Epistemic Uncertainty in Agent-based Modeling

Scott Ferson[1)] and Kari Sentz[2)]
[1)] *Applied Biomathematics, USA, scott@ramas.com*
[2)] *Los Alamos National Laboratory, USA, ksentz@lanl.gov*

Abstract: Traditional approaches to handling uncertainty in agent-based models employ Monte Carlo methods to randomly sample parameters and probabilistically determine whether and how a behavior or interaction rule is realized by an individual agent. A simulation of all agents thereby represents a single realization from among many possible scenarios, and simulations with many replications are used to reveal differential probabilities and the likelihoods of extreme results. Unfortunately, Monte Carlo is a poor way to project epistemic uncertainty through a complex model, and it is an unsatisfying scheme for representing the uncertainty about volitional choices of agents. Adding epistemic uncertainty to agent-based models properly requires the ability to (1) characterize stochastic drivers imprecisely, (2) specify agent attributes and other quantities as intervals, probability distributions, or p-boxes, and (3) execute behavior rules in a way that respects uncertainty in their conditional clauses. When uncertainty makes the truth value of the conditional clause of any rule unclear, the simulation should hold that the rule both fires and does not fire. This may result in subsequent uncertainties elsewhere in the simulation including the status of attributes of agents, even perhaps whether an agent exists or not. These facilities advance agent-based modeling to uncover a more comprehensive picture of the effects of epistemic uncertainty, which can be vastly more important than aleatory uncertainty. We compare this approach with traditional simulation using only Monte Carlo methods to reveal the differences between these two approaches to uncertainty.

Keywords: agent-based models, epistemic uncertainty, Monte Carlo simulation, p-boxes, intervals

1. Introduction

Predictive failures are more common in probabilistic modeling and simulation than they should be. For instance, NASA estimated the risk of catastrophic failure of its Space Shuttle to be 1/10,000 per flight, but its observed failure rate was about 1 per 70 flights (Hamlin et al. 2011; Oberkampf and Roy 2010). The National Weather Service and the Army Corps of Engineers strongly underestimate risks of "100-year floods" and miscommunicate uncertainties about flooding risks (Morss and Wahl 2007; NRC 2006; NWS 1998). Observed failures and near-misses found at the Diablo Canyon Nuclear Power Plant reveal gross understatement of its assessed risks (Lochbaum 2011). Failure cascades in electricity distribution systems are more common than they are forecasted to be (RAWG 2005; USCPSOTF 2006). Probabilistic assessments understating risks in the financial industry precipitated the Great Recession (Savage 2012; Silver 2012). These are not random rare events in a discipline that estimates and controls risks well overall. They are not the unlucky but expected tail events sometimes called "noble failures". They seem instead to be wholesale errors which are the result of pervasive and systemic errors that undermine the credibility of all modeling and simulation efforts. The comment by a prominent engineer that the 2007 bridge collapse in Minneapolis was a "billion to one" event typifies the raw fact that uncertainties are sometimes not well understood in engineering.

What can be done about this situation when such predictive failures are intolerable? We believe that a pervasive cause of these predictive failures is neglecting or mishandling uncertainties. This paper focuses on practical methods for introducing uncertainty into agent-based models because such models are widely used and growing in importance within engineering. The problem of correctly handling uncertainties in this context also seems to be rather subtle.

2. Kinds of Uncertainty

In the past, uncertainty analysis considered the source of uncertainty to be its salient aspect, so modelers talked, for example, about their parametric uncertainty or their model-form uncertainty. A more modern view is that the nature of the uncertainty, rather than its source, is a more important characteristic. We can distinguish between two main forms of uncertainty: epistemic and aleatory. Aleatory uncertainty refers to the variability or stochastic fluctuations in a quantity through time, variation across space, manufacturing or genetic differences among components or individuals, or similar heterogeneity within some ensemble or population. The word aleatory comes from *alea*, the Latin word for dice. This is considered to be a form of uncertainty because the value of the quantity can change each time one looks, and we cannot predict precisely what the next value will be (even though the distribution of values may be known). Epistemic uncertainty, on the other hand, refers to the lack of full knowledge about a quantity that arises from imperfect measurement, limited sampling effort, or incomplete scientific understanding about the underlying processes that govern its value.

