this report describes the MBS. One of the challenge problems (Problem 5a) from the international workshop (Helton and Oberkampf 2004) is used as a demonstration of the MBS and other risk-analysis methodologies. Section 3 provides a "behind- the- scenes" characterization of the challenge-problem solution. This kind of knowledge does not exist for real applications, but insight into the solution is used here to benchmark and provide insight into the various solution approaches discussed. Traditional solution approaches to problems involving variabilities and uncertainties are explored in three cases:

1. Deterministic (Section 4): One- or two-point calculations are aimed at characterizing the "nominal" response of the system or the "bounding response" of the system.

2. Bayesian (Section 5): Variabilities and uncertainties are represented by probability distributions, which are propagated through the model using probabilistic methods.

3. Second-order methods using subjective probabilities to represent uncertainties (Section 6): Variabilities and uncertainties are treated separately, and both are propagated through the model using probabilistic methods.

Two additional cases are discussed that employ the MBS, which overcomes some of the deficiencies identified in the more traditional approaches:

4. Second-order methods using the MBS to represent uncertainties (Section 7): Variabilities and uncertainties are treated separately. Variabilities are represented by frequency distributions, which are propagated through the model using traditional probabilistic methods, while uncertainties are propagated through the model using a nonprobabilistic rule-based methodology.

5. The MBS with distributional relaxation (Section 8): Probability distributions, characterizing variabilities, are "relaxed" into belief distributions, which are propagated through the model using a nonprobabilistic rule-based methodology.

Sections 4 through 8 also point out the strengths and weaknesses in each of the five analysis approaches, and address sensitivity analyses and quantification of the computational burden in the context of each approach. Section 9 considers the regulatory language and the quantification of margins as they relate to the adopted analysis approach. The report ends with Section 10, which summarizes the various approaches and gives some concluding remarks.

2 Proposed Methodology: The MBS

We focus on the representation, propagation, and interpretation of *uncertainties* in the context of high-consequence decisions. The treatment of variabilities using probabilistic tools is well accepted and need not be described here. The combined treatment of uncertainties and variabilities will be illustrated in the context of the challenge problem.

The first and most natural response in the face of large epistemic uncertainties is to seek and define the limits of credibility. Interval analysis is the simplest manifestation of this process. An interval is defined for input quantities such that any value within the interval is considered "possible," with no statement of frequency other than that it is finite (but unknown) at all points within the interval and that the possibility and frequency values occurring outside the interval is zero. The expectation is that all outputs of interest will also be represented as intervals if all the inputs are represented as intervals. This, of course, should be confirmed in any given application.

Interval analyses have an intuitive appeal; however, a rigorous adherence to the definition of input intervals (i.e., zero frequency outside the interval) often leads to such large intervals on inputs and the resulting output metrics that regulatory requirements may be difficult to meet, and any design that does meet the requirements may be overly conservative and unnecessarily costly. Advocates for the rejected technology might then lobby for reduced input intervals, often based on reasonable physical arguments. This can result in an intense debate with critics who argue that rejected values cannot be *positively* excluded. A practical approach to this problem is to allow for graded degrees of belief so that no arguable input is fully excluded. This has traditionally been done using subjective probabilities that were propagated by traditional probabilistic tools; however, we will illustrate later, in Section 5, that this approach is not *information preserving*, i.e., the results imply more than the evidence can support.

The approach adopted here utilizes physically based, graded degrees of belief that can be propagated through the model using a nonprobabilistic rule-based methodology defined later in Table 1. The intent is to envelop all possible inputs and all possible resulting outputs that one consistent with the possible inputs.

Table 1 summarizes the physical scale used for quantifying the construction of input distributions. It will be argued that this same table can also be used to provide a physical interpretation of the output results. The table is motivated, in part, by a desire to facilitate consistency amongst knowledgeable experts on the construction of "input distributions" and the interpretation of "output distributions."

Table 1. Process Characteristics for Graded Degrees of Belief

Belief Scale	Process Characteristics
Bel = 1	Behavior is ***physically reasonable*** and within the limits of credibility.
Bel = 0.1	Behavior is ***physically unreasonable***, but the behavior cannot be positively excluded, i.e., the behavior is still within the limits of credibility. This is a reflection of our will to doubt.[a]
Bel ≤ 0.01	Behavior is ***physically incredible*** and violates well-known reality.[b] Its occurrence can be argued against positively. (Note: There is no need to represent these values in input distributions.)

[a] The transition from Bel = 1 to Bel = 0.1 occurs where the frequency or possibility is arguably zero but not definitively zero.

[b] The transition from Bel = 0.1 to Bel = 0.01 occurs where the frequency or possibility is unquestionably zero.

Note: Section 8 discusses situations where these rigorous interpretations can be relaxed to achieve significant computational savings.

The methodology proposed here is referred to as the *Method of Belief Scales* (MBS) because of the characteristics embodied in Table 1. By intent, the process characteristics convey <u>no</u> sense of *frequency* or likelihood. The intended focus on *limits of credibility* greatly minimizes the "overconfidence" and "anchoring (bias)" that commonly plague expert elicitation processes. The subjective assessments of credibility apply not only to the "values" in question but also to the relevance of the available evidence. Peer review is critical to all aspects of a risk analysis, including the quantification of input distributions, as a validation of subjective judgments, which cannot be wholly avoided (only managed).

The coarse granularity in characterizing belief in Table 1 is appropriate for lack-of-knowledge issues. What incremental *characteristics* of evidence would support construction of a continuous distribution? Typically, analysts who construct continuous subjective (belief) distributions do so by establishing expectation (e.g., a distribution mean) and limits of credibility (e.g., the 5 and 95 percentiles of the distribution) and then *arbitrarily* drawing a continuous distribution through the points. However, what characteristic of evidence, other than a statistically significant database, allows analysts to say, for instance, that they are 37% confident that a parameter value is less than 3.7?

The motivation and construction of Table 1 for the MBS is strongly influenced by a similar table used in the Risk-Oriented Accident Analysis Methodology (ROAAM) (Theofanous 1996); however, there are some significant differences between the MBS and ROAAM tables regarding the construction of input distributions, as noted below. Figure 1 illustrates some possible MBS quantifications of uncertain inputs.

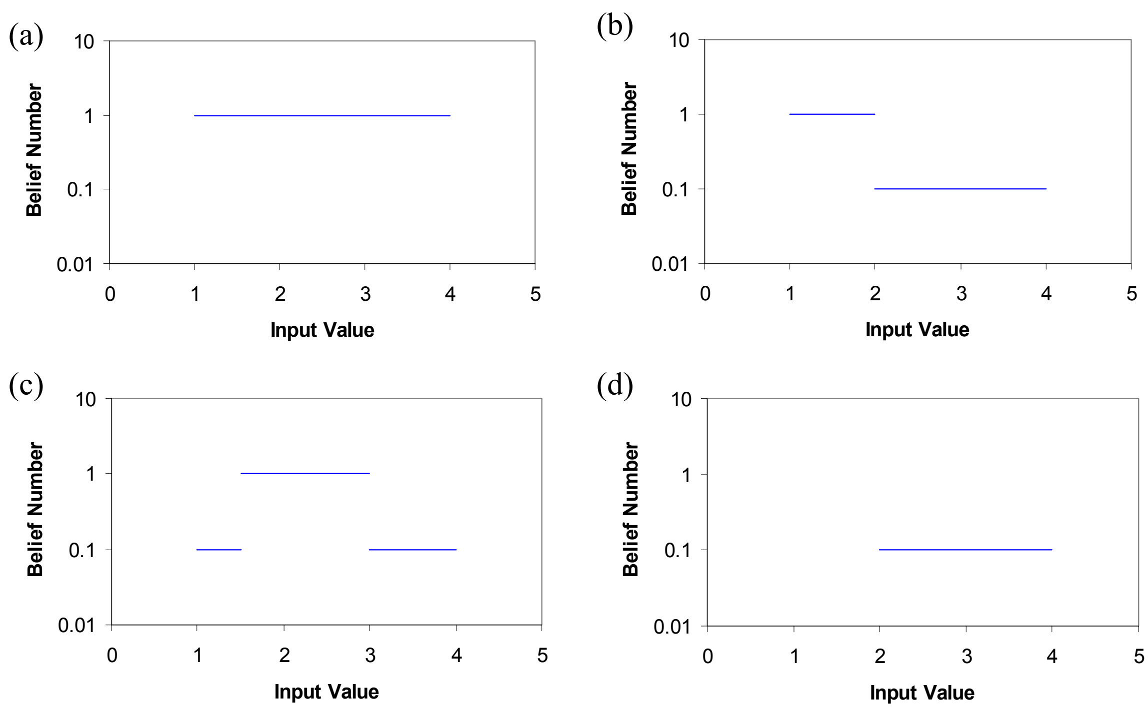


Figure 1. Representative input distributions for the MBS.

The following observations and comments about the input distributions in Figure 1 point out similarities and differences between the MBS and the ROAAM. Input distributions do not need to include all values listed in Table 1 in an ascending/descending fashion, and the distributions do not need to be symmetric. This is the same as in the ROAAM.

At most, there are only two tiers for graded belief with the MBS. The ROAAM allows an additional tier when grading belief.

The belief numbers are assigned to all values on a given subinterval, for example, Bel(1.9) = Bel(2.2) = Bel(2.87) = 1 in Figure 1c. This is significantly different from the ROAAM, where the subjective probabilities are assigned to the whole subinterval interval and the probability of a given value within the interval is incrementally small, i.e., PDF(x)*Δx.

Belief numbers can be additive, superadditive, or subadditive; i.e., the sum of belief numbers does not have to add to unity with the MBS. This is significantly different from the ROAAM, where the subjective probabilities of all intervals must add to unity.

The construction of input distributions according to Table 1 should be based on the physical *characteristics* of the available evidence and not on the quantity of available evidence. Typically, one first seeks evidence to establish minimum and maximum values of the input quantity that do not violate well-known reality or physical laws. This fixes the maximum credible range of input values, and all values within this range could be judged physically reasonable (Bel = 1) if no additional evidence is available. A thermodynamic limitation to energy exchange is an example of how one of these physical limits of credibility might be established.

There may be additional evidence that some credible values for the input are physically unreasonable, further restricting the range of physically reasonable values, i.e., those values assigned Bel = 1. For example, it might be argued that rate limitations preclude the full thermodynamically allowable energy exchange; consequently, the analyst may judge that some credible input is physically unreasonable but that these inputs cannot be positively excluded from the input characterization. These credible, although physically unreasonable, inputs can be maintained in further analyses (as a reflection of our will to doubt) but with diminished belief by assigning Bel = 0.1 to the inputs.

When multiple pieces of evidence need to be combined, aggregation process should be focused on characterizing the evidence with regard to the natural limits of credibility listed in Table 1 and discussed above. Naturally, any relevant measurement should be assigned Bel = 1. The intent is to proactively focus the debate on natural limits of credibility with the expectation that technical consensus can be approached. This is notably different from other aggregation schemes that try to "average" in some sense. Averaging schemes are too sensitive to what is being averaged, and small changes can significantly alter the results.

Sometimes averaging of experts is advocated. More weight is given to certain input when there is greater community consensus on the specific values. Even when all the experts have access to the same evidence (which often is not the case), this process is flawed because the average is too strongly influenced by *who* is represented on the expert panel. Change the character of the panel or the number of panel members and the average of the experts could change significantly, especially when the size of the panel is small. "Stacking the panel" with supporters (or dissenters) of a particular issue is an extreme example of the concern expressed here. The focus is too much on the expert and too little on the characteristics of the evidence.

At other times, averaging of pieces of evidence is advocated. More weight is given to a certain input when there are more pieces of evidence supporting that value. Even if care is taken to ensure the pieces of evidence are independent, this process is flawed unless the evidence is fully relevant data. If the available evidence does not allow definition of limits of credibility (e.g., multiple researchers who have tested over the same subspace of the full application parameter space), then averaging this evidence will underrepresent the range of inputs that should be considered. On the other hand, values at the limits of credibility should not be given more weight relative to other values just because different models, developed at different institutions and by different researchers, all agree that certain values are at the limits of credibility. In addition, when evidence is sparse, the addition of a new piece of evidence (or the failure to include a piece of existing evidence) could significantly alter the average.

We use a rule-based logic structure to assign belief numbers to computed outputs. The belief value assigned to an output quantity is equivalent to the joint belief of all uncertain input choices. Note that this process treats the model as a black box, so it is generally applicable to any model regardless of whether it is simple or complex, linear or nonlinear, continuous or discontinuous, etc. The only issue is whether the "sampling density" of inputs is adequate to explore the important features and define the limits of credibility of the model response. It is important to note that the MBS makes no assumption concerning the statistical independence of input distributions because all combinations of defensible inputs are explored.

Table 2 summarizes the quantification for the joint belief of two uncertain inputs. The rule-based logic is motivated as follows:

If input A and input B are both physically reasonable, then any result (based solely on uncertainties with inputs A and B) must also be physically reasonable.

The joint belief of two inputs cannot be more than the belief of the "weakest link." If input A or input B (but not both) is judged physically unreasonable, then the joint belief of the input pair should be judged physically unreasonable.

If input A or input B is judged physically incredible (i.e., impossible), then the joint belief of the input pair must be judged physically incredible.

Table 2. Joint Belief of Two Uncertain Inputs

Input A	Input B		
	Physically Reasonable Bel(B) = 1.0	Physically Unreasonable Bel(B) = 0.1	Physically Incredible Bel(B) = 0.01
Physically Reasonable Bel(A) = 1.0	Physically Reasonable Bel(A, B) = 1.0	Physically Unreasonable Bel(A, B) = 0.1	Physically Incredible Bel(A, B) = 0.01
Physically Unreasonable Bel(A) = 0.1	Physically Unreasonable Bel(A, B) = 0.1	Physically Incredible Bel(A, B) = 0.01	Physically Incredible Bel(A, B) = 0.001
Physically Incredible Bel(A) = 0.01	Physically Incredible Bel(A, B) = 0.01	Physically Incredible Bel(A, B) = 0.001	Physically Incredible Bel(A, B) = 0.0001

Note that the numerical values for belief are chosen so that the physical descriptions of joint belief are recovered when the belief numbers are multiplied. For example, if Bel(A) = 0.1 and Bel(B) = 0.1 (i.e., each is physically unreasonable), then the joint belief is given by Bel(A, B) = Bel(A) × Bel(B) = 0.1 × 0.1 = 0.01. This belief value is interpreted as physically unreasonable, which is consistent with the definitions in the joint belief table (Table 2) and the table for assigning belief to inputs (Table 1). The procedure can be generalized to find the joint belief of three or more inputs as follows:

$$\mathrm{Bel(joint)} = \Pi_{i=1}^{N} \mathrm{Bel}(\mathrm{input}_i). \tag{1}$$

A second useful feature of the joint belief number is that it keeps track of how many inputs were judged physically unreasonable. For instance, if the joint belief of an input set, ($input_i$; I = 1, *N*), is 0.1, this means that only one of the *N* inputs was judged physically unreasonable; similarly, a joint belief of 0.0001 means that four of the *N* inputs were judged physically unreasonable. Retaining input values that are judged physically unreasonable is a reflection of our will to doubt; however, credibility is lost when many such physically unreasonable inputs are combined. As noted in Table 2 , the recommendation is to discount any interactions of physically unreasonable inputs and to interpret the output results based on joint belief numbers ≤ 0.01 as physically incredible. With this decision logic, the scale of belief numbers described in Table 1 can be used to construct physically based input quantifications and also to physically interpret output results. The discussion to this point (as will be the case throughout the report) has been in terms of characterizing and propagating belief distributions that are associated with model *parameters.* Although this is an important and common class of problems, the MBS is easily extensible to include model-form uncertainties and scenario uncertainties. For example, the analyst might be faced with a situation where there are two alternate plausible models that give widely differing results when extrapolated outside an existing validation database to an application parameter space. One model may perform substantially better against the existing validation database, but the second model cannot be positively excluded. With the MBS, a risk analysis can be performed exercising both models, but with one being belief-weighted higher than the other. A similar situation can occur with scenario definitions. The joint belief of an analysis involving scenario uncertainties, model-form uncertainties, and parameter uncertainties can still be calculated using Equation 1 and the results interpreted according to Table 1.

The regulatory language sometimes dictates that a requirement be satisfied with "high confidence." In the context of the MBS, all processes or inputs that would violate the requirement must be judged physically incredible, i.e., Bel ≥ 0.01. Alternatively, compliance to regulations might be judged based on expectation, i.e., mean or best-estimate values in other contexts. In the context of the MBS, all processes or inputs that would violate the requirement must be judged physically unreasonable or physically incredible, i.e., Bel ≥ 0.1.At this stage, it is worth noting that the numerical values assigned to the physical descriptors are somewhat arbitrary.

Table 3 shows an alternate (geometric) scale. Using the base scale for two input parameters (A, B) that are judged physically unreasonable, the joint belief is Bel(A, B) = Bel(A) × Bel(B) = (0.1) × (0.1) = 0.01, which is interpreted as physically incredible according to Table 3. Using the alternate scale, Bel(A, B) = (1/3) × (1/3) = 1/9, which is also interpreted as physically incredible. Thus, the real focus is on the physical descriptors and not on the numerical scale used to propagate information.

Table 3. Alternate Belief Scale

Belief Scale		Process Characteristics
Base	**Alternate**	
1	1	Physically reasonable
0.1	1/3	Physically unreasonable
0.01	1/9	Physically incredible

3 Algebraic Problem 5a from Epistemic Uncertainty Workshop

The challenge problem sets were crafted to focus on the issues of representation, aggregation, and propagation of uncertainty through mathematical models. With this as a goal, a simple algebraic model of the form $y = (a+b)^a$ was specified for the challenge problems so that computational issues do not unnecessary complicate the analyses. The description of challenge problem 5a is taken from Oberkampf et al. 2004 and repeated here for completeness.

The information concerning *a* is given by three independent sources of information. Each source specifies a closed interval A_i that contains the value of *a*, with all the intervals being consonant, that is, nested:

$$A_1 = [0.5, 0.7], \quad A_2 = [0.3, 0.8], \quad A_3 = [0.1, 1.0]. \tag{2}$$

The information concerning *b* is given by three independent sources of information. Each source agrees that *b* is given by a log-normal distribution: $\ln(b) = N(\mu, \sigma)$. Each source, however, specifies closed (nested) intervals, M_j and S_j, of possible values of the mean μ and standard deviation σ, respectively:

$$\begin{aligned} M_1 &= [0.6, 0.9], \quad M_2 = [0.1, 0.7], \quad M_3 = [0.0, 1.0], \\ S_1 &= [0.3, 0.45], \quad S_2 = [0.15, 0.35], \quad S_3 = [0.1, 0.5]. \end{aligned} \tag{3}$$

By intent, the challenge problems do not specify requirements for the decision context; however, real-world issues typically have such "regulatory requirements." Consequently, the framework for the algebraic challenge problems will be extended here by stating that values of $y > 3$ are considered unacceptable from a regulatory perspective. More specifically, it is assumed that regulatory requirements state that $P(y > 3) < 0.05$, that is, the probability that output *y* is greater than 3 is less than 0.05.

A simple algebraic model such as this provides the luxury of visually exploring the topographical features of the response function over the range of input parameters (see Figure 2). This will not be the case in most real-world applications where function evaluations are the product of thousands of lines of inscrutable code, where function evaluations are expensive, where there are many uncertain inputs, or where significant nonmonotonicity (in at least some of the uncertain inputs) should be anticipated. Although not obvious in Figure 2, the output *y* is slightly nonmonotonic to input *a* for the smaller values of input *b*. The green line denotes the regulatory threshold, i.e., $y = 3$.

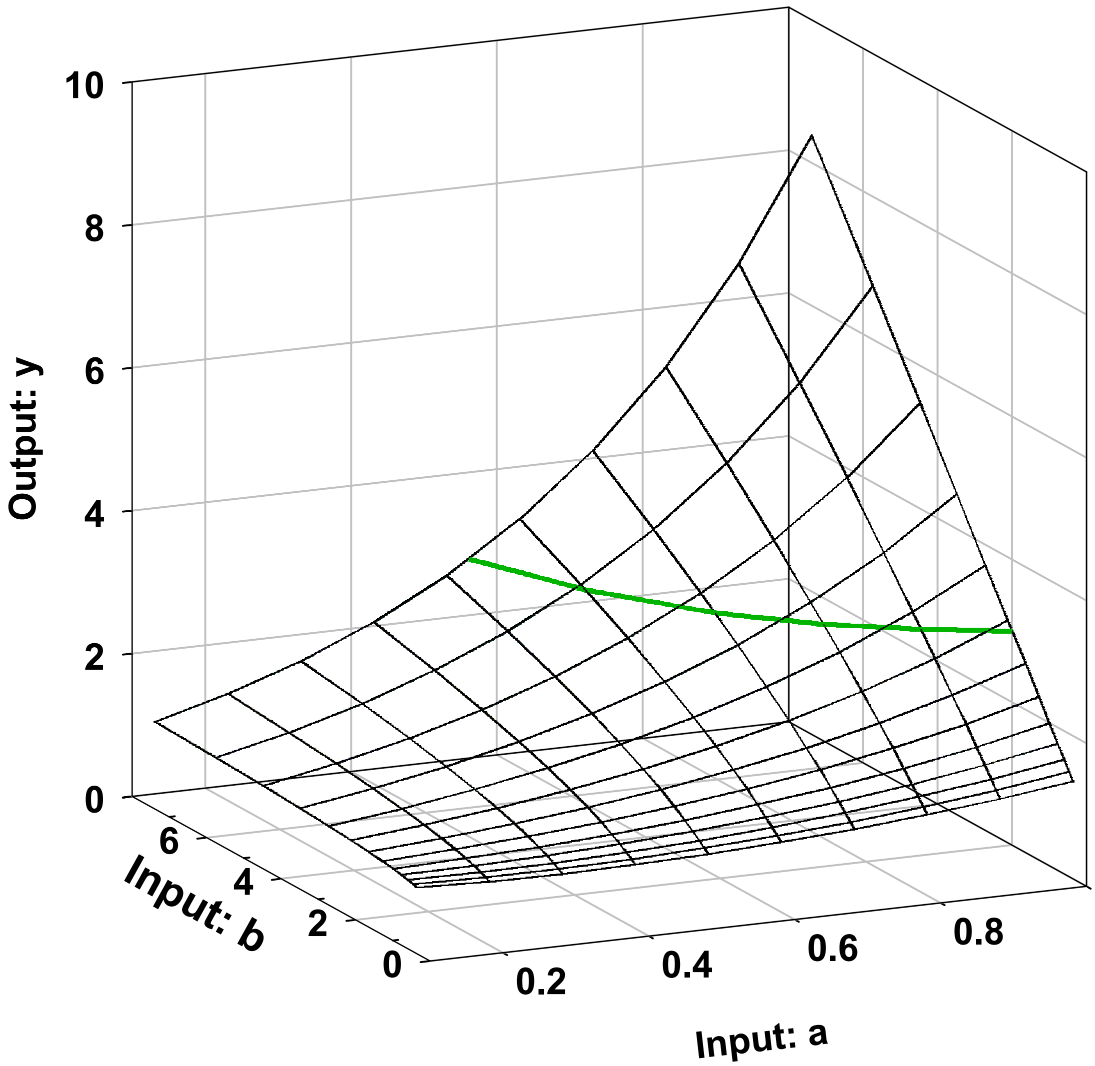


Figure 2. Topological features of the response function. The green line denotes the regulatory threshold, y = 3.

