Uncertainty Quantification Whitepaper

"Uncertainty Quantification and the Department of Homeland Security"

Timothy G. Trucano
Optimization and Uncertainty Estimation Department
Sandia National Laboratories
Albuquerque, New Mexico 87185-0370

tgtruca@sandia.gov

Abstract:

"Uncertainty Quantification" (UQ) is the quantitative characterization and use of uncertainty in information applications. Fundamental to UQ is the recognition that there are two distinctly different types of uncertainty: "variability," which can be quantified in principle using classical probability theory; and "lack of knowledge," which requires more than classical probability theory for its quantification. We readily find both fundamental types of uncertainty in complex technical decision problems. While there are scientific challenges associated with UQ for certain applications, an important body of methods and results exists that is broadly useful. For example, the U.S. has established two important methodological precedents, one in nuclear reactor safety, and the other in operational licensing of the Waste Isolation Pilot Plant, for the deployment of UQ methods in high-consequence decision-making. These precedents, in fact, provide a foundation as well as encouragement for future use of UQ in both traditional and nontraditional decision applications.

Introduction

The need to make high-consequence decisions under severe constraints, including decision time and uncertainty in information, will be an important concern for the U. S. Department of Homeland Security (DHS). Decisions require four key elements: (1) identification of possible events or actions, called scenarios here; (2) assessment of the likelihood of these scenarios; (3) prediction of the consequences of these scenarios; and (4) understanding the level of confidence in the information gathered for the first three elements. In all of these elements, uncertainty in the required information influences the decision making process and complicates the required analyses. Ignoring uncertainty, perhaps using (believed) conservatism in the decision process, is inadequate and dangerous. Instead, fully embracing the influence and consequences of uncertainty in high-consequence DHS decisions is the required approach.

Uncertainty Quantification Whitepaper

What is Uncertainty?

Uncertainty pertains to information that is not definitely ascertainable or fixed, not precisely determined, not dependable or that is vague or indistinct. A decision maker dealing with uncertainty has reduced confidence and assurance in the information, hence in decisions that are dependent upon that information. To the degree that we can quantify uncertainty, it is increasingly feasible to make more reliable and assured decisions.

The community that specializes in ***uncertainty quantification*** (UQ) (Helton, 1994; Kaplan and Garrick, 1981) understands uncertainty should be divided into two categories for purposes of quantification. The first category, called ***variability*** (also called aleatory uncertainty), is uncertainty of the type associated with frequensic (also called objective) probabilistic processes and inference. Variability appears, or may appear, in many classes of information that DHS S&T must deal with. Variability may arise in the specification of inputs for various models (for example, likely wind conditions at a given location); in the conduct of the actual modeling (for example, the use of Monte Carlo methods within a model); and in the development of scenarios (for example, the use of polling data).

Variability as a description of uncertainty is constrained by the requirement that sufficient data exist to characterize the frequensic probabilistic interpretation underlying it. In other words, variability is a product of stochastic behavior, and the accurate characterization of this stochastic behavior is presumed to be given *or possible*.

The other major class of uncertainty that arises in information is ***lack of knowledge*** (also called epistemic uncertainty). Lack of knowledge is of even greater importance than variability in many DHS decision problems, in our opinion. As an illustration, note that lack of knowledge uncertainty is intrinsic to intelligence data. For any single source of such data, one is interested in the integrity and fidelity of the information, as well as how complete it might be. All of these characteristics are subject to lack of knowledge. When more than one source of information is involved, there can be subtle or not so subtle differences, or outright conflict, in the presented information. We are then interested in constructively resolving this conflict, perhaps through subjective weighting, data aggregation, or through the decision to acquire more information.

At a more detailed level, lack of knowledge uncertainty also arises from (1) lack of fidelity in scientific models (for example, uncertainty in the effectiveness of reduced order models); (2) lack of fidelity in the computational implementation of models (for example, uncertainty in the accuracy of coarsely zoned finite difference models); (3) from heterogeneity of required sources of information (for example, combining expert opinion with experimental data); (4) from insufficient information (for example, insufficient data to characterize the stochastic characteristics of a believed variability); (5) from deeper structural questions about scenario development (for example, in the case of polling are the “right” questions being asked); (6) and from the question of existence of unknown unknowns (for example, what is missing in a model that we have not identified).

Uncertainty Quantification Whitepaper

The use of intelligence information in target definition and consequently weapon lethality prediction, say for earth penetrating weapons, provides one example of the interplay between variability and lack of knowledge. Intelligence data may be necessary to develop a target definition for a complex bunker. Specific bunker constructions details may be vague, however. We choose the word “vague” precisely: the available information about the overall site, architecture, and composition of the bunker may involve variability and lack of knowledge factors combined in a complex manner.