These two forms of uncertainty have important differences. Epistemic uncertainty can in principle be reduced by empirical effort; investing more in measurement and study of a system should yield better precision. Aleatory uncertainty, in contrast, can sometimes be better characterized, but cannot generally be reduced by empirical effort. Epistemic uncertainty depends on the observer and the observations made. Aleatory uncertainty does not depend on an observer at all. It exists whether or not anyone witnesses it, like the sound waves emanating from the proverbial tree falling unseen in the forest. Although epistemic and aleatory uncertainty can sometimes be like ice and snow in that their distinction can be difficult to discern through complicating details (and sometimes one can change into the other depending on the scale and perspective of the analyst), the macroscopic differences between these two forms of uncertainty are usually obvious and often significant in practical settings.

3. Agent-based Modeling

Agent-based models (Wooldridge 2009) arose from early research on cellular automata (Gardner 1970) when computing first became widely accessible. These models were generalized to allow individual objects to move about on the plane or in space. Simulations also endowed the objects with goals and the ability to make choices in complex environments, and even emulated how agents might learn from their past interactions and experiences to adapt and evolve within the simulation. As a consequence, agent-based models have the potential to capture emergent behavior that was not conceived *a priori*. Most recently, the models have considered more than a single kind of entity interacting with each other. These include agent-based models, multi-agent simulations, individual-based models, and swarm models. Today, these heterogeneous agent-based models are considered a new paradigm for design in computer science with self-structuring software systems (Shoham and Leyton-Brown 2009), and models for distributed decision-making and problem solving (Weiss 2001).

The practical applications of agent-based models are many and include fleet and supply-chain management, logistics, portfolio planning, forestry, biological resource management, epidemiological modeling, and analysis of coalitions and joint mission planning. Ilachinski (2004) described the use of agent-based models in modeling military conflicts, which is natural because of the importance in that context of nonlinear interactions among elements organized in command-and-control hierarchies of often fluctuating coherence. Simulations using autonomous agents to model individual behaviors represent a bottom-up approach to synthetic modeling. They can complement or perhaps supersede conventional analytic and reductionist approaches, which may be less useful in the study of emergent patterns of self organization that commonly arise in complex systems of systems. Although much effort has been devoted to developing this approach, and to combining it with uncertainty analysis (e.g., Wu et al. 2003), further work is needed, particularly workable schemes to express and propagate not just aleatory but also epistemic uncertainties about agent status and rule outcomes through simulations.

4. What's Wrong with Sampling

A traditional sampling approach to uncertainty projection in engineered systems routinely and often profoundly underestimates the overall uncertainty that should be associated with outcomes (Atwood 1985; Ferson and Ginzburg 1996; Oberkampf and Roy 2010; Ferson and Siegrist 2012; Savage 2012). Figure 1 illustrates this underestimation, showing Monte Carlo results in a tangle of trajectories whose overall uncertainty grossly underexplores the true range of possibilities arising from the epistemic uncertainty about the outcome circumscribed by outer bounds. The actual uncertainty is often much wider, sometimes much larger than that predicted by ordinary probability methods such as Monte Carlo simulation, Latin hypercube simulation, or second-order Monte Carlo. The outer bounds are not necessarily worst-case outcomes, but they may be envelopes of the plausible scenarios in a proper accounting of uncertainty that, unlike traditional Monte Carlo approaches, does not make unjustified assumptions about independence among variables, and does not use equiprobability or uniform distribution assumptions to represent incertitude (empirical ignorance). We can sometimes observe the difference between the Monte Carlo simulations and the outer bounds in Figure 1 to be several orders of magnitude. Sometimes the Monte Carlo results are in the middle of the true possible range of outcomes on some scale as depicted in the figure, but this is not always the case. While the Monte Carlo results may ‘fill up’ the possible range for some intervals in time dependent models, this is rare. When it does happen in time-dependent models, it is often an ephemeral occurrence at the start of a simulation or at pinch points or ranges were the direction of trajectories shift. Monte Carlo sees uncertainty through a glass darkly, in general. Variants of Monte Carlo methods are only marginally different. Latin hypercube sampling, for example, is usually worse at identifying extreme events. Second-order Monte Carlo does generally produce slightly wider results, but these trajectories form a halo around the regular Monte Carlo results, and they also come nowhere near to filling up the range of possible outcomes in most cases.