The statement of evidence for this challenge problem is nonexistent, and the aggregation of evidence according to the physically based belief scales is not possible. The sterile problem statement is biased to weighting schemes for aggregation, while the MBS tries to aggregate based on the *characteristics* of the evidence. The latter is not possible because the challenge problems (by intent) do not specify the rationale of numerical inputs, and this is precisely what must be examined to properly characterize the evidence. A second, and subtler, issue is that the challenge problems talk in terms of independent "sources," while belief scales talk in terms of independent "characteristics" of evidence. With belief scales, it is irrelevant if two independent researchers derive the same limit of credibility. What is relevant is that a defensible argument exists for a limit of credibility, not how *many* experts agree with the assessment.

Not being able to provide a physically based characterization of the evidence, we are faced with making an arbitrary quantification of inputs to demonstrate other features of the methodology. Figure 3 summarizes our choices for input distributions. Focusing here on inputs relevant to the MBS, graded belief was ascribed to inputs for *a*, *M*, and *S*. The total range for each input maps to the broadest interval from the problem specification, while the ranges for "physically reasonable" behavior arbitrarily correspond to the overlap region for all three sources. Sections 4 – 8 will now discuss the pros and cons of various underline approaches, with the emphasis on the WBS, to the challenge problem.

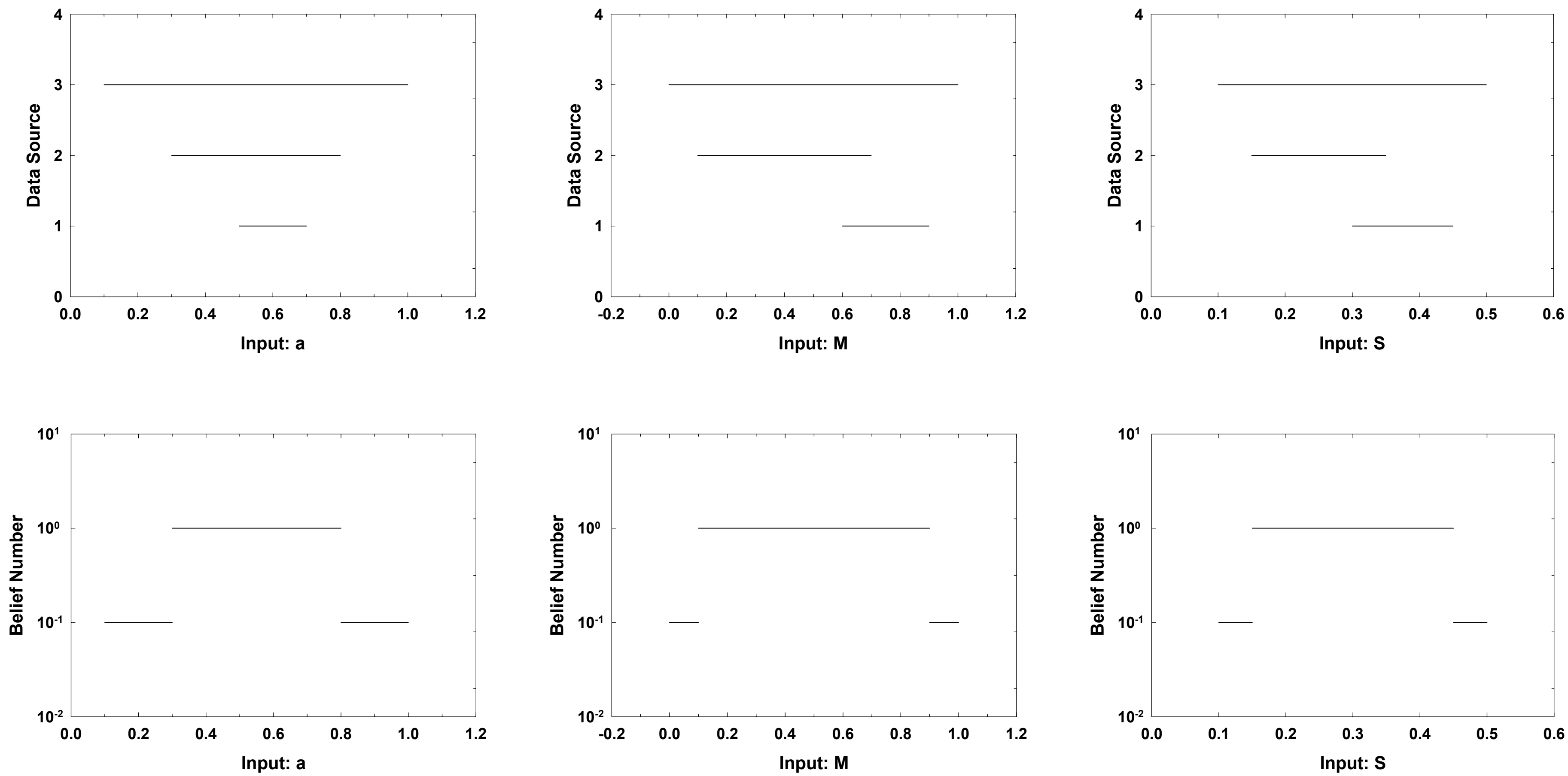


Figure 3. Arbitrary construction of belief distributions from three sources.

4 Deterministic Approach

We do not live in a deterministic world, and it is risky to think that tough (and controversial) regulatory issues can be resolved with a single decisive calculation. This is true regardless of whether the calculation is thought to be best estimate or bounding in nature. Variabilities, and most assuredly uncertainties, abound for any given issue, and they must be acknowledged and addressed if the assessments are to be credible. Applying safety factors to deterministic analyses is one strategy to build in robustness against variabilities and uncertainties. However, such an ad hoc method, which has been codified in some industries, may be inadequate or overly conservative (and more generally indefensible) for other high-consequence or first-of-a-kind issues.

Best-estimate analyses can *never* be decisive (even if the regulatory requirements are met) because these analyses leave unanswered the question of how much impact variabilities and uncertainties have on the computed result. Nominal or mean values for inputs are used for best-estimate analyses; but for nonlinear problems, the computed result based on mean inputs will not be the same as the mean result computed by propagating distributions of inputs through the model (even when the input distributions are symmetric). Thus, best-estimate inputs do not generally produce results that are at the mean of the distribution of output if the distribution were calculated.

Bounding analyses that convolute the extremes with the extremes of the inputs and that pursue a self-consistent fringe of conservative scenarios, although noble in intent, can be an unsuccessful or even an undesirable strategy. Heavy reliance on expert judgment is often required to define scenarios, model forms, and model parameters that produce extreme outputs; such judgments, however, are commonly flawed for highly nonlinear coupled multiphysics problems in complex geometry, which are generally the circumstances surrounding many high-consequence decisions of interest. Confidence is derived from demonstrating an understanding of system behavior, which cannot be justified when judgment-based assertions sweep all other potential states under the rug of assumed “limiting behavior.” In general, extreme responses may *not* be associated with extremes of input, such as resonance behavior, and thus the parameter space must be more densely explored unless it is well established that the output has a known monotonicity to all inputs. Alternatively, confounding high-confidence inputs (e.g., at the 95% level) can produce an overly conservative result (e.g., 99.99%) on an actual output distribution if the distribution were computed. Lastly, a rigorous pursuit of bounds often focuses precious resources on extreme states by diverting resources away from more credible states.

Bounding analyses are vulnerable to some tough programmatic questions when the requirements are not met. Why were so many resources expended pursuing extreme states? Would more “rational” analyses meet the requirements? What uncertainties contribute most to uncertainties in the regulatory metrics? How should remaining resources be expended?

For the challenge problem described in Section 3, the function was evaluated for both nominal and high-confidence (bounding) inputs as noted in Table 4 and in Figure 4. Table 5 shows that the regulatory threshold (y = 3) is violated by a wide margin when high-confidence (bounding) inputs are employed, while an analysis using nominal inputs produces a result that is well below the regulatory threshold.

A great appeal of deterministic analyses is that only one or two function evaluations are required. This computational efficiency comes at the expense of understanding the sensitivity of key outputs to known variabilities and uncertainties. Also lost is the ability to compare intuition or insights with predicted model behavior.

Table 4. Analysis Process for Deterministic Method

Case	Analysis Process
Deterministic	**High-Confidence (Bounding) Inputs** 1. Fix a = 1.0 corresponding to its maximum value. 2. Fix M = 1.0 corresponding to its maximum value. 3. Fix S = 0.5 corresponding to its maximum value. 4. Evaluate ln(b) = 1.8224 corresponding to the 95th percentile of $N(M, S)$. 5. Evaluate y_{hc} = 7.1867 **Nominal Inputs** 1. Fix a = 0.55 corresponding to its midrange value of Bel = 1.0. 2. Fix M = 0.5 corresponding to its mid-range value of Bel = 1.0. 3. Fix S = 0.3 corresponding to its midrange value of Bel = 1.0. 4. Evaluate ln(b) = 0.5 corresponding to the 50th percentile of $N(M, S)$. 5. Evaluate y_{nom} = 1.5424.

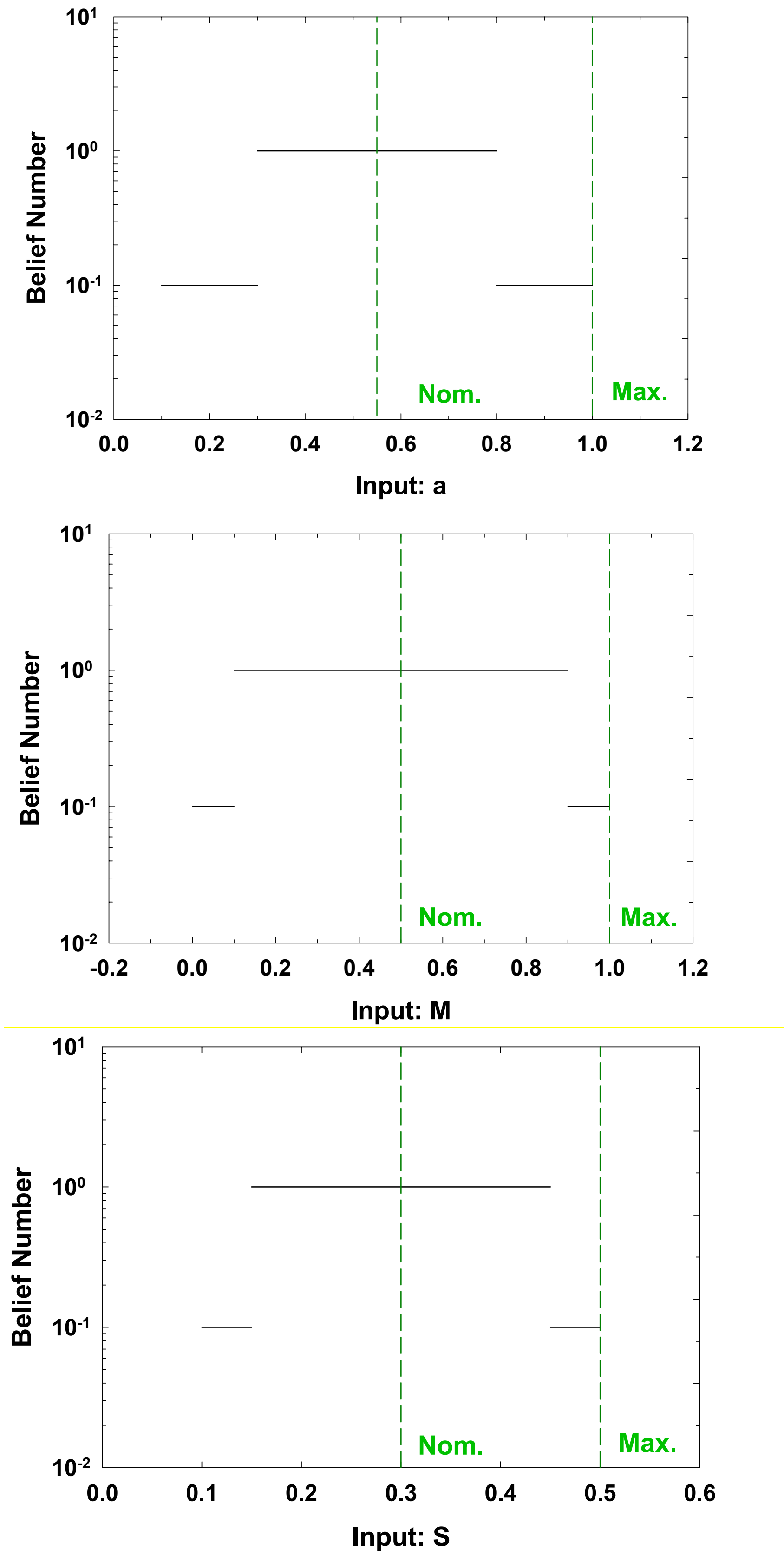


Figure 4. Nominal and upper-bound inputs (denoted by green lines and labels) for deterministic analyses.

Table 5. Analysis Results for Deterministic Method

Case	Regulatory Decision	Sensitivity	# FEVs
Deterministic	Bounding analysis: y = 7.1867 (> 3) • Regulatory requirement not met Nominal analysis: y = 1.5424 (< 3) • Regulatory requirement met	Not Applicable	2

5 Bayesian (and Frequensic) Methods

Bayesian methods are part of the rich tradition of probabilistic methods. Variabilities and uncertainties are both represented by *prior* (probability) distributions. Non-informative priors (i.e., uniform distributions for continuous input and equal weights for discrete inputs) are commonly employed when information is sparse. Input distributions are forward propagated using traditional probabilistic methods such as Monte Carlo. Bayesian methods also provide formalism for updating *priors* for input distributions based on model comparison with new observed system behaviors.

Input distributions for the MBS are not additive as required by the more traditional probabilistic analyses considered here. To facilitate comparisons, the MBS distributions of Section 3 were "normalized" to provide appropriate input distributions for the probabilistic analyses. This is done by specifying trilinear cumulative probability distributions (noted in Figure 5 by blue lines and labels) such that 90% of the probability is contained in the interval where behavior was judged "physically reasonable." To proceed further with an evaluation, we must assume that the parameter distributions are all statistically independent, even though there is no evidence in the problem statement to support this assumption.

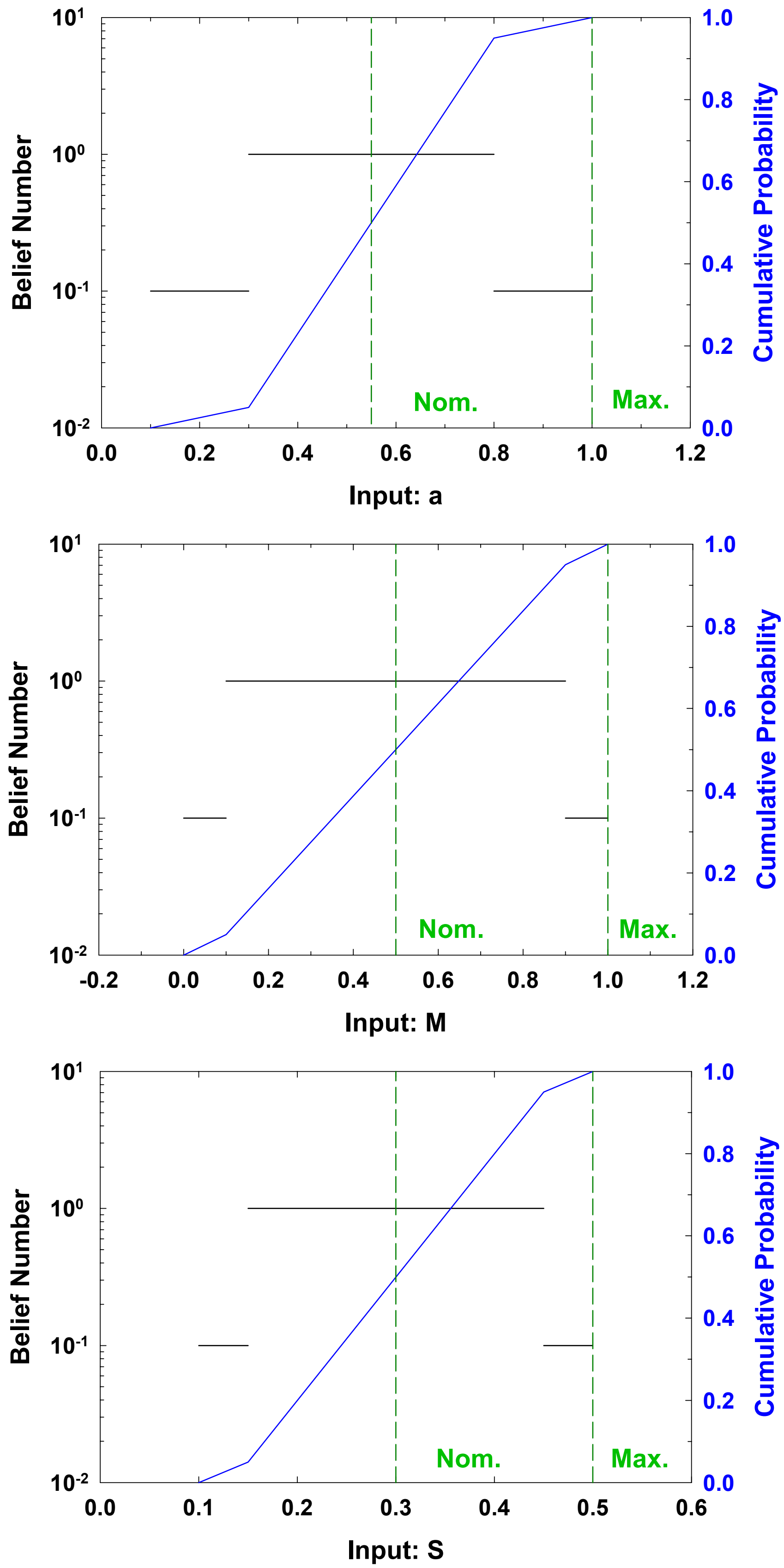


Figure 5. Input distributions for Bayesian analysis.

We note at this point that there is a family of distributions for ln(*b*) because of the uncertainty in the mean and the standard deviation of the distribution. This is illustrated in Figure 6 where 25 realizations of the ln(*b*) distribution are shown corresponding to 25 {*M*, *S*} doublets derived from 25 independent Monte Carlo samples of *M* and *S*. To proceed with a Bayesian description of ln(*b*), we must collapse this family of distributions to a single effective distribution that reflects both the variability of ln(*b*) and the uncertainty in its mean and its standard deviation. This is commonly done by averaging the cumulative probabilities vertically for a fixed values of ln(*b*). The "effective" distribution for ln(*b*) is shown in Figure 6 as the blue curve, and this distribution is also normal with $M' = 0.50158$ and $S' = 0.41140$. These are only estimates because we have averaged across a finite sample of the infinite family of ln(*b*) distributions. We note that the estimated mean of the effective distribution is essentially equal to the expected value (0.5) of the distribution for *M*. This is not the case with the estimated standard deviation for the effective distribution (0.41140), which is 37% larger than the expected value of *S* (0.3); consequently, uncertainty is represented here as an enhanced variance in the ln(*b*) distribution.

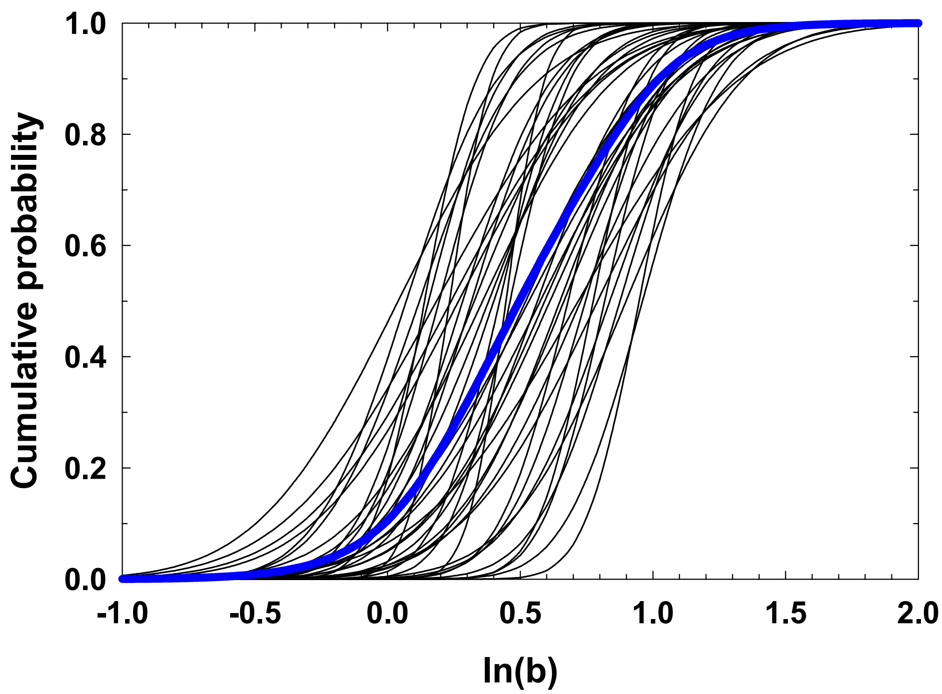


Figure 6. Effective distribution of ln(*b*) reflecting both variability in ln(*b*) and uncertainty in the mean and the standard distribution for ln(*b*).