- The type of concrete used in the construction may be uncertain. However, it may be adequate in computational analysis of potential lethality of a given weapon against the defined bunker to quantify this uncertainty by applying variability to the various parameters in a computational concrete model.
- The site and architecture, such as the depth, the characterization of the surrounding geology and the extent of the bunker, are lack of knowledge issues. Even if one attempts to place probability distributions on factors governing these issues, say with second order probability (see below), expert opinion is likely to be important in determining the probabilities and must therefore be included in the uncertainty quantification (Keeney and van Winterfeldt, 1991).
- In this kind of lethality assessment a conservative bounding analysis, based on conservative assumptions about depth, geology, and bunker extent, may be applied. Nevertheless, such conservative analysis then does not adequately reflect the uncertainty underlying the assessment of the bunker within the available intelligence data. As Helton has emphasized (Helton, 1994), probabilistic performance assessment (in this example, an analysis of the probability of bunker kill) requires identifying and quantifying uncertainty, not conservative bounding.
- Countermeasures, such as disruptive obstacles and false bunkers, are also a lack of knowledge issue to the extent that they are ill defined.
- Finally, assessing the destruction of a specified bunker in the event of an attack can be a lack of knowledge issue.

The combination of variability and lack of knowledge in this example is more prominent to the degree that surgical precision is required in an airstrike and direct ground engagement is unavailable.

Uncertainty Quantification (UQ)

Uncertainty must be quantified in order to use it systematically in decision-making processes. We briefly discuss quantification for both variability and lack of knowledge to emphasize issues that we think are relevant to DHS.

As suggested above, quantification of variability is already specified. This is achieved through frequensic probabilistic formalisms. We assume that sufficient information exists, or that sufficient information can be gathered, to characterize stochastic variability to perform probabilistic inference. For example, in a decision process that uses a computational model where one input is wind conditions at a given location, the UQ proceeds quite directly. Wind conditions are described in terms of a frequensic probability distribution based on empirical observation (sampling); this distribution is propagated through the computational model, typically using sampling-based statistical procedures that produce an ensemble of model calculations; the resulting frequensic probability distributions of relevant model outputs are statistically summarized; this summary is then used in the decision. The key notion here is that objective probabilistic information influences the use of the model both in terms of inputs and of outputs.

This ideal is severely limited by our ability to characterize the input uncertainty, propagate it accurately (Helton and Davis, 2000; Kleijnen, 2002; Oberkampf, 2002) and usefully summarize the results, especially when there may be *many* such inputs and the model may be very complex and computationally demanding. An important philosophical point, however is that if the associated uncertainty is believed to be variability, then these difficulties are viewed as simply influencing the accuracy of the UQ, not as questioning the nature of the uncertainty or the accepted process of UQ.

Quantifying lack of knowledge uncertainty is fundamentally harder. In the example quantification process above for variability, the first step of quantifying the uncertainty in a form that is suitable for propagation through the model is more challenging. There are three current options for accomplishing this task. First, ***second-order probability*** can be applied, where the uncertainty is still characterized by frequency distributions, but the choice of distribution is itself now viewed as stochastic. Second, ***Bayesian methods*** (Press, 2003), used here specifically in the sense of probabilistically quantifying subjective (non-frequensic) information, may be applied. For example, Bayesian methods are used to provide a probabilistic quantification of expert opinion or for attaching a probabilistic quantification to the likelihood of unobserved events (e.g. what was the "likelihood" of the 9-11 event prior to September 11, 2001?). Bayesian methods seek to attach probabilities to questions like "What is the probability that given intelligence information is correct?" Bayesian methods in addition provide systematic methods for updating these quantifications based on new information. Third, there are what we will refer to as ***generalized probability theories***, typically evidence theory, possibility theory, and fuzzy probability (Klir and Yuan, 1995; Ferson and Ginzburg, 1996; Dubois et al, 2000; Klir and Wierman, 1998; Ben-Haim, 2001). These concepts in some sense all generalize classical probability to set-valued probability distributions (rather than numerical distributions); but they also thus require certain fundamental changes in the allowed inference methods.

For all of these cases, propagating the resulting quantified uncertainty through a model is less straightforward than for variability, and inference about the resulting output uncertainty is more complex. Second-order probability characterizations are propagated using the same sampling formalisms as used for propagating first-order probability, but with associated increase in scope and computational burden to handle the probabilistic choice of distributions (Helton, 1994). Bayesian methods demand a heavy computational burden even for straightforward models; the computational challenges increase for the complex models that DHS deploys. Propagation of generalized probability is currently a research problem for these kinds of applications (Helton, Johnson, and Oberkampf, 2003). Finally, we comment that the choice of uncertainty representation can have a dramatic impact on computational complexity in even simple problems. Averbackh (2001) presents an example of an uncertain optimization problem that transforms from being polynomially hard to NP-hard due to a change in uncertainty representation.