Figure 7 shows the results of the Bayesian analysis conducted according to the process outlined in Table 6. How then should a decision maker interpret these results? The probability of exceeding the regulatory threshold is 0.0149, which is less than the requirement that $P < 0.05$. A naive decision maker might be inclined to interpret Figure 7 as saying that the system, tested a very large number of times, would yield the distribution of results shown in Figure 7 and that only 149 failures would be expected in

10,000 system tests. This frequensic interpretation would be incorrect (even if the model is indisputably valid) because none of the inputs were based on a statistically significant database (fully relevant to the application), which is required for a frequensic interpretation of inputs and outputs.

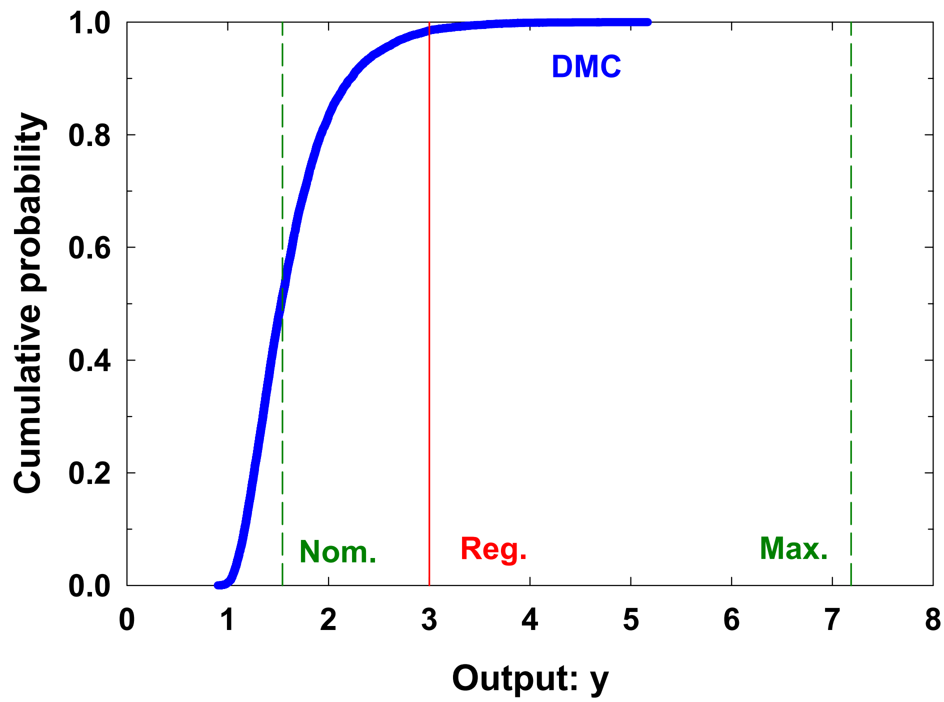


Figure 7. Results of analyses using frequensic or Bayesian methods.

Table 6. Analysis Process for Bayesian Method

Case	Analysis Process
Bayesian	1. Randomly sample *a* and (effective) ln(*b*) 10,000 times. 2. Evaluate y_i for each input doublet $\{a_i, b_i\}$. 3. Construct a cumulative probability distribution for the ensemble of 10,000 y_is. 4. Evaluate the probability that $y > 3$.

A Bayesian would train the decision maker to interpret both the input and output distributions as measures of belief rather than frequency. Consequently, the decision maker would conclude that his belief in output values exceeding "3" should be very low, which is (at least qualitatively) the same conclusion he would reach with the frequensic interpretation, only now for the correct reasons. Training the decision maker is always critical, but the dual use of "probability" to represent frequency as well as belief will always be a source of confusion to decision makers and many analysts as well.

Bayesian methods represent objective information (frequency distributions) and subjective information (belief distributions) with traditional probability distributions, and these methods propagate such distributions using traditional probabilistic methods such as Monte Carlo. Bayesian methods have been criticized because the relative contributions of objective and subjective inputs cannot be examined separately. A decision maker examining the results of Figure 8 cannot, in general, discern whether the results are based on high-quality objective information or on unsubstantiated judgment for the inputs.

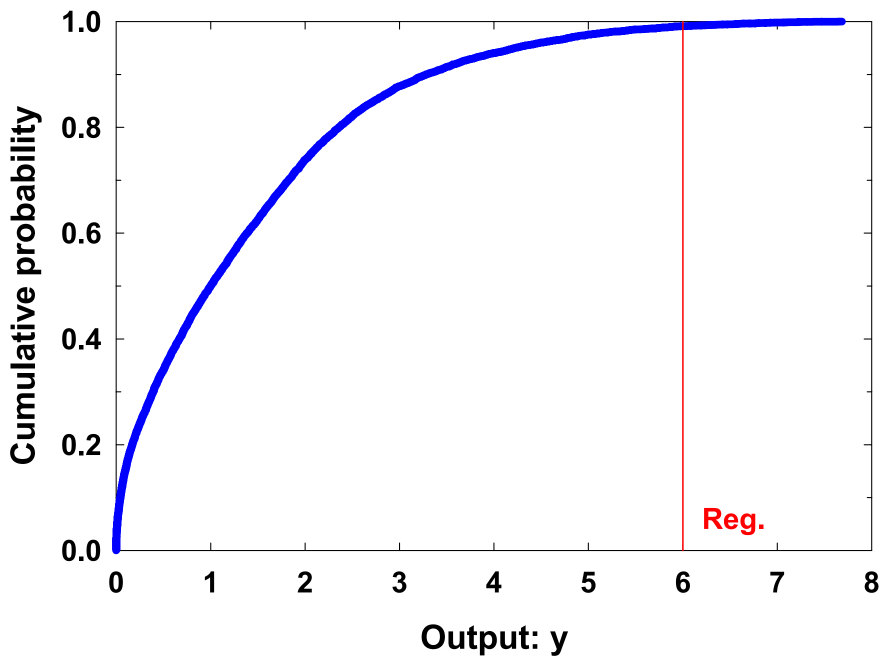


Figure 8. Function $y = b^a$ subject to probabilistic propagation of inputs a = [1, 3] and b = [0,1] sampled from their respective intervals.

In the more general case where some inputs are objective (frequensic) and some are subjective (belief), the interpretation of results becomes even more ambiguous. Regardless of the interpretation (frequensic or belief), it is also ambiguous as to whether the regulatory requirement is actually met. Was it the intent of the requirement that objective (frequensic) data or subjective beliefs form the basis for acceptance? The regulator may not know what language to use in hard problems. This ambiguity in the regulatory language is common and will be discussed further in Section 9.

Bayesian methods are well established, and they have been used in some high-consequence decisions, such as for dikes and seawalls (Speijker et al. 2000), outside the United States. One drawback of Bayesian methods is that they are not *information preserving*, that is, they can imply more belief about computed results than is warranted based on supporting evidence. This is best illustrated by considering an even simpler algebraic problem, $y = b^a$, so that the arithmetic is absolutely transparent. The function is positive monotonic in both inputs. Assume that available evidence can only support that input values for $a = [1, 3]$ and $b = [0, 2]$ are within the specified intervals. The distribution of output values shown in Figure 8 can be obtained by assuming uniform probability distributions on the input intervals (i.e., invoking the maximum entropy principle) and by propagating inputs through the model using any probabilistic technique like Monte Carlo. Consider that output values greater than 6 are "bad" from a regulatory perspective. How should the decision maker interpret the results?

The decision maker might conclude that $P(y > 6) = 0.0089$ and that this conclusion is "conservative" because inputs were characterized by uniform distributions. After all, the maximum entropy principle states that uniform distributions embody the least information of all possible probability distributions. We suggest that this conclusion is logically inconsistent with the available input information. Any output on the interval [0, 8] is *physically reasonable*, and there is no basis to judge some outputs as more "likely" or more "believable." As a note, the MBS applied to this problem produces the logically consistent result.

Bayesian methods provide a well-accepted formalism for updating input distributions and updating computed results as more information becomes available. Consequently, the Bayesian method is best viewed as an iterative process that is assumed to converge to the correct distribution of output when a sufficient mix of relevant input and output information becomes available. The presumption is that a statistically significant characterization of dominant inputs is even possible given practical constraints on time, money, or even technology. Unfortunately, the first- and other low-order passes through the process might not produce conservative results.[3] Even worse, if only one pass through the process is possible, there is no way of telling whether the results are conservative or not. The logical extension is to insist that "issues" addressed through Bayesian methods must be iterative and that sufficient cycles be exercised to ensure a converged result, i.e., a decision to use Bayesian methods comes with a programmatic mortgage. In reality, it often takes a heroic effort involving large resource expenditures (time and money) to perform a single peer-reviewed assessment of a complex issue. A decision maker might find little motivation (particularly if the process is schedule driven) to justify the continuance of large resource expenditures based on a frequensic or Bayesian interpretation of the results presented in Figure 8.

Bayesian updating can be very computationally expensive and is most justified when initial information is sparse but substantial new information is expected within the time frame of the study. The value of Bayesian updating is not justified when initial information is sparse and only limited new information may become available. In this case, the assumed "prior distributions" will dominate and will not be significantly altered

[3] If a statistically significant database was subsequently generated to characterize the *actual* values (or distributions) for the inputs to the $y = b^a$ problem, then it is possible (and still consistent with the bounding intervals) that these quantified inputs would lie close to the upper bounds of the original intervals. Consequently, the probability of exceeding the regulatory threshold could be substantially higher than the initial estimate using uniform distributions.

by limited new information. Likewise, the value of updating is not justified when substantial information exists at the outset. In this case, any new information is only confirmatory and can have little impact on the “priors.” In many applications, the only practical path forward may be to perform final assessments by forward propagating the best available information without first exercising the Bayesian formalism for updating the priors.

It is useful to compare the distribution of results derived here for the challenge problem with the nominal and high-confidence point values derived in Section 4 and noted in green in Figure 7. First, we note that the deterministic output associated with nominal inputs ($y = 1.5429$) corresponds to the 51.3 percentile of the full distribution shown in Figure 6. Here, deterministic output ($y = 1.5424$) based on nominal inputs very nearly equals the median output ($y = 1.5294$) associated with the full distribution. The deterministic output ($y = 7.1867$) based on high-confidence inputs corresponds to a hyperconservative point ($P >> 99.99$th percentile) on the full distribution.

Monte Carlo methods were used here as the tool of choice for probabilistic propagation. Such methods are considered the “gold standard,” but they are not computationally efficient, particularly when regulatory requirements are established in terms of low probability levels. The current results are based on 10,000 Monte Carlo simulations, and, as such, can be viewed as providing a converged estimate of $P(y > 3)$. Such an enormous number of function evaluations are not possible for the codes and applications that are being addressed by the ASC program.

The probability that output y is greater than three, i.e., $P(y > 3)$, can be calculated using a smaller number of Monte Carlo samples, like 100, that is more representative of ASC applications. Of course, this is only an estimate of $P(y > 3)$, and different ensembles of 100 samples each would produce a slightly different estimate of $P(y > 3)$, as noted in the top part of Figure 9. Given 25 estimates of $P(y > 3)$ based on 100 samples per estimate, we can estimate our confidence (as a statement of numerical convergence) that the regulatory requirement is met (see bottom of Figure 9). This is analogous to solution verification in the discretized solution of partial differential equations. In this case, we are ~100% confident that $P(y > 3)$ is less than 0.05, and we would thus conclude that the requirement has been met.

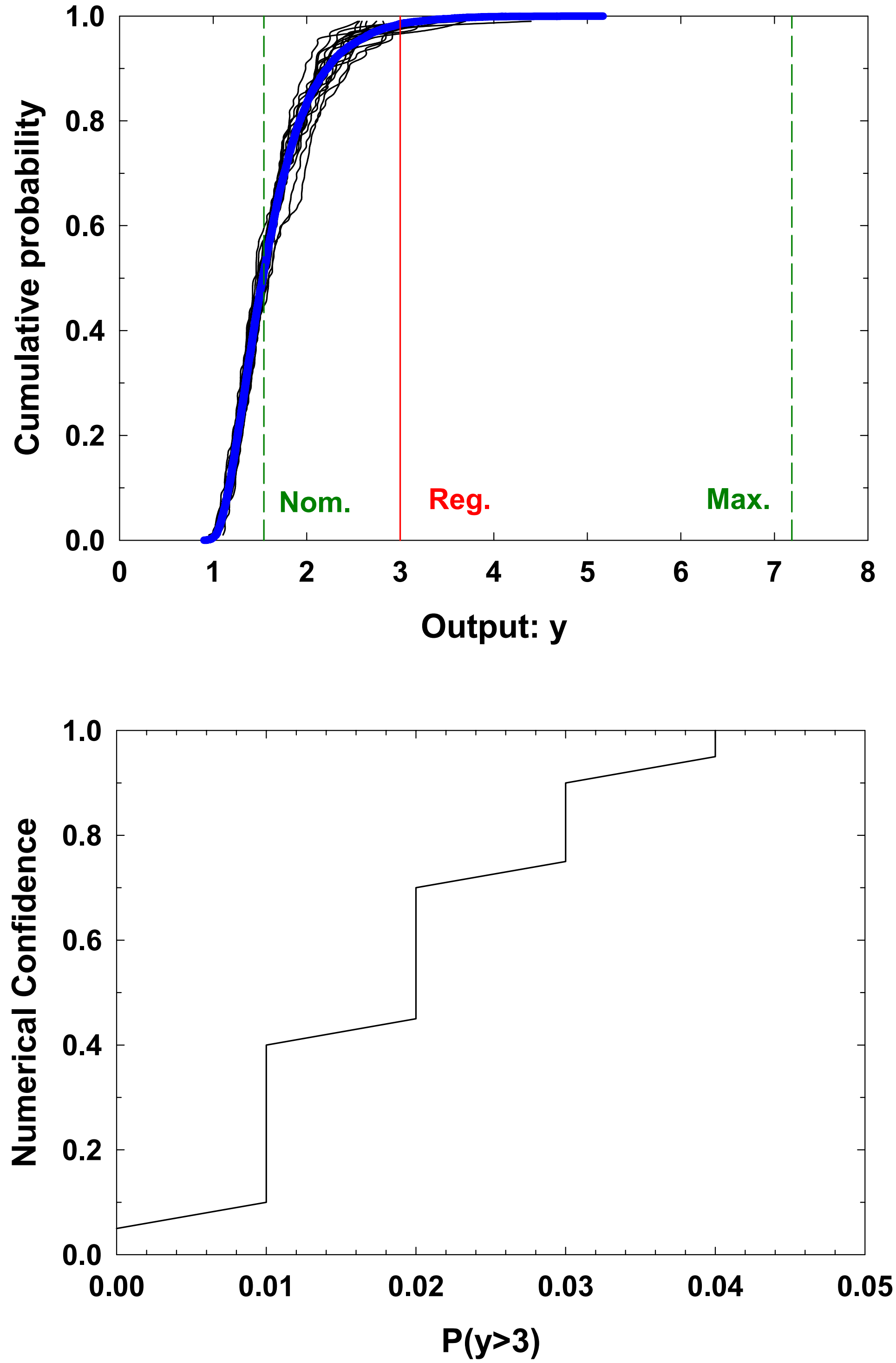


Figure 9. Confidence in estimates of $P(y > 3)$ for undersampled distributions.

Two concerns should be immediately obvious. What if the regulatory requirement was set at 10^{-6} instead of 0.05 and only 100 function evaluations could be afforded? One possible path forward would be to fit an assumed distribution, e.g., a normal distribution, to the 100 results and calculate $P(y > 3)$ under the assumption of this distribution form. One should (as a postprocessing step requiring no new function evaluations) account for uncertainties in the computed probability resulting from standard errors in the mean and the standard deviation of the fit distribution. More importantly, representational errors have been introduced by assuming a normal distribution; and other distributions such as a beta distribution might reasonably fit the 100 results, which could have a dominant impact on the estimated $P(y > 3)$. However, with time and effort and no new function evaluations, the uncertainty in $P(y > 3)$ due to these representational errors can also be quantified.

As a second concern, we have stated that we had a computational budget of 100 function evaluations; but in order to estimate our confidence in $P(y > 3)$, we have executed 25 ensembles of 100 function evaluations each! This, of course, is not generally practical; however, bootstrap methods (Davison and Hinkley 1997) can be used to estimate confidence in $P(y > 3)$ with no new function evaluations beyond the original 100. In bootstrap, a statistically significant number of ensembles of 100 each are created by sampling from the original set of 100 *with replacement*. Each ensemble provides a different estimate of $P(y > 3)$, and the set of such estimates can be used to estimate confidence in the computation of $P(y > 3)$. If required, an assumed distribution can be fit to the results of the original ensemble of 100 function evaluations, and the bootstrap process can be executed by repeating the sampling process from the assumed distribution. As before, this introduces representational errors that must also be quantified, but this can be done with no new function evaluations.

Alternate methods for probabilistic propagation exist; however, for highly nonlinear problems involving large numbers of uncertain inputs, the computational burden for these more sophisticated methods can also be enormous relative to what is practical. In any case, the issue of numerical convergence must be addressed with any method.

One common strategy is to replace the physics model with a surrogate model such as a response surface. The response surface, which is cheap to evaluate, can then be evaluated as many times as the particular propagation strategy requires. As an example of these methods, consider a simple second-order polynomial with all interaction terms. Such a model (where N is the number of terms in the polynomial) requires as a minimum $(N^2 + N + 2)/2$ function evaluations (and more is much preferable) to evaluate the constants. Consequently, this technique can accommodate at most 13 inputs if the computational budget is 100 full physics evaluations. Response-surface functions with so many terms are not very intuitive and often exhibit undesirable behavior. As an alternate strategy, function evaluations can be used to support a sensitivity study with a goal of (hopefully) identifying a small number of inputs that dominate uncertainty in the outputs. A response surface can then be constructed around this reduced set of inputs. This latter strategy is also effective when the number of uncertain inputs is too large to uniquely determine all the fitting constants needed for a surrogate model with the affordable computational budget.

Surrogate models are important tools in risk analyses; however, such tools may not be effective if the range of inputs is broad and the behavior is complex, highly nonlinear, significantly nonmonotonic, or discontinuous. In all cases, surrogate models introduce additional "representational" errors that should be quantified.

The current results are based on 10,000 Monte Carlo simulations and, as such, can be viewed as providing a well-converged estimate of $P(y > 3)$. Such an enormous number of function evaluations are not possible for the codes and applications that are being addressed by the ASC program. Another challenge to the computational budget exists if there is some need to change one or more of the input distributions, e.g., an error is discovered in the input distributions or new information becomes available. In this case, the ensemble of calculations must be repeated to get a credible and consistent result.

Sensitivity of the output of interest to the various uncertain inputs is important to understand. Such information can be used to efficiently focus research activities. The simplest, and often useful, approach is to first create scatterplots of the output versus each of the uncertain inputs. The results for the Bayesian approach are shown in Figure 10. The scatterplots suggest that the output y is monotonic to both input a and input b. We know from the topology of the output function, as shown in Figure 2, that the output function is slightly nonmonotonic in input a for small values of input b, but this small nonmonotonicity cannot be resolved in the scatter plots. Sensitivity scatterplots are powerful for several reasons: (1) they provide a basis to judge model consistency, (2) they provide a subjective check on the assumptions of more-quantitative sensitivity techniques, and (3) they require no new function evaluations above what was needed to make the regulatory assessment. The effectiveness of the human eye at determining patterns also contributes to the powerfulness of these scatterplots.

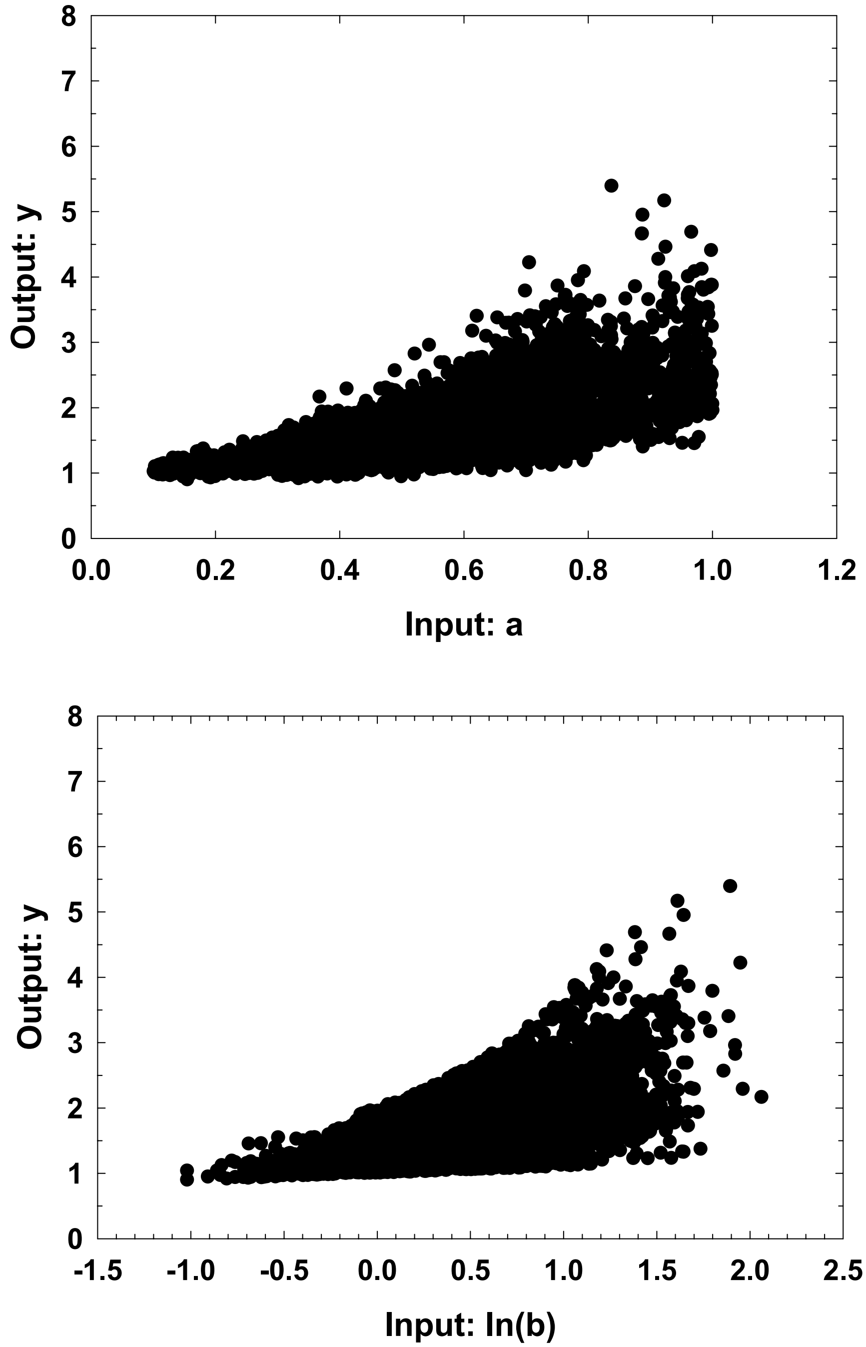


Figure 10. Sensitivity analysis for Bayesian approach.

Because the output y appears linearly monotonic to inputs a and b, we can employ more quantitative techniques to assess sensitivity based on linear regression and correlation (Saltelli et al. 2000). The relative importance of predictor variables (x_i) is obtained by looking at their coefficients (β_i) in a linear regression equation (i.e., 1st order polynomial) where all the variables are in their standardized form (i.e., referenced to the mean and normalized by the standard deviation):

$$\frac{y-\bar{y}}{\sigma_y}=\sum_i \beta_i \frac{x_i-\bar{x}_i}{\sigma_{xi}} . \tag{4}$$

The beta coefficients (β_i) are equivalent to the partial correlation coefficients, $r(y, x_i)$. The square of a partial correlation coefficient physically represents the fraction of the variability in output y that can be explained by a linear association with input x_i; consequently, the relative importance of any two predictor variables can be obtained by taking the ratio of the squares of their respective βs.

The assumption of linearity can be checked quantitatively by computing the coefficient of determination,

$$R^2=\frac{\sum_i (\hat{y}_i-\bar{y})^2}{\sum_i (y_i-\bar{y})^2} , \tag{5}$$

where y_i is the raw response at x_i and $\hat{y}_i$ is the predicted response using the linear regression equation noted above. The quantity R^2 gives the proportion of total variation in the output y that can be explained by a linear regression through the predictor variables.

Values of R^2 near unity imply that the output y is well represented as a linear combination of the uncertain inputs; and for the current problem (see Table 7), the assumption of linearity is justified. Table 7 shows that the sensitivity of output y to uncertainties in input a is only slightly greater than the sensitivity of output y to uncertainties in input ln(b).

Table 7. Analysis Results for Bayesian Method

Case	Regulatory Decision	Sensitivity	# FEVs
Bayesian	$P(y > 3) = 0.0149$ Requirement met without confidence statement	Sensitivity of y to a, ln(b) $R^2 = 0.882$ $\beta_a^2 = 0.491$ $\beta_{\ln(b)}^2 = 0.391$	10,000

In summary, Bayesian methods are not information preserving, initial iterations may not be conservative, and it is difficult to prove convergence to the right answer. Probabilistic propagation requires a large number of function evaluations to ensure numerical convergence, particularly if very small probability levels must be resolved. There are options for performing analyses with a limited computational budget, but the numerical convergence error and any representational errors must be evaluated and may ultimately dominate the results. Any single function evaluation is meaningless. It is only in the context of an ensemble, i.e., a distribution, that meaning can be derived.

6 Second-Order Methods Using Subjective Probabilities to Represent Epistemic Uncertainties

When the distinction between variabilities and uncertainties is not maintained, the deleterious events associated with the system, the likelihood of such events, and the confidence with which both likelihood and consequences can be estimated can become commingled in a way that makes it difficult to draw useful insights (Helton et. al, 2000). Second-order methods are one approach to address this concern. Second-order methods are founded on an explicit conceptualization of an ensemble of frequency (output) distributions: one output (frequency) distribution for each realization of uncertain inputs. Variabilities on inputs contribute to the output frequency distribution, and uncertainties on inputs or model form are associated with the multiple realizations of (output) frequency distributions. Consequently, the impact of variabilities and uncertainties is represented separately and should be elements of any high-consequence decision process.

The reassessment of the risk from nuclear power plants conducted by the U.S. Nuclear Regulatory Commission (NUREG-1150) provides an example of a very large analysis in which an extensive effort was made to separate stochastic and subjective uncertainty (Helton and Breeding, 1993) This analysis was instituted in response to criticisms (Lewis et. Al, 1978) that the Reactor Safety Study (US NRC, 1975)) had inadequately characterized the uncertainty in its results Similarly, the Environmental Protection Agency's (EPA's) standard for the geologic disposal of radioactive waste[4] can be interpreted as requiring (1) the estimation of a complementary cumulative distribution function (CCDF), which arises from the different disruptions that could occur at a waste disposal site and is thus a summary of the effects of stochastic uncertainty, and (2) the uncertainty associated with the estimation of this CCDF, with this uncertainty deriving from a lack of knowledge on the part of the analysts involved and thus providing a representation for the effects of subjective uncertainty.

Reactor safety and geologic disposal of radioactive wastes are both high-consequence issues in the highest national interest. Performance assessment (risk analysis) was central to both of these activities in establishing that regulatory guidelines and requirements were met. Sandia led the performance assessment efforts for both applications, which were subject to the highest level of national peer review. Thus, second-order methods represent the mainstream of risk analyses for high-consequence, complex systems, and Sandia is in a unique position to adapt these technologies to the needs of its weapons certification efforts.

In the second-order approach as applied to date, uncertain (epistemic) inputs are represented by subjective probabilities (subject to the usual rules of probability distributions) that are propagated using traditional probabilistic methods. The representation and propagation of epistemic uncertainties using traditional probabilistic

[4] Waste Isolation Pilot Plant (WIPP) (Helton et al. 2000)

methods is similar to the Bayesian approach; only here, the contributions of variabilities and uncertainties are separated so that the contribution of each is made explicit.

For the current problem, the same trilinear cumulative probability distributions constructed for the Bayesian analysis were used (see Figure 5). Figure 11 shows the results of the second-order analysis conducted according to the process outlined in

Table 8. The analysis produces an ensemble of frequency distributions (represented as cumulative distribution functions [CDFs]), with one CDF for each representation of uncertainty. The dense clustering of CDFs corresponds to regions of higher belief. For

Case	**Analysis Process**
Second-order analysis using subjective probabilities to represent uncertainties	1. Sample *a, M,* and *S* 100 times using direct Monte Carlo sampling to create a set of input triplets reflecting uncertainty. a. Sample ln(b) 1,000 times (conditional on one realization of *M* and *S)* to create an input set representing variability. *b.* Evaluate *y* for each input set representing variability (conditional on one realization of *a, M,* and *S).* *c.* Construct a cumulative probability distribution for *y* (conditional on one realization of *a, M,* and *S)* *d.* Evaluate the probability that $y > 3$ from the cumulative distribution (conditional on one realization of *a, M,* and *S)* 2. Repeat steps 1a to 1d for each of the 100 realizations of uncertainty. *3.* Construct a cumulative probability distribution for $P(y > 3)$.

each frequency distribution, the probability of exceeding the regulatory requirement can be recorded, and a distribution of such values can also be created. Consequently, we can state that we are 95% confident that $P(y > 3)$ is less than the regulatory requirement of 0.05.

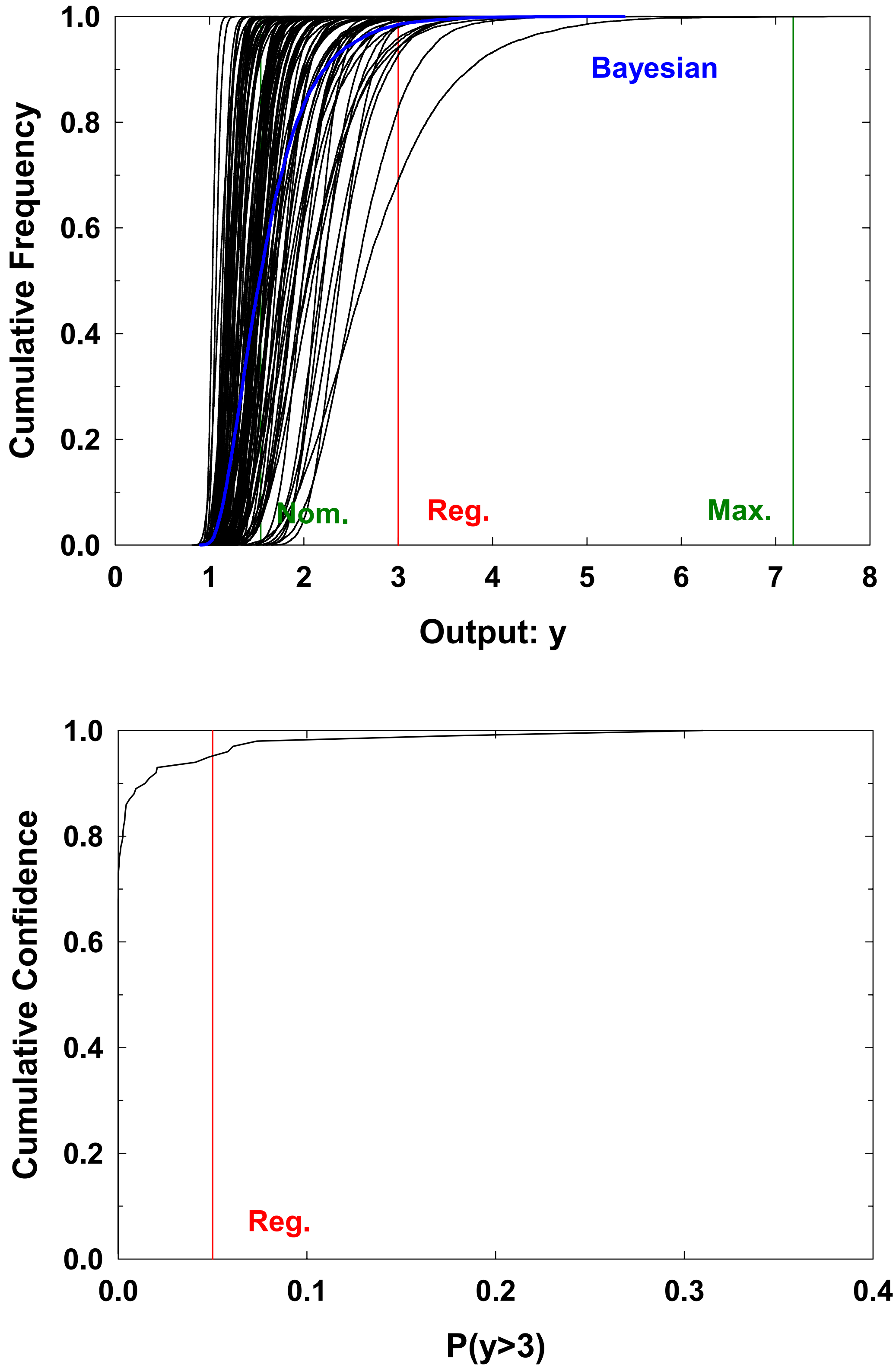


Figure 11. A second-order approach using subjective probabilities to represent epistemic uncertainties separately from aleatory uncertainties.

Table 8. Analysis Process Associated with Second-Order Analysis Using Subjective Probabilities to Represent Uncertainties

Case	Analysis Process
Second-order analysis using subjective probabilities to represent uncertainties	4. Sample *a*, *M*, and *S* 100 times using direct Monte Carlo sampling to create a set of input triplets reflecting uncertainty. e. Sample ln(*b*) 1,000 times (conditional on one realization of *M* and *S*) to create an input set representing variability. f. Evaluate *y* for each input set representing variability (conditional on one realization of *a*, *M*, and *S*). g. Construct a cumulative probability distribution for *y* (conditional on one realization of *a*, *M*, and *S*) h. Evaluate the probability that $y > 3$ from the cumulative distribution (conditional on one realization of *a*, *M*, and *S*) 5. Repeat steps 1a to 1d for each of the 100 realizations of uncertainty. 6. Construct a cumulative probability distribution for $P(y > 3)$.

Although aleatory and epistemic uncertainties are treated separately, the epistemic portion is characterized by subjective probability distributions that are propagated through the model using probabilistic methods. The criticism leveled against the Bayesian approach is applicable here as well; that is, such methods are not *information preserving*.

The ensemble of frequency distributions provides a more complete description of uncertainty compared to the previous Bayesian analysis. Figure 11 shows that there are many frequency distributions substantially to the right of the Bayesian distribution (noted in blue). For the Bayesian distribution, $P(y > 3) = 0.0149$. Here, approximately 5% of the frequency distributions produce $P(y > 3) > 0.05$; and for the most extreme frequency distribution, $P(y > 3) = 0.3$! It is thus critical to properly characterize the confidence or "belief" in such extreme distributions. Despite the expansive range of frequency distributions, none noticeably intersect the value ($y = 7.1867$) obtained from a bounding deterministic analysis, again illustrating how seemingly simple bounding analyses can produce hyperconservative results.

Sensitivity of the output y to uncertainties in inputs a, M, and S is addressed through a series of scatterplots shown in Figure 12. Sixty-eight percent of the frequency distributions had no values in excess of the regulatory threshold of $y = 3$. These values have been excluded from the scatterplot and the following regression analysis. The value of R^2 is small (Table 9), which suggests that $\log_{10}[P(y > 3)]$ is not well represented by a linear function of each of the uncertain inputs.

Case	Regulatory Decision	Sensitivity	# FEVs
Second-order analysis using subjective probabilities to represent uncertainties	High confidence, $P_{95}(y > 3) = 0.048$ • Requirement met Expectation, $P_{50}(y > 3) \sim 0$ • Requirement met	Sensitivity of $\log_{10}[P(y > 3)]$ to a, M, S $R^2 = 0.295$ $\beta_a^2 = 0.175$ $\beta_M^2 = 0.101$ $\beta_S^2 = 0.019$	1 million

Qualitatively from the scatterplots and quantitatively (within the limitations of the linear regression analysis), we see that uncertainties in input a are more important than uncertainties in either of the distribution parameters M or S and that uncertainties in M are more important that uncertainties in S within the distribution for ln(b). Knowledge that input b can be represented as a lognormal distribution adds a lot of information, and the sensitivity analysis suggests that overall risk is best reduced by focusing on improving information about input a.

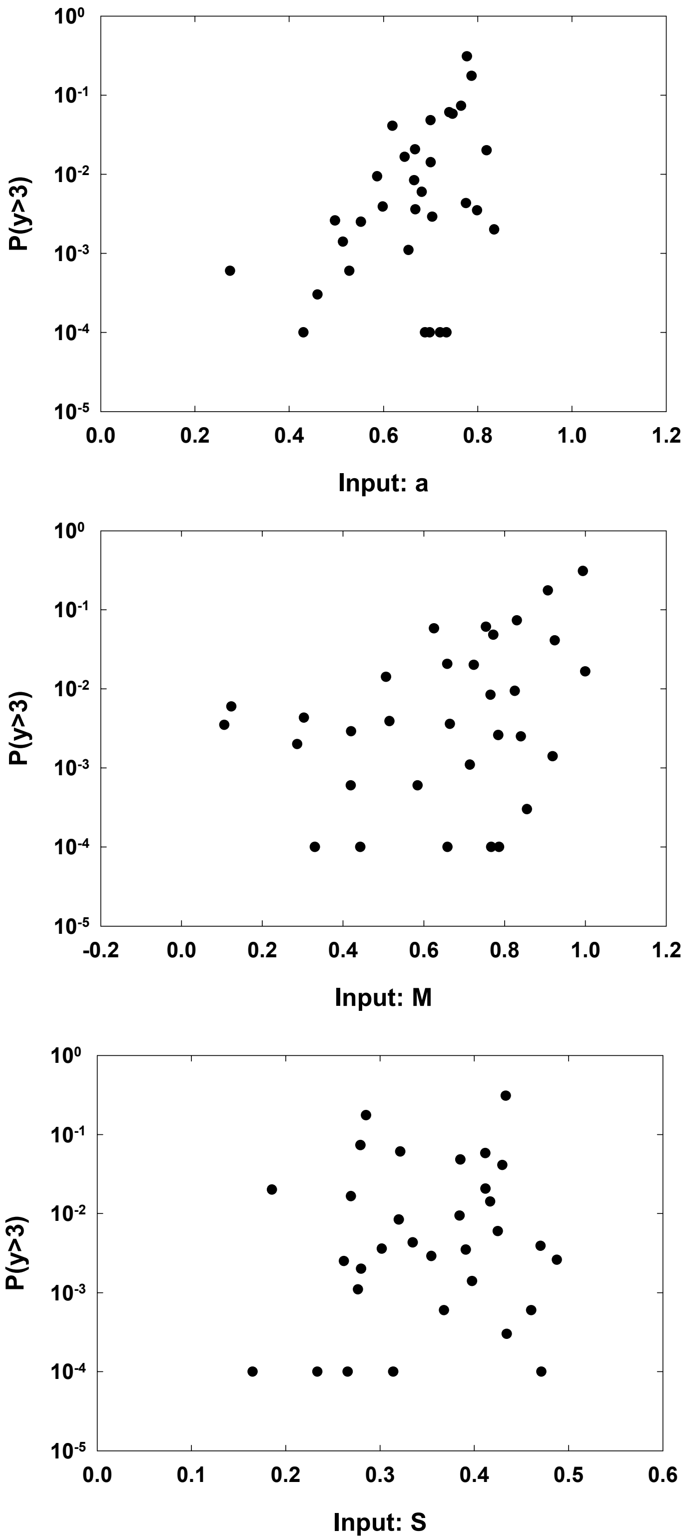


Figure 12. Sensitivity analysis for various modeling approaches.

Table 9. Analysis Results for Second-Order Analysis Using Subjective Probabilities to Represent Uncertainties

Second-order analyses are most appropriate for high-consequence issues where neither aleatory uncertainties nor epistemic uncertainties clearly dominate. The issue of adequate numerical convergence is greatly compounded with a probabilistic second-order analysis such as this. The computational burden here is orders of magnitude greater than a first-order probabilistic analysis (Section 5). Here, we represented the epistemic uncertainties with 100 realizations, each of which involved 10,000 function evaluations to represent the aleatory uncertainties, resulting in a total of 1 million function evaluations! One hundred realizations of epistemic uncertainty might be acceptable for estimating the mean of the confidence distribution, but it likely does not provide an adequate estimate of the 95th percentile. Clearly, second-order probabilistic analyses are not generally practical in a brute-force sense for our anticipated applications. As with the first-order Bayesian analysis, surrogate models may offer a practical path forward, recognizing that such approaches introduce additional "representational" errors that must be quantified.

The author is aware of important weapons applications in which a second-order analysis only requires large code calculations for the outer epistemic loop; the inner aleatory loop can be accomplished with information supplied solely from the outer loop and the aleatory characterization of component failure criteria. In these special cases, the computational burden (in terms of expensive function evaluations) of a second-order analysis becomes comparable to that of a first-order analysis.

Another challenge to the computational budget exists if there is some need to change one or more of the input distributions, e.g., if an error is discovered in the input distributions or new information becomes available. This is also a problem with the Bayesian approach, and it is particularly disheartening here because of the compounded computational burden. In this case, the ensemble of calculations must be repeated to get a credible and consistent result.

In summary, second-order methods greatly improve the clarity and utility of high-consequence risk analyses. Except for certain limiting cases, this improvement comes at the expense of greatly increased computational burden. Practical approaches to address the computational expense exist, but they introduce additional representation errors that must be addressed in the decision context along with physical variabilities and uncertainties. The representation and propagation of epistemic uncertainties with probabilistic methods is subject to the same concerns expressed for Bayesian methods, that is, such treatments are not information preserving.

7 Second-Order Method Using the MBS to Represent Epistemic Uncertainties

Alternate approaches have been proposed for characterizing and propagating epistemic uncertainty. These include interval analysis, possibility functions, belief/plausibility functions (evidence theory), and equivalence-class random variables, which all use some form of nonprobabilistic propagation. The MBS has roots in interval analysis and also uses nonprobabilistic propagation. There are some subtle differences that will be noted when possible, but there has been no attempt here to establish the formal relationship (if any) of the MBS with any of these other approaches.