There are important examples where propagation of quantified lack of knowledge has been successfully performed for very high consequence decisions. In particular, in the U.S. for both nuclear reactor safety assessments and in the licensing process for the WIPP facility second-order probability has been applied (Breeding et al, 1992; Helton et al, 2003). In both of these examples, the work was performed within a tight legal regulatory structure. This also demonstrates that UQ involving lack of knowledge can be effectively deployed in complex technical frameworks.

We state one caution about "unknown unknowns." UQ cannot create information from nothing. To the extent that we do not know or cannot specifically express something understood about unknown unknowns, we cannot quantify this fundamental uncertainty. From the perspective of informing important DHS decisions, UQ implements our fundamental desire to express *identified* variability and lack of knowledge uncertainty within a methodology that allows scientific assessment of their relative and absolute impact. UQ will not quantify truly unknown unknowns.

UQ Challenges for DHS

The DHS ASC program has already emphasized the importance of UQ and we do not need to reaffirm that point here. Rather, we simply emphasize several aspects of UQ for DHS that naturally follow from our discussion above:

- <u>Characterization of uncertainty</u> in information that DHS must deal with is of enormous importance. This characterization is complex if for no other reason than the vast scope of information that might play an important role in various DHS decision processes. It is likely that lack of knowledge uncertainty, will be the most prominent factor in large classes of DHS information. Therefore, the formalism for achieving this characterization is a key issue. The choice of formalism is inextricably tied to the problem being addressed by the decision, and to the nature of the decision process.

  We also speculate that critical DHS decisions will involve the use of data with heterogeneous uncertainty, a mixture of both variability and lack of knowledge that should not be completely partitioned. (Such heterogeneity probably is of the essence in intelligence data.) Formalism for UQ with heterogeneous uncertainty is immature in our opinion. Current approaches operate by either forcing the separation of uncertainty into variability and lack or knowledge or strictly subsume what we call heterogeneous uncertainty within the lack of knowledge category. Current literature (Sentz and Ferson, 2000; Chen, 200; Ben-Haim, 2001) discusses combining heterogeneous uncertainty in information.

- <u>Propagation of uncertainty</u>, for example through coupled multi-model simulation frameworks, is important to DHS. For example, to simulate economic consequences of an aerosol biological attack minimally requires a physical transport model (partial differential equation-based), an epidemiological model (possibly agent-based), and an economic consequence model (again, possibly agent-based but different than the epidemic model). For high-consequence decisions, validity of each model is a lack of knowledge uncertainty, while the predictions of the models for given inputs are likely to be intrinsically variable because of the use of stochastic solution methods. "Propagating" uncertainty in this situation is nontrivial and presumably has unique requirements based on specific DHS needs.

- Understanding the output for models incorporating uncertainty will be a challenge. Typically, for high-consequence decision making it can be difficult to understand adequately what a single complex physical model is telling us. The problem magnifies for the mass of simulation data arising from active quantification of uncertainty. Such simulation data should probably be viewed as another application of the data mining and other informatics tools of broader interest to the ASC program.

  For example, it is essential to visualize complex data sets in computational fluid dynamics to achieve intuitive understanding of the information. In time-dependent 3-D simulations, data mining tools like pattern recognition methods and immersive environments have been under study to move beyond simple summary measures of information (such as time-history of a velocity at a specified location). The problem of intuiting large-scale complex information in a single calculation is increased if we consider an ensemble of such calculations, as might be generated using UQ of variability in the underlying problem. We now need to go beyond simple statistical summary measures (such as means and variances) to develop intuitive understanding of the ensemble. There are similar examples in informatics problems, for example arising in the transformation of deterministic nodes and edges in a graph to variable quantities. Interest in the visualization of uncertain information in computational physics is just beginning [Pang, 2001].

- High-consequence decision processes are rather strongly coupled to the modeling process when uncertainty is quantified. One illustration of this principle is the current attention in the DOE NNSA Stockpile Stewardship Program given to "Quantitative Margins and Uncertainties" (QMU). The challenge will be to recognize this coupling, and identify key features that must be attended to.

  Here is one example of what we mean. UQ on the face of it, at least for variability, or for second-order probability, is straightforward computationally. However, part of the uncertainty to be quantified is the validity of the models themselves. Understanding this uncertainty influences the decisions that must be made about whether or not the models are acceptable ("good enough") for use in the given decision problems. This is a problem of ***qualification*** that, in fact, depends strongly on the details of the UQ. In this sense, UQ may provide quantitative "margins" of usefulness or acceptability of the model.