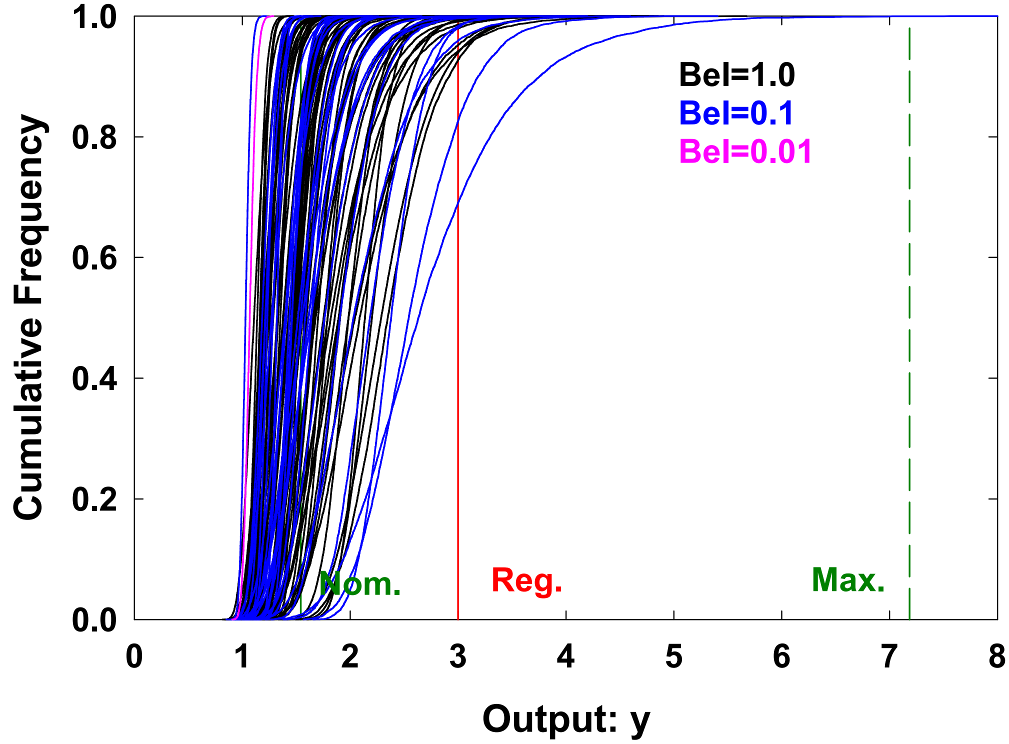
The approach taken here, as described in Table 10, is similar to that used in Section 6; that is, we generate an ensemble of frequency distributions, with each frequency distribution corresponding to a different realization of uncertainty. In Section 6, greater belief was associated with regions in which a greater density of frequency distributions existed. In this section, the attribute of belief is conveyed by associating a belief number (in the sense of the MBS) with each of the frequency distributions. It is important to reiterate that the MBS makes no assumption about the statistical independence of input distributions because all combinations of defensible inputs are explored.

Table 10. Analysis Process Associated with Second-Order Analysis Using the MBS to Represent Uncertainties

Case	Analysis Process
Second-order analysis using the MBS to represent uncertainties	1. Sample *a*, *M*, and *S* 100 times using Latin Hypercube Sampling (LHS) to create a set of input sets reflecting uncertainty. a. Evaluate the joint belief of the uncertain input set. b. Sample ln(*b*) 1,000 times (conditional on one realization of *M* and *S*) to create an input set representing variability. c. Evaluate *y* for each input set representing variability (conditional on one realization of *a*, *M*, and *S*). d. Construct a cumulative probability distribution for *y* (conditional on one realization of *a*, *M*, and *S*). Assign the joint belief of uncertain inputs to each cumulative distribution. e. Evaluate the probability that $y > 3$ from the cumulative distribution (conditional on one realization of *a*, *M*, and *S*). Assign the joint belief of uncertain inputs to each $P(y > 3)$. 2. Repeat steps 1a to 1e for each realization of uncertainty. 3. Construct a cumulative probability distribution for $P(y > 3)$.

We can invoke one of the unique features of the MBS to illustrate the method for the problem at hand. Without generating any new function evaluations we can work directly with the existing function evaluations previously generated in Section 6. In Section 6, each frequency distribution was associated with a triplet (a_i, M_i, S_i) representing one realization of uncertainty in *a*, *M*, and *S*. Based on our assessment of a belief distribution for each parameter (see Figure 3), we can associate a belief number to each individual parameter in the triplet and a joint belief to the triplet as a whole. This is the belief number associated with the frequency distribution derived from the triplet.

The results of the analysis are shown in Figure 13. The upper part of the figure displays an ensemble of frequency distributions that is exactly the same as that shown in Figure 11 of Section 6; only here, each frequency distribution is color coded according to its belief number. The lower part of the figure displays a "distribution (i.e. scatterplot)" that was created by calculating $P(y > 3)$, the probability of exceeding the regulatory requirement, for each frequency distribution and associating each probability with a belief number. Note that the probability calculated for each frequency distribution inherits the belief number associated with the particular frequency distribution.



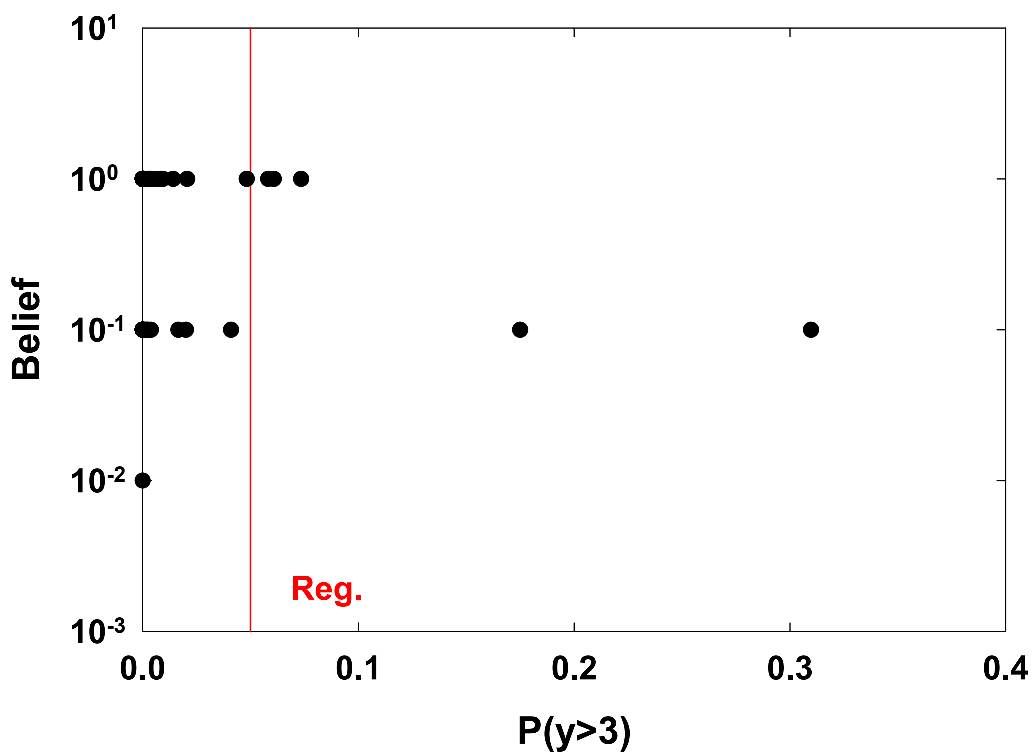


Figure 13. Second-order analysis using the MBS to represent uncertainties.

Note that the belief numbers in the scatterplot, or lower part, of Figure 13 appear in three tiers. The first tier (labeled 10°) corresponds to Bel = 1.0. This tier defines a range of values for P(y > 3) that are derived from only physically reasonable inputs. The second tier (labeled 10^{-1}) corresponds to Bel = 0.1. The second tier defines a range of values for P(y > 3) that is derived from inputs where one is judged physically unreasonable and all the others are judged physically reasonable. Likewise, the third tier (labeled 10^{-2}) corresponds to the single case where all the input is judged physically unreasonable, and thus P(y > 3) is judged physically incredible. Note that the ranges overlap and that they are not all nested. This means that some values of P(y > 3) can be achieved with inputs that have a joint belief of 1.0 or a joint belief of 0.1 or a joint belief of 0.01. Intuitively the highest belief values prevail in judging regulatory compliance. Note also that there is a range of values for P(y >3) that can only be achieved with some degree of diminished belief.

For the current problem, there are some physically reasonable and some physically unreasonable values of $P(y > 3)$ that exceed the regulatory requirement of 0.05; consequently, we <u>cannot</u> state that the regulatory requirement is met in any sense. This conclusion contrasts with the conclusion reached by using either the Bayesian approach or the subjective separatist approach. The cause of these differing conclusions is rooted in the fact that probabilistic characterization and propagation of epistemic inputs is not information preserving, i.e., implying more information than the evidence can justify.

Exploration of sensitivity through scatterplots, regression, and correlation is independent of the method used to represent and propagate uncertainties for the regulatory context. Here, the same function evaluations used in the previous second-order analysis are used; therefore, the representation and interpretation of sensitivity information is also the same. The results are repeated here as Table 11.

Table 11. Analysis Results for Second-Order Analysis Using the MBS to Represent Uncertainties

Case	Regulatory Decision	Sensitivity	# FEVs
Second-order analysis using the MBS to represent uncertainties	High confidence, $f_{max}(y > 3, \text{Bel} = 0.1) = 0.3098$ • Requirement not met Expectation, $f_{max}(y > 3, \text{Bel} = 1) = 0.0735$ • Requirement not met	Sensitivity of $\log_{10}[P(y > 3)]$ to a, M, and S $R^2 = 0.295$ $\beta_a^2 = 0.175$ $\beta_M^2 = 0.101$ $\beta_S^2 = 0.019$	1 million

The computational burden of a second-order probabilistic/MBS analysis can often be prohibitive unless a suitable surrogate model can be developed or for special situations where the aleatory loop does not require function evaluations as noted in the previous section. Here, we also represented the epistemic uncertainties with 10^2 realizations, but unlike the previous case, these 100 realizations are decisively adequate for determining that the regulatory requirement was not met.

With the MBS, all 100 realizations of epistemic uncertainty need not have been calculated if the analyst rank-ordered the 100 realizations of uncertain inputs according to his judgment of greatest challenge to the regulatory requirement. One calculation (or a small number of calculations) could have revealed that the regulatory requirement was not met with sufficient confidence. The remaining calculations would not be required unless sensitivity information was required; and even then, a reduced set might have been adequate.

The MBS has another attribute that could lead to significant computational savings. With the MBS, every calculation has value *in itself*, and belief numbers can be assigned or updated after the calculations are performed (just as we did in this section). Consequently, no new function evaluations may be necessary if adjustment in the belief distributions becomes necessary (because some error is discovered or new information becomes available) after the initial calculations are performed! The MBS has potential for additional computational savings, but this discussion will be deferred to the next section.

Note that the ensemble of curves shown previously in Figure 13 may not include the extreme distributions that envelop the regions of Bel = 1.0 and Bel ≥ 0.1 because it is highly unlikely that the sampling scheme would select input sets leading to the extreme states. Enveloping curves can be estimated with insight derived from the sensitivity analysis, which suggests that $P(y > 3)$ is positive monotonic to input a and input M and insensitive to input S. Based on these insights, Table 12 shows input for four additional calculations that lead to approximate enveloping distributions, which are shown in Figure 14. The expected value of each enveloping curve is listed in Table 13 and compared to bounds on expected values obtained by other participants of the Epistemic Uncertainty Workshop who also worked challenge problem 5a (Ferson et al. 2004). In particular, the bounds of expected values corresponding to Bel ≥ 0.1 compare almost identically to the range of expected values obtained by Ferson and Hajagos (2004).

Table 12. Inputs Leading to Approximate Enveloping Curves with Bel = 1.0 and Bel = 0.1

Input *a*	Bel *a*	Input *M*	Bel *M*	Input *S*	Bel *S*	Joint Bel
1	.1	.9	1	.45	1	.1
.8	1	.9	1	.45	1	1
.3	1	.1	1	.15	1	1
.1	.1	.1	1	.15	1	.1

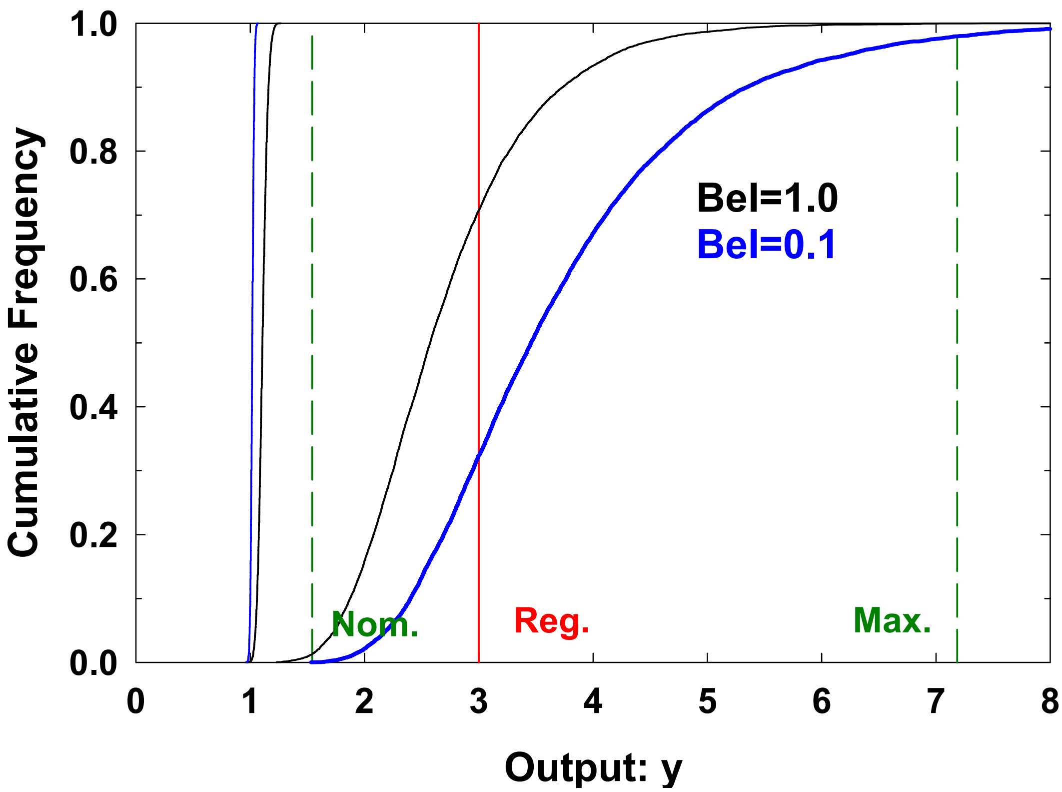


Figure 14. Approximate enveloping distributions.

Table 13. Comparison of Bounds on Expected Values

MBS	Kozine and Utkin 2004	DeCooman and Troffaes 2004	Ferson and Hajagos 2004
Bel = 1.0 {1.11, 2.72} Bel ≥ 0.1 {1.02, 3.72}	{1.45, 2.82}	{1.54, 2.19}	{1.05, 3.79}

8 The MBS and Distribution Relaxation

Second-order methods using the MBS provide a general treatment of variabilities and uncertainties in any risk analysis. Sometimes, there is broad consensus[5] that uncertainties dominate all variabilities; consequently, the second-order analysis can be reduced to a first-order analysis where only epistemic uncertainties remain. This greatly reduces the computational burden. It will become evident that this limiting case is a subset of a broader treatment provided by this section.

The computational burden of a second-order analysis is a significant drawback. Previously, we discussed the possible reduction of computational burden through the use of surrogate models. Here, we explore an alternate approach, which we call *distributional relaxation*. The goal is to reduce the second-order analysis to a first-order analysis while still preserving the enveloping nature of the MBS. The interval-like nature of the MBS also affords the opportunity for additional computational savings.

Whereas a Bayesian represents epistemic uncertainties with (subjective) probability distributions, here we propose representing aleatory uncertainties with belief distributions in the sense of the MBS. The latter is referred to as distribution relaxation because less information is implied by the resulting belief distribution than the evidence would support. Unlike the Bayesian approach, this leads to a more conservative result.

Aleatory uncertainties exist in the challenge problem because ln(b) was represented as a normal distribution, $N(M, S)$, while epistemic uncertainties existed (in part) because there was uncertainty in the distribution parameters (M, S). To develop the concept of distributional relaxation, consider first a simpler (purely aleatory) case where M and S are known precisely so that the distribution for ln(b) is completely prescribed as a frequency distribution. As a practical matter, we would not be interested in values of ln(b) that are below the 5th percentile or above the 95th percentile. Even if these extreme inputs always lead to outputs exceeding the regulatory threshold of 3, they would do so with a cumulative frequency that is less than the requirement of 5%. Hence, the normal distribution could be represented by an interval with limits equal to the 5th and 95th percentiles of the normal distribution. Although values outside the limits of the interval might be physically realizable, they are of no consequence in assessing regulatory compliance.

We can now return to the challenge problem where the normal distribution exhibits epistemic uncertainty in its statistics; consequently, we must examine an ensemble of distributions for ln(b). Four points (with a corresponding belief value) can characterize the three intervals for the belief distribution for M; likewise, four points can characterize the belief distribution for S. There are 16 combinations that must be explored, 4 of which are *physically incredible* (Bel ≤ 0.01). The remaining 12 cumulative curves are plotted in Figure 15 (curves are plotted between the 5th and 95th percentiles) with color coding to indicate which curves are *physically reasonable* (black) and which curves are *physically*

[5] Subjective judgments cannot be wholly eliminated from any risk analysis.

unreasonable (blue). The belief distribution is constructed as indicated by the dotted lines connecting the upper and lower figures. For example, the lower bound of the physically reasonable interval corresponds to the minimum ln(b) for the black curves (at the 5% level) and the maximum ln(b) of the black curves (at the 95% level). It is interesting to note that construction of the belief distribution is dominated by uncertainty in the mean M, which is consistent with the sensitivity analysis conducted for the second-order methods (see Figure 12 and Table 9).

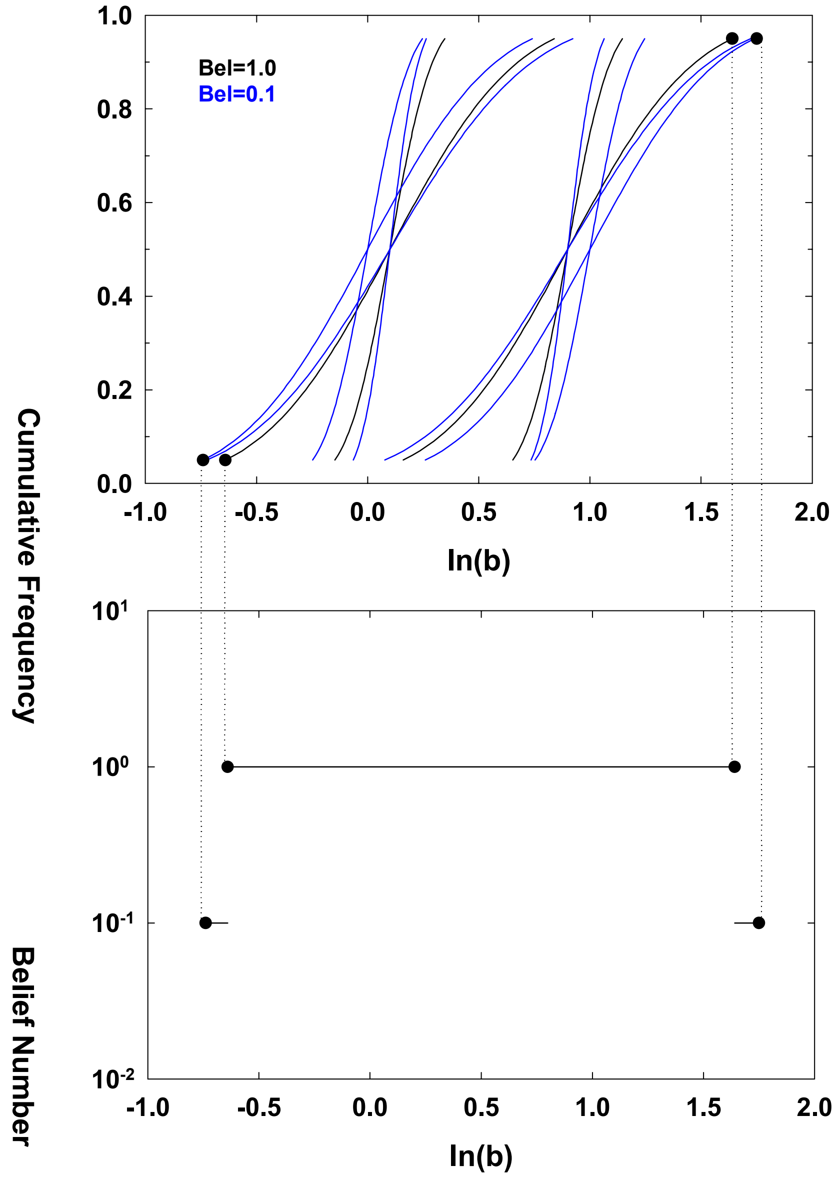


Figure 15. Distributional relaxation. Blue lines correspond to Bel = 0.1, and black lines correspond to Bel = 1.

Figure 16 and Table 15 shows the results of the distributional-relaxation analysis conducted according to the process outlined in Table 14. Here, we have employed a factorial design to illustrate computational designs different from sampling-based schemes.[6] The analyses suggest that output values greater than the regulatory threshold are *physically reasonable*; therefore, we cannot conclude that the regulatory requirement is met. Very large values of *y*, although *physically unreasonable*, cannot be positively excluded. These physically unreasonable values are not as large as those derived for the deterministic bounding analysis; and with the MBS, we understand explicitly that these large values are only possible with diminished belief. Even larger values are possible, but only under combinations of inputs deemed *physically incredible*.

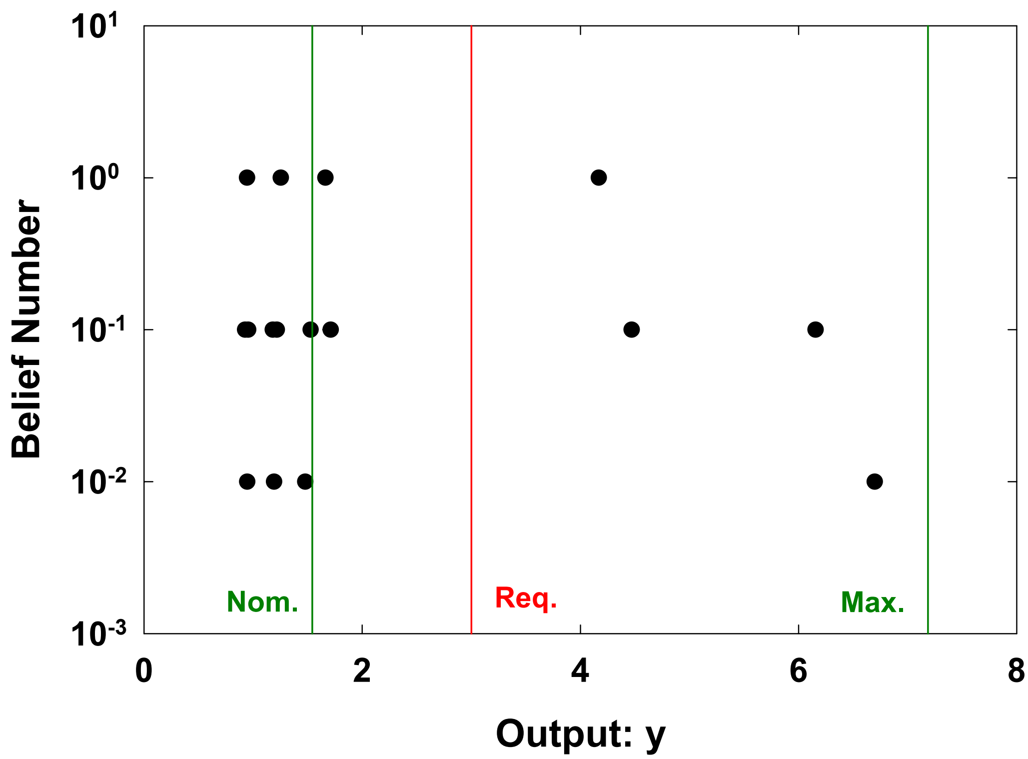


Figure 16. Analysis results for the MBS with distribution relaxation.

[6] When using probabilistic methods, meaning can only be drawn from the characteristics, e.g., mean and variance, of the ensemble of results. In the MBS, every single result has full meaning in itself. Structured analysis matrices, such as factorial designs, and unstructured analysis matrices, derived from sampling-based schemes, are useful ways of exploring the response space; but these methods, when employed in the context of the MBS, do not imply the same statistical attributes as they do when applied in the context of more traditional probabilistic methods.

Table 14. Analysis Process Associated MBS with Distribution Relaxation

Case	Analysis Process
MBS	1. Relax the probability distributions into belief distributions. 2. Select a and $\ln(b)$ input combinations using a four-level factorial design (16 input sets). 3. Evaluate the joint belief for each set of uncertain inputs. Evaluate y for each input set. Assign the joint belief of uncertain inputs to each y. 4. Construct the belief distribution from the set of 16 output ys. 5. Evaluate the belief that $y > 3$.

The scatterplots shown in Figure 17 address sensitivity of the output y to uncertainties in inputs a and $\ln(b)$ and key results are returned in Table 15. Figure 17 suggests that $\log_{10}(y)$ is a monotonic function of each of the uncertain inputs; however, the small R^2 value suggests that the relationship is not linear. Within the limitations of linearity, the β^2 statistics suggest that uncertainties in input a dominate uncertainties in $\ln(b)$. The global perspective of a scatterplot cannot detect the local non-monotinicity to input a that is known to exist (see section 3).

Table 15. Comparative Summary of Analysis Results

Case	Regulatory Decision	Sensitivity	# FEVs
MBS	$\mathrm{Bel}(y > 3) = 1$ • High confidence (i.e., Bel ≤ 0.01), requirement not met • Expectation (Bel ≤ 0.1), requirement not met	Sensitivity of $\log_{10}(y)$ to a, $\ln(b)$ $R^2 = 0.428$ $\beta_a^2 = 0.426$ $\beta_{\ln(b)}^2 = 0.002$	12

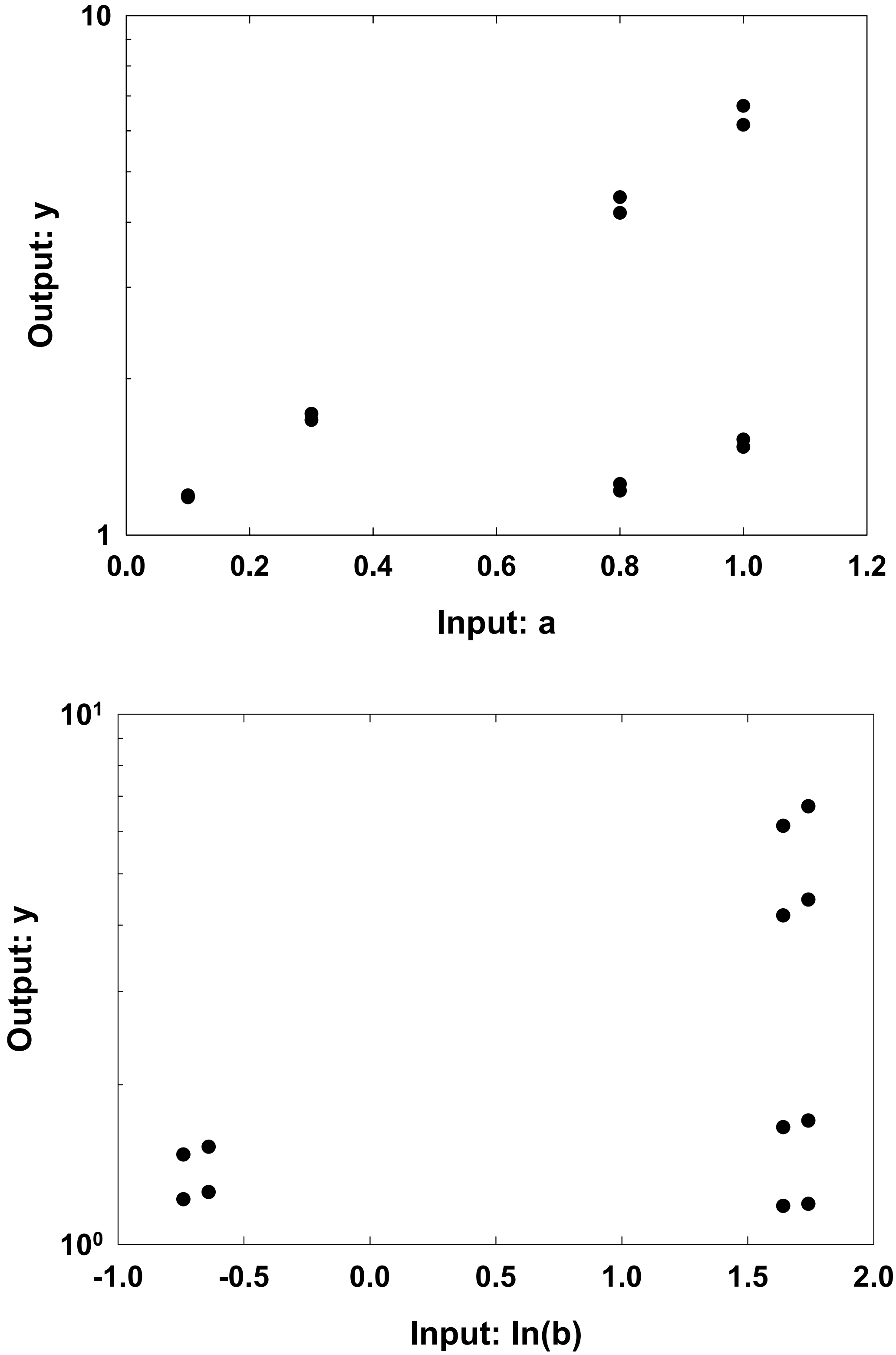


Figure 17. Sensitivity analysis for the MBS with distribution relaxation.

Table 16 summarizes the individual calculations. One advantage of the MBS is that the joint belief of various input combinations is known before the function evaluation is required, thus providing a means to screen out input sets of diminishing value. Consequently, the four input sets with a joint belief of 0.01 need not have been evaluated, and this is true regardless of the strategy employed to generate the input sets. For instance, we might have generated inputs sets through sampling (as opposed to the factorial design) of the input distributions. The joint belief of each input set could then be calculated, and those with Bel ≤ 0.01 could have been eliminated from the pool requiring function evaluation. This ability to screen out calculations of limited value is a unique advantage of the MBS.

Table 16. Problem 5a with Distribution Relaxation

a	Bel(*a*)	ln(*b*)	*b*	Bel(*b*)	*y*	Bel(*y*)
0.1	0.1	–0.7402	0.4770	0.1	0.9465	0.01
0.1	0.1	–0.6402	0.5272	1	0.9544	0.1
0.1	0.1	1.6402	5.1562	1	1.1805	0.1
0.1	0.1	1.7402	5.6985	0.1	1.1922	0.01
0.3	1	–0.7402	0.4770	0.1	0.9271	0.1
0.3	1	–0.6402	0.5272	1	0.9447	1
0.3	1	1.6402	5.1562	1	1.6637	1
0.3	1	1.7402	5.6985	0.1	1.7116	0.1
0.8	1	–0.7402	0.4770	0.1	1.2161	0.1
0.8	1	–0.6402	0.5272	1	1.2541	1
0.8	1	1.6402	5.1562	1	4.1685	1
0.8	1	1.7402	5.6985	0.1	4.4694	0.1
1	0.1	–0.7402	0.4770	0.1	1.4770	0.01
1	0.1	–0.6402	0.5272	1	1.5272	0.1
1	0.1	1.6402	5.1562	1	6.1562	0.1
1	0.1	1.7402	5.6985	0.1	6.6985	0.01

Factorial designs have the advantage that the limits of credibility for the outputs can be computed exactly if the output is monotonic to all the uncertain inputs, and limits of credibility can be computed approximately if the output is monotonic to all *dominant* inputs. Assuming monotonicity, and if all uncertain inputs are represented by an interval (Bel = 1), and it is not known whether the output is positive or negative monotonic, then 2^N function evaluations are required to exactly determine the limits of credibility for the output interval. Six uncertain input intervals can be addressed for a function budget of 64 evaluations.

If, on the other hand, all inputs are represented by two tiers of belief (e.g., Figure 3), that can be characterized by four input levels that span the limits of credibility (two with Bel = 1.0 and two with Bel = 0.1), then 4^N function evaluations are required to compute all interactions. Only three uncertain variables can be addressed with 64 function evaluations. However, some of the interactions will have Bel ≤ 0.01, and the MBS allows these physically incredible cases to be identified beforehand and eliminated from the calculation matrix. These cases need not be computed; therefore, only 32 function evaluations are needed to address three uncertain inputs. In general, the four-level design noted above will generate $(N+1)2^N$ combinations with Bel $= 1$ or Bel $= 0.1$; consequently, four uncertain inputs can be completely explored with 80 function evaluations (as opposed to 256 function evaluations for the full factorial design). The four-level factorial designs (truncated to cases with Bel $= 1$ and Bel $= 0.1$) can also approximate limits of credibility when the output function in nonmonotonic to uncertain inputs. This is because two points are evaluated at values intermediate to the extremes.

Monte Carlo is an alternate and useful means for generating input sets (as was done for the case using subjective probabilities in a second-order analysis). Latin Hypercube Sampling (LHS) is a variance-reducing class of Monte Carlo methods (Helton and Davis, 2000) that has the advantage of better representing the input ranges when the number of samples is relatively small (as would be the case in many of the anticipate weapons applications). LHS also has the advantage of more densely representing input ranges. For example, if we generate 50 LHS samples, then each uncertain parameter will have 50 values spanning its input range (as opposed to the 4 values for the factorial design). Thus, if the output exhibits a resonance behavior to one of the inputs, then this strategy has a better opportunity to capture the effect. On the other hand, if the extremes of the output are dominated by combinations of the extremes of the inputs, then the sparse sampling of interactions afforded by LHS may not result in an adequate estimate of the extreme outputs.

The MBS lends itself to a hybrid strategy. Initially, 50 function evaluations could be used to scope out the response and hopefully conclude that only a few of many inputs dominate uncertainty. These 50 function evaluations may be sufficient to make the regulatory decision (usually you can show that the regulatory requirement is <u>not</u> met). A strategy as outlined in Table 17 can be employed if more refinement is required to better resolve limits of output credibility when compliance to requirements is in doubt. The strategy, which leverages insights obtained from the initial 50 function evaluations, involves $2(N+1)$ additional function evaluations, where N is the number of uncertain inputs. As a practical matter, however, N could be interpreted as the number of *dominant* inputs because all others could be set to nominal values without significantly altering the output. Thus, only eight additional function evaluations are needed if 3 out of an initial 35 inputs are found to dominate the output response.

When the above strategy is successful, the MBS allows more information to be extracted from the limited full-physics function evaluations *without having to rely on surrogate models* and the additional representational errors they introduce to the analysis. When the above strategy is unsuccessful, it will be for those cases where dominant contributors to uncertainty cannot be identified, but these are the situations where it is likely to be most difficult to identify a suitable surrogate model.

Table 17. Strategy to Efficiently Estimate Limits of Credibility for Output *Y*

To Better Define:	Strategy	No. FEVs
Minimum output y with Bel = 0.1	• Pick the first x_i and select the value with Bel = 0.1 that minimizes y. • For all other x_is, pick the value with Bel = 1.0 that minimizes y. • Evaluate y for this set. • Repeat for all x_is.	N
Minimum output y with Bel = 1.0	• Pick the value with Bel = 1.0 for each x_i that minimizes y. • Evaluate y for this set.	1
Maximum output y with Bel = 1.0	• Pick the value with Bel = 1.0 for each x_i that maximizes y. • Evaluate y for this set.	1
Maximum output y with Bel = 0.1	• Pick the first dominant x_i and select the value with Bel = 0.1 that maximizes y. • For all other x_is, pick the value with Bel =1.0 that maximizes y. • Evaluate y_a for this set. • Repeat for all dominant x_is.	N

9 Regulatory Language and the Quantification of Margins

The basic language of regulation can be expressed as

Regulatory Metric (Assessed) < Regulatory Metric (Required).

Table 18 shows that the regulatory metric is expressed in terms of response, reliability, reliability at confidence, or confidence depending on the level of sophistication and generality brought to the analysis. Logically, the analysis methodology is dictated by the regulatory language; but in practice, the analysis community often works directly with the regulatory community to tailor the regulatory language to be consistent with the proposed methodology. This was certainly the case with WIPP for the geologic disposal of transuranic radioactive wastes, and with NUREG-1150 for reactor safety. Ambiguity (or conflict) arises when new methods become available. This is best illustrated through some examples.

Consider first a requirement on the safety of nuclear weapons in abnormal environments:[7]

> The probability of a premature nuclear detonation of a warhead due to warhead component malfunctions, … shall not exceed: … 1 in 10^6 per warhead exposure or accident. (Walske 1968)

We recognize this requirement as one of quantified reliability (QR): PIND[8] $< 10^{-6}$. This requirement was instituted in 1968. Modern risk analysts might ask if the Walske requirement should be interpreted as a frequency requirement (the expected value of many millions of such events is less than 10^{-6}), or as a requirement in belief (Bayesian perspective), or whether confidence intervals should be applied to such high-consequence events. Such confusion occurs because of the dual definition of the word “probability”: objective (frequency) and subjective (belief). Sufficient clarification (in the interest of “purity”) of the regulatory language, for example, frequensic interpretation, may preclude the possibility of demonstrating compliance through any other feasible analyses.

Conflict can also occur when the regulatory language is too specific. For example, we might have required for the challenge problem that $P(y > 3) < 0.05$ with 95% confidence. This language would fall in the class of “quantified reliability at confidence.” Such language is specific to a second-order approach using subjective probabilities to represent epistemic uncertainties. If an analyst wished to apply an alternate methodology, such as the MBS, then the regulatory language must be changed, which might involve interpreting “95% confidence” as “high confidence” (i.e., Belief ≥ 0.1).

---

[7] e.g., a nuclear weapon exposed to a large fuel fire associated with an airplane crash

[8] Probability of Inadvertent Nuclear Detonation (PIND)

Table 18. Regulatory Language and the Quantification of Margins

Regulatory Language		Quantification of Margins for Challenge Problem $\text{Margin} = \frac{\text{Regulated Value}}{\text{Assessed Value}}$
Regulatory Description	**Examples from Challenge Problem**	
Quantified Threshold (QT) Deterministic • Requirement *not met* with high confidence • Requirement *met* with best estimate	High Confidence $y_{max} = 7.1867$ Best Estimate $y_{nominal} = 1.5424$	High Confidence Margin = $3/y_{max} = 0.42$ Best Estimate Margin = $3/y_{nominal} = 1.95$
Quantified Reliability (QR) Frequensic or Bayesian • Requirement *met* with best estimate	$P(y > 3) = 0.0149$ Bayesian "belief" interpretation	Margin = $0.05/P(y > 3) = 3.36$
Quantified Reliability at Confidence (QRC) Second Order with Subjective Probabilities • Requirement *met* with best estimate	High Confidence $P_{95}(y > 3) = 0.048$, i.e., 95% confidence Best Estimate $P_{50}(y > 3) \sim 0$, i.e., 50% confidence	High Confidence Margin = $0.05/P_{95}(y > 3) = 1.04$ Best Estimate Margin = $0.05/P_{50}(y > 3) >> 1$
Quantified Reliability at Confidence (QRC) Second Order with Belief Scales • Requirement *not met* with high confidence • Requirement *not met* with best estimate	High Confidence $f_{max}(y > 3; \text{Bel} = 0.1) = 0.3098$, i.e., *physically unreasonable* Best Estimate $f_{max}(y > 3; \text{Bel} = 1) = 0.0735$, i.e., *physically reasonable*	High Confidence Margin = $0.05/f_{max}(y > 3; \text{Bel} = 0.1) = 0.16$ Best Estimate Margin = $0.05/f_{max}(y > 3; \text{Bel} = 1) = 0.68$
Quantified Confidence (QC) Belief Scales • Requirement *not met* with high confidence	High Confidence $y_{max}(\text{Bel} = 0.1) = 6.1562$, i.e., *physically unreasonable* Best Estimate $y_{max}(\text{Bel} = 1) = 4.1689$, i.e., *physically reasonable*	High Confidence Margin = $3/y_{max}(\text{Bel} = 0.1) = 0.49$ Best Estimate Margin = $3/y_{max}(\text{Bel} = 1) = 0.72$

Note that some risk-averse analysts are of the notion that high-consequence decisions should not be made unless we know everything about everything. The implication is that compliance with "conservative regulatory requirements" must be demonstrated with "high confidence." Such a perspective could lead to *analysis paralysis*, which is not constructive or even in the spirit of many high-consequence decision contexts. Take for instance the licensing of WIPP (Helton et al. 1996) for transuranic nuclear wastes. The relevant regulations (40 CFR 191.13) state the following: "Performance assessments need not provide complete assurance that the requirements of 191.13(a) will be met. . . what is required is a reasonable expectation, on the basis of the record before the implementing agency, that compliance with 191.13(a) will be achieved." Lastly, we note that the regulatory position taken by many agencies (Helton and Breeding 1996; Apostolakis and Guedes Soares 2004) is one of risk-informed regulation rather than risk-based regulation. This is an implicit acknowledgement that no body of evidence is perfect in every sense. A risk-informed approach acknowledges that there will always be programmatic and technical "intangibles" and that high-consequence decisions are ultimately assertion based; peer review thus becomes a critical element in validating the modeling approach and judging the adequacy of the evidence. Lastly, we note that both WIPP and reactor safety have employed second-order analyses using subjective probabilities to represent epistemic uncertainties. Compliance to requirements was judged based on the mean values (not the high-confidence values) of the resulting epistemic distributions.

Demonstrated compliance with a regulatory requirement, even if epistemic uncertainties are explicitly addressed as part of the requirement language, is often not enough to satiate the curiosity of decision makers because of the moral and legal liabilities that exist with high-consequence decisions. Decision makers often want to see some quantification of "margins," which is a measure of "ease" by which the requirement is met. Implicit in this curiosity is a realization that any risk analyses can only hope to quantify what is 'known" to some degree. Lurking in the shadows is the specter that some significant unknown-unknown could invalidate the conclusions of the study, and margin is a subjective hedge against such potential disasters.

There is no universally accepted definition of "margin." From a deterministic perspective, margin can be defined in terms of a difference, as in

$$\text{Margin} = \text{Required Value} - \text{Assessed Value},$$

or a ratio, as in

$$\text{Margin} = \frac{\text{Required Value}}{\text{Assessed Value}}.$$

Consequently, a positive value or a value exceeding unity denotes that the requirement is met, and increased comfort is derived for larger margins. There is a distinct advantage to defining margins in terms of a ratio because unity is a universal benchmark against which the magnitude of margin can be judged regardless of the nature of the regulatory metric and its associated units. In some disciplines, a safety factor applied to the ratio form of margin is prescribed by regulations as an experienced-based attempt to build in robustness against variabilities and uncertainties. Ad hoc requirements for margin-on-the-margin may be inadequate or overly conservative (and more generally indefensible) for other high-consequence or first-of-a-kind issues.