  Concern about whether a model is "good enough" in traditional computational science tends to drive conservatism in the ultimate use of the models that may be in opposition to the actual needs of DHS in various circumstances. An example of

> conservatism is "Do not use the model." We presume that this is not a suitable guidance to DHS for the most part. One way to regard this problem is to ask what kind of information would have been required *prior* to 9/11/2001 to authorize a decision to ground *all* U.S. commercial air traffic. Incorporation of uncertainty raises the importance of the qualification process and its use in DHS decisions.

How and in what sense the UQ is itself validated is an interesting question. For example, when is poor validation of a quantified variability a lack of knowledge uncertainty, or simply an accuracy problem that may be addressed by gathering more data? The choice of interpretation in this case influences the choice of corrective action to improve our knowledge.

Acknowledgement

Detailed discussions and joint work with Jon Helton, Martin Pilch and William Oberkampf of Sandia over the past seven years underlie the content of this paper. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

1. I. Averbakh (2001), "On the Complexity of a Class of Combinatorial Optimization Problems With Uncertainty," Mathematical Programming, Series A, Vol. 90, 263-272.
2. Y. Ben-Haim (2001), Information Gap Decision Theory – Decisions Under Severe Uncertainty, Academic Press, New York.
3. R. J. Breeding, J. C. Helton, E. D. Gorham, and F. T. Harper (1992), "Summary Description of the Methods Used in the Probabilistic Risk Assessments for NUREG-1150," Nuclear Engineering and Design, Vol. 135, 1-27.
4. Y. Y. Chen (2000), "Fuzzy Analysis of Statistical Evidence," IEEE Transactions on Fuzzy Systems, Vol. 8, No. 6, 796-799.
5. D. Dubois, H. T. Nguyen and H. Prade (2000), "Possibility Theory, Probability and Fuzzy Sets," in Fundamentals of Fuzzy Sets, ed. D. Dubois and H. Prade, Kluwer, Boston, 343-438.
6. S. Ferson and L. Ginzburg (1996), "Different Methods are Needed to Propagate Ignorance and Variability," Reliability Engineering and System Safety, Vol. 54, No. 2-3, 133-144.
7. J. C. Helton (1994), "Treatment of Uncertainty in Performance Assessments for Complex Systems," Risk Analysis, Vol. 14, No. 4, 483 – 511.

8. J. C. Helton and F. J. Davis (2000), “Sampling-Based Methods for Uncertainty and Sensitivity Analysis,” Sandia National Laboratories report SAND99-2240.
9. J. C. Helton (2003), “Mathematical and Numerical Approaches In Performance Assessment for Radioactive Waste Disposal: Dealing With Uncertainty,” in Modeling Radioactivity in the Environment, ed. E. Marian Scott, Elsevier, pp. 353-390.
10. J. C. Helton, J. D. Johnson and W. L. Oberkampf (2003), “An Exploration of Alternative Approaches to the Representation of Uncertainty in Model Predictions,” Preprint.
11. S. Kaplan and B. J. Garrick (1981), “On the Quantitative Definition of Risk,” Risk Analysis, Vol. 1, No. 1, 11-27.
12. R. Keeney and D. von Winterfeldt (1991), “Eliciting Probabilities from Experts in Complex Technical Problems,” IEEE Transactions on Engineering Management, Vol. 38, No. 3, 191-201.
13. J. P. C. Kleijnen (1997), “Sensitivity Analysis and related analyses: A review of some statistical techniques,” J. Stat. Comput. Simul., Vol. 57, No. 1-4, 111-142.
14. G. J. Klir and B. Yuan (1995), Fuzzy Sets and Fuzzy Logic – Theory and Applications, Prentice-Hall, New Jersey.
15. G. J. Klir and M. J Wierman (1998), Uncertainty-Based Information, Springer-Verlag, New York.
16. W. L. Oberkampf, et al (2002), “Error and Uncertainty in Modeling and Simulation,” Reliability Engineering and System Safety, Vol. 75, No. 3, 333-357.
17. A. Pang (2001), “Tools for Comparative and Uncertainty Visualization,” presentation at Sandia National Laboratories.
18. S. J. Press (2003), Subjective and Objective Bayesian Statistics – Principles, Models, and Applications, Wiley-Interscience, New York
19. K. Sentz and S. Ferson (2002), “Combination of Evidence in Dempster-Shafer Theory,” Sandia National Laboratories report SAND2002-0835.