Another common approach is to define margin in the context of a confidence ratio (CR):

$$CR = \frac{\mu_{\text{Required Value}} - \mu_{\text{Assessed Value}}}{\sqrt{\sigma^2_{\text{Required Value}} + \sigma^2_{\text{Assessed Value}}}}$$

when variabilities dominate, and

$$CR = \frac{\text{Nominal Required Value} - \text{Nominal Assessed Value}}{\text{Uncertainty}}$$

when uncertainties dominate. The former definition is common in the reliability literature (Modarres 1992), and the latter has been proposed in the context of quantifying design margins and uncertainties (QMU) (Sharp and Wood-Schultz 2003; Logan and Nitta 2002) in the certification of physics package performance of nuclear weapons. Compliance to requirements is judged in terms of a specified value of the CR. In the reliability context, specifying a critical CR is equivalent to specifying an acceptable probability that the assessed value does not exceed the required value.[9] In the uncertainty dominated case, a CR > 1 is required to ensure compliance to requirements.

Although logically self-consistent, the CR approach has two drawbacks. First, the CR approach has not been applied to cases where aleatory and epistemic uncertainties are of comparable magnitudes. Second, the CR approach can be confusing because it is possible to have a positive margin and still not meet requirements. This is because, in this context, margin is defined in terms of the difference of mean or nominal values, and variabilities or uncertainties must be accounted for separately. Consequently, a new metric, is required to combine the concept of "best estimate" spread with the concept of variability or uncertainty to judge compliance to requirements.

We propose here a definition of margin that is always a direct measure of compliance to regulatory requirements,

[9] Strictly speaking, this is only true when the distributions are normal.

$$\text{Margin} = \frac{\text{Required Value of Regulatory Metric}}{\text{Assessed Value of Regulatory Metric}},$$

so that margin has values greater than unity when the regulations are met.[10] This definition can be written more explicitly for the specific cases of deterministic, variability dominated,[11] comparable variabilities and uncertainties, and uncertainty dominated respectively:

$$\text{Margin} = \frac{\text{Required Deterministic Value}}{\text{Assessed Deterministic Value}},$$

$$\text{Margin} = \frac{\text{Required Reliability}}{\text{Assessed Reliability}},$$

$$\text{Margin} = \frac{\text{Required Reliability at Specified Confidence}}{\text{Assessed Reliability at Assessed Confidence}},$$

$$\text{Margin} = \frac{\text{Required Value at Specified Confidence}}{\text{Assessed Value at Assessed Confidence}}.$$

Table 18, presented previously, evaluates these definitions in the context of specific analysis approaches for the challenge problem. The rank ordering of "high-confidence" margins (from the greatest to the smallest) is as follows:

1. Bayesian methods where variabilities and uncertainties are both represented by probabilities and propagated using probabilistic methods in a first-order analysis

2. Second-order methods where variabilities and uncertainties are represented by probabilities and both are propagated with probabilistic methods

3. Second-order methods where uncertainties are represented by the MBS.

4. The MBS with distributional relaxation in which variabilities are first represented by belief distributions and then propagated with all other uncertainties using the methods of the MBS

5. Deterministic bounding

[10] This formulation is appropriate to the common case where a regulated value should *not* be exceeded. There are cases where an assessed value should always exceed the required value. For these cases, the numerator and denominators should be swapped in this and the following expressions such that margin is always greater than unity when the requirement is met.
[11] This includes the Bayesian approach where both variabilities and uncertainties are represented by probabilities and both are propagated using first-order probabilistic methods.

The most appropriate answer is dictated by the nature of the available evidence. We assert that second-order methods using the MBS to address epistemic uncertainties are the most appropriate for the challenge problem. Bayesian methods and second-order methods with subjective probabilities arguably underrepresent the risk, while the MBS with distributional relaxation and the bounding deterministic analyses envelop the risk. MBS with distributional realization is preferable for high-consequence decisions, provided the results are not hyperconservative, and can be accommodated in a practical manner.

A favorable margin is adequate for making risk-informed decisions. Although unknown-unknowns are unquantifiable, there still remains an obligation to exercise "due diligence" in their identification and management. Some approaches to minimizing the risk of "surprise" from unknown-unknowns follow.

1. **Factor of safety:** This is akin to *requiring* a "margin on the margin." Even if a factor of safety is not required, it is useful to explore how far you have to push things before requirements are not met. Are we close to cliffs or points that would stimulate large undesirable swings in system response? Built-in redundancy is also a hedge against the unknown-unknown.

2. **Peer review:** Someone else may know what is unknown to me. Independent peer review by senior subject-matter experts from stakeholder organizations can identify unrecognized "issues."

3. **Organizational memory:** Some unknown-unknowns are should-have-been-knowns. Deliberate review of operational history and lessons-learned documents is also useful. This is an attempt to benefit from "organizational memory."

4. **Phenomena bifurcation:** A conscious employment of heuristics can be useful. For example, have you thought about instabilities or other phenomena that could cause physical phenomena to bifurcate, such as temperatures of key structures exceeding their melting point or the consequences of material interactions?

5. **Hierarchal testing:** The goal here is to allow Mother Nature an opportunity to reveal her unknown-unknowns through increasingly complicated testing. This can take the form of systematic hierarchical validation activities and system-level testing with real hardware under application-relevant conditions with emphasis on quality diagnostics at multiple locations.

6. **High-fidelity modeling and simulation:** The heuristics of even the best subject-matter experts are fallible in the face of nonlinear, coupled multiphysics in complex geometries. A systems-level approach, coupled with high-fidelity modeling (geometry and physics), can reveal the consequences of undesirable complex interactions.

7. **Verification:** Within the decision context, require explicit application-relevant evidence of code verification (are you solving the equations right?), solution verification (is the discretization adequate?), and input/output verification (are you processing information correctly?) whenever modeling and simulation is playing a critical role.

Clearly, strong safety and V&V cultures are critical if the risk of “surprise” is to be minimized.

10 Summary and Conclusions

Modeling and simulation is playing an increasing role in supporting weapons qualification and other tough regulatory decisions, which are typically characterized by *variabilities* and *uncertainties* in the scenarios, input conditions, failure criteria, model parameters, and even model form. *Variability* exists when there is a statistically significant database that is fully relevant to the application. *Uncertainty*, on the other hand, is characterized by some degree of ignorance. An algebraic problem was used to illustrate how various risk methodologies address variability and uncertainty in a regulatory context. These traditional risk methodologies include deterministic (bounding) methods, probabilistic methods (including frequensic and Bayesian perspectives), and second-order methods where variabilities and uncertainties are separated.

The deterministic approach is the most computationally efficient of all the methods addressed in this report. However, the deterministic approach, even if addressed from the perspective of bounding analyses, is the most programmatically risky. This approach is a risky proposition because of the high reliance on expert judgment to define worst-case scenarios and model inputs for nonlinear systems controlled by multiphysics in complex geometry, because compliance with requirements may be intractable (with practical resources) or unnecessarily costly and because no insight into sensitivity is obtained. Computational efficiency comes at the expense of an adequate understanding of contributors to uncertainty because an adequate sensitivity study cannot be formed with the limited computational budget.

Bayesian methods are well established in the technical community; however, the representation of variabilities epistemic uncertainties with (subjective) probability distributions, and the use of probabilistic methods to propagate subjective distributions, can lead to results that are not logically consistent with available knowledge. Bayesian methods are not information preserving (i.e., they can imply more than the available evidence can support), initial iterations may not be conservative, and it is difficult to prove convergence to the right answer. Probabilistic (forward) propagation requires a large number of function evaluations to ensure numerical convergence, particularly if very small probability levels must be resolved. There are options for performing analyses with a limited computational budget, but the numerical convergence error and any representational errors must be evaluated and may ultimately dominate the results. Any single function evaluation is meaningless; it is only in the context of an ensemble, i.e., a distribution, that meaning can be derived. Bayesian updating can be prohibitively expensive for the problems being addressed by the ASC program; and in many cases, the value of updating does not justify the added computational cost. With a Bayesian analysis, sensitivity analyses (critical to any risk analysis) confound the rank order of reducible (epistemic) and irreducible (aleatory) uncertainties.

State-of-the-art risk assessments in the United States separate the treatment of variabilities and uncertainties with a second-order analysis approach. Second-order methods are founded on an explicit conceptualization of an ensemble of frequency (output) distributions: one output (frequency) distribution for each realization of uncertain inputs. Variabilities on inputs contribute to the output frequency distribution, and uncertainties on inputs or model form are associated with the multiple realizations of (output) frequency distributions. Consequently, the impact of variabilities and uncertainties are represented separately and should be elements of any high-consequence decision process. Although treated separately, uncertainties are still commonly represented with (subjective) probability distributions and propagated with probabilistic methods. Second-order methods greatly improve the clarity and utility of high-consequence risk analyses. Except for certain limiting cases, this improvement comes at the expense of greatly increased computational burden. Practical approaches to address the computational expense exist, but such approaches introduce additional representation errors that must be addressed in the decision context along with other variabilities and uncertainties. The representation and propagation of epistemic uncertainties with probabilistic methods is subject to the same concerns expressed for Bayesian methods, i.e., such treatments are not information preserving. Traditional methods of sensitivity analysis can be employed, but the focus is solely on reducible (epistemic) uncertainties.

The MBS was developed as a means to logically aggregate uncertain input information and to propagate that information through the model to a set of results that are scrutable, easily interpretable by the nonexpert, and logically consistent (or conservative) with the available input information. The MBS can be used in a second-order analysis to represent and propagate the epistemic contributions to uncertainties. The MBS, particularly in conjunction with traditional sensitivity analyses, has the potential to be the most computationally efficient of all the risk methodologies because every function evaluation has full meaning in itself. Traditional methods of sensitivity analysis can be employed, but the focus is solely on reducible (epistemic) uncertainties.

Logically, the analysis methodology is dictated by the regulatory language; but in practice, the analysis community often works directly with the regulatory community to tailor the regulatory language to be consistent with the risk methodology. On the other hand, such focusing of the language to a specific risk methodology creates conflict for those who wish to pursue alternate methodologies. Margins, although not universally defined in a consistent manner, are defined here in a way that facilitates interpretation of the results, regardless of the risk methodology used. We generally define margin as

$$\text{Margin} = \frac{\text{Required Value of Regulatory Metric}}{\text{Assessed Value of Regulatory Metric}},$$

which is easily tailored to the MBS and all the methodologies discussed in this report.

References

Apostolakis, G.E., 2004. "How Useful is Quantitative Risk Assessment?", Risk Analysis, Vol 24, No. 3: 515-520.

Davison, A. C., and D. V. Hinkley. 1997. *Bootstrap Methods and Their Application*. Cambridge, England: Cambridge University Press.

De Cooman, G. , and M. C. M. Troffaes. 2004. "Coherent Lower Previsions in Systems Modelling: Products and Aggregation Rules." *Reliability Engineering & System Safety* 85, nos. 1–3 (July–September): 103–112.

Ferson, S., and J. G. Hajagos. 2004. "Arithmetic with Uncertain Numbers: Rigorous and (Often) Best Possible Answers." *Reliability Engineering & System Safety* 85, nos. 1–3 (July–September): 135–152.

Ferson, S., C. A. Joslyn, J. C. Helton, W. L. Oberkampf, and K. Sentz. 2004. "Summary from the Epistemic Uncertainty Workshop: Consensus Amid Diversity." *Reliability Engineering & System Safety* 85, nos. 1–3 (July–September): 355–369.

Helton, J. C., and R. J. Breeding. 1993. "Calculation of Reactor Accident Safety Goals," *Reliability Engineering and System Safety* 39, no. 2:129–158.

Helton, J. C., D. R. Anderson, G. Basabilvazo, H. –N. Jow, and M. G. Marietta. 2000. "Conceptual Structure of the 1996 Performance Assessment for the Waste Isolation Pilot Plant." *Reliability Engineering & System Safety* 69 (September): 151–165.

Helton, J. C., and F. J. Davis. 2000. *Sampling-Based Methods for Uncertainty and Sensitivity Analysis*. SAND99-2240. Albuquerque, NM: Sandia National Laboratories.

Helton, J. C., and F. J. Davis. 2001. *Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems*. SAND2001-0417. Albuquerque, NM: Sandia National Laboratories.

Helton, J.C., and F. J. Davis. 2003. "Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems," *Reliability Engineering & System Safety* 81. no. 1 (July): 23–69.

Helton, J. C., J. D. Johnson, and W. L. Oberkampf. 2004. "An Exploration of Alternative Approaches to the Representation of Uncertainty in Model Predictions," *Reliability Engineering & System Safety* 85, nos. 1–3 (July–September): 39–71.

Helton, J.C., and W. L. Oberkampf. 2004. "Guest Editorial." *Reliability Engineering & System Safety* 85, nos. 1–3 (July–September): 1–10.

Kozine, I. O., and L. V. Utkin. 2004, "An Approach to Combining Unreliable Pieces of Evidence and their Propagation in a System Response Analysis," *Reliability Engineering & System Safety*, 85, nos. 1–3 (July–September): 103–112.

Kusnezov, D. and J. Soudah, Advanced Simulation and computing Program Plan FY05, Sept, 2004, NA-ASC-102R-04-Vol. 1-Rev. D.

Lewis ET.Al., 1978. Rish Assessment Review Group Report to the US Nuclear Regulatory Commission, Nureg/CR-0400.

Logan, R.W. and C.K. Nitra, 2002. Verification & Validation (V&V) Methodology and Quantitative Reliability at confidence (QRC): Basis for an Investment Strategy, UCRL-ID-150874, November.

Modarres, M. 1992. *What Every Engineer Should Know about Reliability and Risk Analysis*. New York: Marcel Dekker.

Oberkampf, W. L., J. C. Helton, C. A. Joslyn, S. F. Wojtkiewicz, and S. Ferson. 2004. "Challenge Problems: Uncertainty in System Response Given Uncertain Parameters." *Reliability Engineering & System Safety* 85, nos. 1–3 (July–September): 11–19.

Red-Horse, J. R., and A. S. Benjamin. 2004. "A Probabilistic Approach to Uncertainty Quantification with Limited Information." *Reliability Engineering & System Safety 85*, nos. 1–3. (July–September): 183–190.

Saltelli, A., K. Chan, and E. M. Scott, eds. 2000. *Sensitivity Analysis. Probability and Statistics Series*. New York: John Wiley & Sons,.

Sharp, D. H., and M. M. Wood-Schultz. 2003. "QMU and Nuclear Weapons Certification." *Los Alamos Science*, no. 28.

Speijker, L. J. P., J. M. Van Noortwijk, M. Kok, and R.M. Cooke. 2000. "Optimal Maintenance Decisions for Dikes." *Probability in the Engineering and Informational Sciences* 14, no. 1 (January): 101–121.

Theofanous, T. G. 1996. "On the Proper Formulation of Safety Goals and Assessment of Safety Margins for Rare and High-Consequence Hazards." *Reliability Engineering and System Safety* 54, nos. 2–3 (November–December): 243–257.

US NRC, 1975. WASH-1400: Reactor Safety Study, NUREE-75/014

Walske, C. 1968. "Standards for Warhead and Bomb Premature Probability MC Paragraphs." Letter from Atomic Energy Commission Military Liaison Committee Chairman Walske to Bgen Edward B. Giller, March 14.

Distribution

Charles E. Anderson
Southwest Research Institute
P. O. Drawer 28510
San Antonio, TX 78284-0510

Bilal Ayyub (2)
Department of Civil Engineering
University of Maryland
College Park, MD 20742-3021

Ivo Babuska
TICAM
Mail Code C0200
University of Texas at Austin
Austin, TX 78712-1085

S. Balachandar
Theoretical and Applied Mechanics Dept.
216 Talbot Laboratory, MC 262
104 S. Wright St.
University of Illinois
Urbana, IL 61801

Osman Balci
Department of Computer Science
Virginia Tech
Blacksburg, VA 24061

J. Michael Barton
HQ Army Developmental Test Command
Technology Management Div., Rm 245
314 Longs Corner Road
Aberdeen Proving Ground, MD 21005-5005

Steven Batill (2)
Dept. of Aerospace & Mechanical Engr.
University of Notre Dame
Notre Dame, IN 46556

Ted Belytschko (2)
Department of Mechanical Engineering
Northwestern University
2145 Sheridan Road
Evanston, IL 60208

John Benek
AFRL/VAAC
2210 Eighth St.
Wright-Patterson AFB, OH 45433

James Berger
Inst. of Statistics and Decision Science
Duke University
Box 90251
Durham, NC 27708-0251

Jay Boris (2)
Laboratory for Computational Physics and Fluid Dynamics
Naval Research Laboratory
Code 6400
4555 Overlook Ave, SW
Washington, DC 20375-5344

Pavel A. Bouzinov
ADINA R&D, Inc.
71 Elton Avenue
Watertown, MA 02472

Mark Brandyberry
Computational Science and Engineering
2264 Digital Computer Lab, MC-278
1304 West Springfield Ave.
University of Illinois
Urbana, IL 61801

John A. Cafeo
General Motors R&D Center
Mail Code 480-106-256
30500 Mound Road
Box 9055
Warren, MI 48090-9055

Andrew Cary
The Boeing Company
MC S106-7126
P.O. Box 516
St. Louis, MO 63166-0516

James C. Cavendish
General Motors R&D Center
Mail Code 480-106-359
30500 Mound Road
Box 9055
Warren, MI 48090-9055

Chun-Hung Chen (2)
Department of Systems Engineering &
Operations Research
George Mason University
4400 University Drive, MS 4A6
Fairfax, VA 22030

Wei Chen (2)
Department of Mechanical Engineering
Northwestern University
2145 Sheridan Road, Tech B224
Evanston, IL 60208-3111

Kyeongjae Cho (2)
Dept. of Mechanical Engineering
MC 4040
Stanford University
Stanford, CA 94305-4040

Hugh Coleman
Department of Mechanical &
Aero. Engineering
University of Alabama/Huntsville
Huntsville, AL 35899

Raymond Cosner (2)
Boeing-Phantom Works
MC S106-7126
P. O. Box 516
St. Louis, MO 63166-0516

Thomas A. Cruse
AFRL Chief Technologist
1981 Monahan Way
Bldg. 12, Room 107
Wright-Patterson AFB, OH 45433-7132

Department of Energy (5)
Attn: Kevin Greenaugh, NA-115
D. Kusnezov, NA-114
Jamileh Soudah, NA-114
K. Sturgess, NA-115
J. Van Fleet, NA-113
Forrestal Building
1000 Independence Ave., SW
Washington, DC 20585

Prof. Urmila Diwekar (2)
University of Illinois at Chicago
Chemical Engineering Dept.
810 S. Clinton St.
209 CHB, M/C 110
Chicago, IL 60607

Isaac Elishakoff
Dept. of Mechanical Engineering
Florida Atlantic University
777 Glades Road
Boca Raton, FL 33431-0991

Ashley Emery
Dept. of Mechanical Engineering
Box 352600
University of Washingtion
Seattle, WA 98195-2600

Scott Ferson
Applied Biomathematics
100 North Country Road
Setauket, New York 11733-1345

Joseph E. Flaherty (2)
Dept. of Computer Science
Rensselaer Polytechnic Institute
Troy, NY 12181

John Fortna
ANSYS, Inc.
275 Technology Drive
Canonsburg, PA 15317

Marc Garbey
Dept. of Computer Science
Univ. of Houston
501 Philipp G. Hoffman Hall
Houston, Texas 77204-3010

Roger Ghanem
254C Kaprielian Hall
Dept. of Civil Engineering
3620 S. Vermont Ave.
University of Southern California
Los Angles, CA 90089-2531

Mike Giltrud
Defense Threat Reduction Agency
DTRA/CPWS
6801 Telegraph Road
Alexandria, VA 22310-3398

James Glimm (2)
Dept. of Applied Math & Statistics
P138A
State University of New York
Stony Brook, NY 11794-3600

James Gran
SRI International
Poulter Laboratory AH253
333 Ravenswood Avenue
Menlo Park, CA 94025

Ramana Grandhi (2)
Dept. of Mechanical and Materials
Engineering
3640 Colonel Glenn Hwy.
Dayton, OH 45435-0001

Bernard Grossman (2)
The National Institute of Aerospace
144 Research Drive
Hampton, VA 23666

Sami Habchi
CFD Research Corp.
Cummings Research Park
215 Wynn Drive
Huntsville, AL 35805

Raphael Haftka (2)
Dept. of Aerospace and Mechanical
Engineering and Engr. Science
P. O. Box 116250
University of Florida
Gainesville, FL 32611-6250

Achintya Haldar (2)
Dept. of Civil Engineering
& Engineering Mechanics
University of Arizona
Tucson, AZ 85721

Tim Hasselman
ACTA
2790 Skypark Dr., Suite 310
Torrance, CA 90505-5345

George Hazelrigg
Division of Design, Manufacturing
& Innovation
Room 508N
4201 Wilson Blvd.
Arlington, VA 22230

David Higdon
Inst. of Statistics and Decision Science
Duke University
Box 90251
Durham, NC 27708-0251

Richard Hills (2)
Mechanical Engineering Dept.
New Mexico State University
P. O. Box 30001/Dept. 3450
Las Cruces, NM 88003-8001

F. Owen Hoffman (2)
SENES
102 Donner Drive
Oak Ridge, TN 37830

Luc Huyse
Southwest Research Institute
6220 Culebra Road
P. O. Drawer 28510
San Antonio, TX 78284-0510

George Ivy
Northrop Grumman Information Technology
222 West Sixth St.
P.O. Box 471
San Pedro, CA 90733-0471

Rima Izem
Sience and Technology Policy Intern
Board of Mathematical Sciences and their Applications
500 5th Street, NW
Washington, DC 20001

Ralph Jones (2)
Sverdrup Tech. Inc./AEDC Group
1099 Avenue C
Arnold AFB, TN 37389-9013

Alan Karr
Inst. of Statistics and Decision Science
Duke University
Box 90251
Durham, NC 27708-0251

J. J. Keremes
Boeing Company
Rocketdyne Propulsion & Power
MS AC-15
P. O. Box 7922
6633 Canoga Avenue
Canoga Park, CA 91309-7922

K. D. Kimsey
U.S. Army Research Laboratory
Weapons & Materials Research Directorate
AMSRL-WM-TC 309 120A
Aberdeen Proving Gd, MD 21005-5066

B. A. Kovac
Boeing - Rocketdyne Propulsion & Power
MS AC-15
P. O. Box 7922
6633 Canoga Avenue
Canoga Park, CA 91309-7922

Chris Layne
AEDC
Mail Stop 6200
760 Fourth Street
Arnold AFB, TN 37389-6200

W. K. Liu (2)
Northwestern University
Dept. of Mechanical Engineering
2145 Sheridan Road
Evanston, IL 60108-3111

Robert Lust
General Motors, R&D and Planning
MC 480-106-256
30500 Mound Road
Warren, MI 48090-9055

Sankaran Mahadevan (2)
Dept. of Civil &
Environmental Engineering
Vanderbilt University
Box 6077, Station B
Nashville, TN 37235

Hans Mair
Institute for Defense Analysis
Operational Evaluation Division
4850 Mark Center Drive
Alexandria VA 22311-1882

W. McDonald
NDM Solutions
1420 Aldenham Lane
Reston, VA 20190-3901

Gregory McRae (2)
Dept. of Chemical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139

Michael Mendenhall (2)
Nielsen Engineering & Research, Inc.
605 Ellis St., Suite 200
Mountain View, CA 94043

Juan Meza
High Performance Computing Research
Lawrence Berkeley National Laboratory
One Cyclotron Road, MS: 50B-2239
Berkeley, CA 94720

Dr. John G. Michopoulos
Naval Research Laboratory,
Special Projects Group, Code 6303
Computational Mutliphysics Systems Lab
Washington DC 20375, USA

Sue Minkoff (2)
Dept. of Mathematics and Statistics
University of Maryland
1000 Hilltop Circle
Baltimore, MD 21250

Max Morris (2)
Department of Statistics
Iowa State University
304A Snedecor-Hall
Ames, IW 50011-1210

Rafi Muhanna
Regional Engineering Program
Georgia Tech
210 Technology Circle
Savannah, GA 31407-3039

R. Namburu
U.S. Army Research Laboratory
AMSRL-CI-H
Aberdeen Proving Gd, MD 21005-5067

NASA/Ames Research Center (2)
Attn: Unmeel Mehta, MS 229-3
David Thompson, MS 269-1
Moffett Field, CA 94035-1000

NASA/Glen Research Center (2)
Attn: John Slater, MS 86-7
Chris Steffen, MS 5-11
21000 Brookpark Road
Cleveland, OH 44135

NASA/Langley Research Center (7)
Attn: Dick DeLoach, MS 236
Michael Hemsch, MS 286
Tianshu Liu, MS 238
Jim Luckring, MS 286
Joe Morrison, MS 128
Ahmed Noor, MS 369
Sharon Padula, MS 159
Hampton, VA 23681-0001

C. Needham
Applied Research Associates, Inc.
4300 San Mateo Blvd., Suite A-220
Albuquerque, NM 87110

A. Needleman
Division of Engineering, Box D
Brown University
Providence, RI 02912

Robert Nelson
Dept. of Aerospace & Mechanical Engr.
University of Notre Dame
Notre Dame, IN 46556

Efstratios Nikolaidis (2)
MIME Dept.
4035 Nitschke Hall
University of Toledo
Toledo, OH 43606-3390

D. L. O'Connor
Boeing Company
Rocketdyne Propulsion & Power
MS AC-15
P. O. Box 7922
6633 Canoga Avenue
Canoga Park, CA 91309-7922

Tinsley Oden (2)
TICAM
Mail Code C0200
University of Texas at Austin
Austin, TX 78712-1085

Michael Ortiz (2)
Graduate Aeronautical Laboratories
California Institute of Technology
1200 E. California Blvd./MS 105-50
Pasadena, CA 91125

Dale K. Pace
4206 Southfield Rd
Ellicott City, MD 21042-5906

Alex Pang
Computer Science Department
University of California
Santa Cruz, CA 95064

Chris Paredis
School of Mechanical Engineering
Georgia Institute of Technology
813 Ferst Drive, MARC Rm. 256
Atlanta, GA 30332-0405

Chris L. Pettit
Aerospace Engineering Dept.
MS-11B
590 Holloway Rd.
Annapolis, MD 21402

Allan Pifko
2 George Court
Melville, NY 11747

Joseph Powers
Dept. of Aerospace and Mechanical
Engr.
University of Notre Dame
Notre Dame, IN 46556-5637

Cary Presser (2)
Process Measurements Div.
National Institute of Standards
and Technology
Bldg. 221, Room B312
Gaithersburg, MD 20899

Gerald R. Prichard
Principal Systems Analyst
Dynetics, Inc.
1000 Explorer Blvd.
Huntsville, AL 35806

Thomas A. Pucik
Pucik Consulting Services
13243 Warren Avenue
Los Angles, CA 90066-1750

W. Rafaniello
DOW Chemical Company
1776 Building
Midland, MI 48674

Chris Rahaim
Rotordynamics-Seal Research
3302 Swetzer Road
Loomis, CA 95650

J. N. Reddy
Dept. of Mechanical Engineering
Texas A&M University
ENPH Building, Room 210
College Station, TX 77843-3123

John Renaud (2)
Dept. of Aerospace & Mechanical Engr.
University of Notre Dame
Notre Dame, IN 46556

Grant Reinman
Pratt & Whitney
400 Main Street, M/S 162-01
East Hartford, CT 06108

Patrick J. Roache
1215 Apache Drive
Socorro, NM 87801

Tim Ross (2)
Dept. of Civil Engineering
University of New Mexico
Albuquerque, NM 87131

Chris Roy (2)
Dept. of Aerospace Engineering
211 Aerospace Engineering Bldg.
Auburn University, AL 36849-5338

J. Sacks
Inst. of Statistics and Decision Science
Duke University
Box 90251
Durham, NC 27708-0251

Sunil Saigal (2)
Carnegie Mellon University
Department of Civil and
Environmental Engineering
Pittsburgh, PA 15213

Larry Sanders
DTRA/ASC
8725 John J. Kingman Rd
MS 6201
Ft. Belvoir, VA 22060-6201

Len Schwer
Schwer Engineering & Consulting
6122 Aaron Court
Windsor, CA 95492

Paul Senseny
Factory Mutual Research Corporation
1151 Boston-Providence Turnpike
P.O. Box 9102
Norwood, MA 02062

E. Sevin
Logicon RDA, Inc.
1782 Kenton Circle
Lyndhurst, OH 44124

Mark Shephard (2)
Rensselaer Polytechnic Institute
Scientific Computation Research Center
Troy, NY 12180-3950

Tom I-P. Shih
Dept. of Mechanical Engineering
2452 Engineering Building
East Lansing, MI 48824-1226

T. P. Shivananda
Bldg. SB2/Rm. 1011
TRW/Ballistic Missiles Division
P. O. Box 1310
San Bernardino, CA 92402-1310

Don Simons
Northrop Grumman Information Tech.
222 W. Sixth St.
P.O. Box 471
San Pedro, CA 90733-0471

Munir M. Sindir
Boeing - Rocketdyne Propulsion & Power
MS GB-11
P. O. Box 7922
6633 Canoga Avenue
Canoga Park, CA 91309-7922

Ashok Singhal (2)
CFD Research Corp.
Cummings Research Park
215 Wynn Drive
Huntsville, AL 35805
R. Singleton
Engineering Sciences Directorate
Army Research Office
4300 S. Miami Blvd.
P.O. Box 1221
Research Triangle Park, NC 27709-2211

W. E. Snowden
DARPA
7120 Laketree Drive
Fairfax Station, VA 22039

Bill Spencer (2)
Dept. of Civil Engineering
and Geological Sciences
University of Notre Dame
Notre Dame, IN 46556-0767

G. R. Srinivasan
Org. L2-70, Bldg. 157
Lockheed Martin Space & Strategic Missiles
1111 Lockheed Martin Way
Sunnyvale, CA 94089

Fred Stern
Professor Mechanical Engineering
Iowa Institute of Hydraulic Research
The University of Iowa
Iowa City Iowa 52242

D. E. Stevenson (2)
Computer Science Department
Clemson University
442 Edwards Hall, Box 341906
Clemson, SC 29631-1906

Tim Swafford
Sverdrup Tech. Inc./AEDC Group
1099 Avenue C
Arnold AFB, TN 37389-9013

Ben Thacker
Southwest Research Institute
6220 Culebra Road
P. O. Drawer 28510
San Antonio, TX 78284-0510

Fulvio Tonon (2)
Geology and Geophysics Dept.
East Room 719
University of Utah
135 South 1460
Salt Lake City, UT 84112

Robert W. Walters (2)
Aerospace and Ocean Engineering
Virginia Tech
215 Randolph Hall, MS 203
Blacksburg, VA 24061-0203

Leonard Wesley
Intellex Inc.
5932 Killarney Circle
San Jose, CA 95138

Justin Y-T Wu
8540 Colonnade Center Drive, Ste 301
Raleigh, NC 27615

Ren-Jye Yang
Ford Research Laboratory
MD2115-SRL
P.O.Box 2053
Dearborn, MI 4812

Simone Youngblood (2)
DOD/DMSO
Technical Director for VV&A
1901 N. Beauregard St., Suite 504
Alexandria, VA 22311

M. A. Zikry
North Carolina State University
Mechanical & Aerospace Engineering
2412 Broughton Hall, Box 7910
Raleigh, NC 27695

Foreign Distribution

Yakov Ben-Haim (2)
Department of Mechanical Engineering
Technion-Israel Institute of Technology
Haifa 32000
ISRAEL

Luis Eca (2)
Instituto Superior Tecnico
Department of Mechanical Engineering
Av. Rovisco Pais
1096 Lisboa CODEX
PORTUGAL

Charles Hirsch (2)
Department of Fluid Mechanics
Vrije Universiteit Brussel
Pleinlaan, 2
B-1050 Brussels
BELGIUM

Igor Kozin (2)
Systems Analysis Department
Riso National Laboratory
P. O. Box 49
DK-4000 Roskilde
DENMARK

Nina Nikolova - Jeliazkova
Institute of Parallel Processing
Bulgarian Academy of Sciences
25a "acad. G. Bonchev" str.
Sofia 1113
Bulgaria

K. Papoulia
Inst. Eng. Seismology & Earthquake
Engineering
P.O. Box 53, Finikas GR-55105
Thessaloniki
GREECE

Dominique Pelletier
Genie Mecanique
Ecole Polytechnique de Montreal
C.P. 6079, Succursale Centre-ville
Montreal, H3C 3A7
CANADA

Vincent Sacksteder
Via Eurialo 28, Int. 13
00181 Rome
Italy

Malcolm Wallace
Computational Dynamics Ltd.
200 Shepherds Bush Road
London W6 7NY
UNITED KINGDOM

Department of Energy Laboratories

Los Alamos National Laboratory (49)
Mail Station 5000
P.O. Box 1663
Los Alamos, NM 87545
Attn: Peter Adams, MS B220
Mark C. Anderson, MS T080
Cuauhtemoc Aviles-Ramos, MS P946
Terrence Bott, MS K557
Jerry S. Brock, MS F663
D. Cagliostro, MS F645
C. Chiu, MS F600
David L. Crane, MS P946
John F. Davis, MS B295
Helen S. Deaven, MS B295
Barbara DeVolder, MS B259
Scott Doebling, MS T080
Sunil Donald, MS K557
S. Eisenhawer, MS K557
Dawn Flicker, MS F664
George T. Gray, MS G755
Alexandra Heath, MS F663
Francois Hemez, MS T006
R. Henninger, MS D413
Kathleen Holian, MS B295
Darryl Holm, MS B284
James Hyman, MS B284
Cliff Joslyn, MS B265
James Kamm, MS D413
S. Keller-McNulty, MS F600
Joseph Kindel, MS B259
Ken Koch, MS F652
Douglas Kothe, MS B250
Jeanette Lagrange, MS D445
Jonathan Lucero, MS C926
Len Margolin, MS D413
Kelly McLenithan, MS F664
Mark P. Miller, MS P946
John D. Morrison, MS F602
Karen I. Pao, MS B256
James Peery, MS F652
M. Peterson-Schnell, MS B295
Douglas Post, MS F661 X-DO
William Rider, MS D413

Tom Seed, MS F663
Kari Sentz, MS F600
David Sharp, MS B213
Richard N. Silver, MS D429
Ronald E. Smith, MS J576
Christine Treml, MS J570
David Tubbs, MS B220
Daniel Weeks, MS B295
Morgan White, MS F663
Alyson G. Wilson, MS F600

Lawrence Livermore National Laboratory (23)
7000 East Ave.
P.O. Box 808
Livermore, CA 94550
Attn: Thomas F. Adams, MS L-095
Steven Ashby, MS L-561
John Bolstad, MS L-023
Peter N. Brown, MS L-561
T. Scott Carman, MS L-031
R. Christensen, MS L-160
Evi Dube, MS L-095
Henry Hsieh, MS L-229
Richard Klein, MS L-023
Roger Logan, MS L-125
C. F. McMillan, MS L-098
C. Mailhiot, MS L-055
J. F. McEnerney, MS L-023
M. J. Murphy, MS L-282
Daniel Nikkel, MS L-342
Cynthia Nitta, MS L-096
Peter Raboin, MS L-125
Edward Russell, MS L-631
Kambiz Salari, MS L-228
David D. Sam, MS L-125
Joe Sefcik, MS L-125
Peter Terrill, MS L-125
Charles Tong, MS L-560
Carol Woodward, MS L-561

Argonne National Laboratory
Attn: Paul Hovland
Mike Minkoff
MCS Division
Bldg. 221, Rm. C-236
9700 S. Cass Ave.
Argonne, IL 60439

Sandia Internal Distribution

1	MS 0511	1021	K. D. Meeks
1	MS 1152	1642	M. L. Kiefer
1	MS 1186	1674	R. J. Lawrence
1	MS 0525	1734	P. V. Plunkett
1	MS 0525	1734	R. B. Heath
1	MS 0525	1734	S. D. Wix
1	MS 0429	2100	J. S. Rottler
1	MS 0437	2100	R. C. Hartwig
1	MS 0447	2111	P. Davis
1	MS 0447	2111	P. D. Hoover
1	MS 0479	2113	J. O. Harrison
1	MS 0487	2115	P. A. Sena
1	MS 0453	2130	H. J. Abeyta
1	MS 0481	2132	M. A. Rosenthal
1	MS 0427	2134	R. A. Paulsen
1	MS 0509	2300	M. W. Callahan
1	MS 1138	2345	G. K. Froehlich
1	MS 0512	2500	T. E. Blejwas
1	MS 0437	2830	J. M. McGlaun
1	MS 0647	2912	D. R. Olson
1	MS 0634	2951	K. V. Chavez
1	MS 0768	4140	A. L. Camp
1	MS 0757	4142	R. D. Waters
1	MS 0759	4145	G. D. Wyss
1	MS 1115	4200	D. S. Miyoshi
1	MS 0670	5526	J. R. Weatherby
1	MS 0735	6115	S. C. James
1	MS 0751	6117	R. M. Brannon
1	MS 0751	6117	L. S. Costin
1	MS 0751	6117	A. F. Fossum
1	MS 0708	6214	P. S. Veers
1	MS 1138	6221	G. E. Barr
1	MS 1138	6222	S. M. DeLand
1	MS 1138	6222	P. G. Kaplan
1	MS 1138	6223	L. M. Claussen
1	MS 1138	6223	A. L. Hodges
1	MS 1137	6224	G. D. Valdez
1	MS 0490	6252	J. A. Cooper
1	MS 1395	6820	M. J. Chavez
1	MS 1395	6821	J. W. Garner
1	MS 1395	6821	E. R. Giambalvo
1	MS 0778	6851	P. N. Swift
1	MS 0778	6851	B. W. Arnold
1	MS 0778	6851	R. J. MacKinnon
1	MS 0776	6852	T. Hadgu
1	MS 0776	6852	M. K. Knowles
1	MS 0776	6852	J. S. Stein
1	MS 1399	6853	M. G. Marietta
1	MS 0776	6853	R. P. Rechard
1	MS 1399	6853	P. Vaughn
1	MS 0748	6861	D. G. Robinson
1	MS 0727	6870	D. A. Powers
1	MS 1146	6871	P. J. Griffin
1	MS 1379	6957	A. Johnson
1	MS 9001	8000	P. E. Nielan
1	MS 9004	8100	R. H. Stulen
1	MS 9007	8200	D. R. Henson
1	MS 9202	8205	R. M. Zurn
1	MS 9014	8242	A. R. Ortega
1	MS 9042	8752	C. D. Moen
1	MS 9409	8754	J. R. Garcia
1	MS 9409	8754	W. A. Kawahara
1	MS 9161	8763	E. P. Chen
1	MS 9405	8763	R. E. Jones
1	MS 9161	8763	P. A. Klein
1	MS 9405	8763	R. A. Regueiro
1	MS 9042	8774	J. J. Dike
1	MS 9003	8900	K. E. Washington
1	MS 9003	8940	C. M. Hartwig
1	MS 9003	8960	J. L. Handrock
1	MS 9217	8962	P. D. Hough
1	MS 9217	8962	K. R. Long
1	MS 9217	8962	M. L. Martinez-Canales
1	MS 0384	9100	T. C. Bickel
1	MS 0824	1500	T.Y. Chu
1	MS 0825	1500	A.C. Ratzel
1	MS 0826	1500	D. K. Gartling
1	MS 0825	1510	W. Hermina
1	MS 0834	1512	J. E. Johannes

1	MS 0834	1512	S. J. Beresh
1	MS 1310	1513	S. N. Kempka
1	MS 0834	1514	K. S. Chen
1	MS 0834	1514	R. R. Rao
1	MS 0834	1514	P. R. Schunk
1	MS 0825	1515	B. Hassan
1	MS 0825	1515	M. Barone
1	MS 0825	1515	F. G. Blottner
1	MS 0825	1515	D. W. Kuntz
1	MS 0380	1542	G. M. Reese
1	MS 0382	1543	J. R. Stewart
1	MS 0382	1543	K. M. Aragon
1	MS 0382	1543	K. D. Copps
1	MS 0382	1543	H. C. Edwards
1	MS 0382	1543	G. D. Sjaardema
1	MS 0321	1400	W. J. Camp
1	MS 0318	1400	G. S. Davidson
1	MS 1110	1410	D. E. Womble
1	MS 0370	1411	S. A. Mitchell
1	MS 0370	1411	R. A. Bartlett
1	MS 0370	1411	S. S. Collis
1	MS 0370	1411	M. S. Eldred
1	MS 0370	1411	D. M. Gay
1	MS 0370	1411	L. P. Swiler
5	MS 0370	1411	T. G. Trucano
1	MS 0370	1411	B. G. van Bloemen Waanders
1	MS 0310	1412	M. D. Rintoul
1	MS 1110	1412	S. J. Plimpton
1	MS 1110	1414	J. DeLaurentis
1	MS 1110	1414	R. B. Lehoucq
1	MS 1110	1415	S. K. Rountree
1	MS 1111	1415	B. A. Hendrickson
1	MS 1110	1415	R. Carr
1	MS 1110	1415	W. E. Hart
1	MS 1111	1415	P. Knupp
1	MS 1110	1415	V. J. Leung
1	MS 1110	1415	C. A. Phillips
1	MS 0318	1416	J. E. Nelson
1	MS 1109	1416	R. J. Pryor
1	MS 0321	4300	R. W. Leland
1	MS 0370	1422	J. A. Ang
1	MS 1110	1423	N. D. Pundit
1	MS 0817	1422	D. W. Doerfler
1	MS 0376	1421	T. D. Blacker
1	MS 0822	1424	D. R. White
1	MS 0822	1424	C. F. Diegert
1	MS 0316	1902	P. Yarrington
1	MS 0378	1431	R. M. Summers
1	MS 0378	6861	S. P. Burns
1	MS 0378	1431	R. R. Drake
1	MS 0378	1431	A. C. Robinson
1	MS 0378	1433	T. Voth
1	MS 0378	1431	M. K. Wong
1	MS 0370	1431	M. E. Kipp
1	MS 0378	1431	S. A. Silling
1	MS 0378	1435	P. A. Taylor
1	MS 0316	1420	S. S. Dosanjh
1	MS 0316	1433	D. R. Gardner
1	MS 0316	1433	S. A. Hutchinson
1	MS 1111	1433	A. G. Salinger
1	MS 0316	1433	J. N. Shadid
1	MS 0310	1435	J. B. Aidun
1	MS 0310	1435	H. P. Hjalmarson
1	MS 0139	5500	M. O. Vahle
1	MS 0139	1904	R. K. Thomas
1	MS 0139	2614	S. E. Lott
1	MS 0428	00200	D. D. Carlson
1	MS 0428	12340	V. J. Johnson
1	MS 0638	12341	D. L. Knirk
1	MS 0638	12341	D. E. Peercy
1	MS 0829	12337	W. C. Moffatt
1	MS 0829	12337	J. M. Sjulin
1	MS 0829	12337	B. M. Rutherford
1	MS 0829	12337	F. W. Spencer
1	MS 0405	12300	T. R. Jones
1	MS 0405	12346	S. E. Camp
1	MS 0830	12335	K. V. Diegert
1	MS 1030	12870	J. G. Miller
1	MS 1221	05000	J. L. McDowell
1	MS 1002	06630	P. D. Heermann
1	MS 1170	06640	R. D. Skocypec
1	MS 1176	06642	D. J. Anderson
1	MS 1176	06642	J. E. Campbell
1	MS 1176	06643	R. M. Cranwell
1	MS 1153	05131	L. C. Sanchez
1	MS 1169	01300	J. R. Lee
1	MS 1179	01341	L. Lorence
1	MS 1162	05422	W. H. Rutledge
1	MS 9018	8945-1	Central Technical Files
2	MS 0899	9616	Technical Library