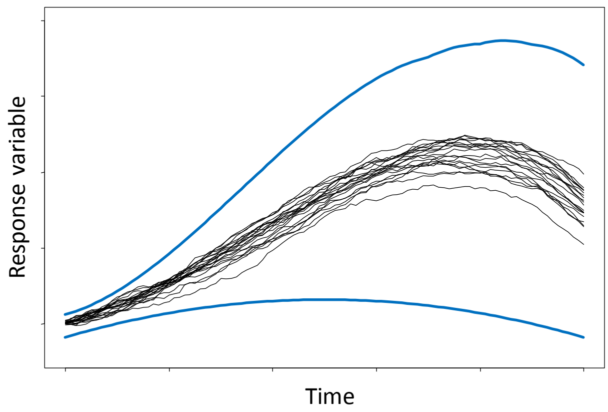

Figure 1. Discrepancy between the true possible range (outer bounds) and Monte Carlo realizations.

Often, what strategists and planners actually need to know is the range of possible outcomes, mostly because they are especially concerned with the extreme cases. There are the outer bounds shown in Figure 1. Unfortunately, Monte Carlo simulation is well known to be a very poor technique for uncovering this range. David Alderson (Alderson et al. 2012) of the Naval Postgraduate School has noted that “Random sampling is terrible at finding worst-case scenarios, although terrorists are pretty good at it.” He was talking about estimating the reliability of transport networks and supply chains and their susceptibility to malevolent disruption, but the observation is actually quite general.

The problem is related to what is sometimes called the “curse of dimensionality”. It can be illustrated with a very simple example. Suppose we are interested in estimating the possible range of the sum of 30 uncertain values each constrained to the unit interval. It is clear that this range is an interval between zero and 30. Random sampling, in which uniform random deviates independently sampled from each of the 30 unit intervals are added together, reveals a range that is almost always much smaller. If there are 100 such sample summations adding 30 deviates together, the output range of the sum is roughly one third of what it should be. Even with a million replications, the observed range of sums is only half as wide as it should be. Using different distributions other than uniform or even systematic rather than uniform sampling for the addends does not help widen the result appreciably. This underestimation worsens dramatically as the number of inputs grows.

Some analysts argue that the larger bounds are unreasonable characterizations of uncertainty and insist that the narrower tangle of Monte Carlo outputs for the example above is actually correct. They say that it would be silly to suppose that all 30 of the addends could be extreme in the same direction at the same time. Such analysts will often be surprised by the outcomes of real-world systems, because this constellation of extremes can indeed be realized in the outcomes of real-world systems. Disasters and extreme events are often born when multiple things go wrong together. Before the 2007 financial crisis, many analysts asserted—and apparently also really believed—that it would be essentially impossible for a significant fraction of the hundreds of mortgages bundled together in securities to all default. But the financial analysts neglected to account for the dependence relationships in risk among these mortgages that made them far from independent. Contracting home markets affected lots of mortgages at the same time, and there were unfortunate positive feedbacks that made the situation worse. The more mortgages that defaulted, the harder it became for any homeowner to stay above water. These examples are very simple instances of the general problem. When the elements become more complex than mortgages that either default or do not, and the functions relating them together become more complicated than simple addition, the potential for dependencies to play a dominant role in the overall risk of failure is both more likely and harder to foretell.

Why do regular probabilistic analyses often badly underestimate risks? There are several reasons, but the most common and significant ones are

- Unjustified use of independence assumptions or overly precise dependence assumption,
- Inappropriate use of equiprobability and uniformity assumptions,
- Underestimation of the uncertainties of original measurements,
- Modeling volitional choices with distributions as though they are random,
- Using averaging as an aggregation, and
- Making assumptions for the sake of mathematical convenience (wishful thinking).

For example, standard approaches based on probability theory almost always assume independence among variables, even when there is little evidence or argument to support such an assumption. In fact, when there is no information about a dependency, many consider this to be a reason to assume independence. This is profoundly wrong in any risk-analytic or uncertainty-analytic context, and the consequences of this are underestimated risks and incorrect assessments of uncertainty. The standard probabilistic approaches grossly understate the uncertainty when they fail to account for the correlations, dependencies, and interactions among the components and among stochastic driver variables (Hartung et al. 1985; Hickman et al. 1983; Hokstad and Rausand 2008). These connections that join together the elements of the model are frequently neglected in the modeling exercise. In principle, each such connection should be characterized. There are two reasons that such a task is generally extremely hard: First, there are often many dependencies to worry about. If there are n elements in a system, there may be up to $(n^2 – n)/2$ pairwise interactions between those elements, and this is just the beginning. There may also be triple-wise or even higher-level interactions as well, which are essentially never even considered in most models. The second reason it is difficult to address these interconnections in modeling is that dependence between stochastic variables can be complex. If we’re lucky, one inter-variable dependence can be fully characterized by a simple real value like a correlation coefficient, but this is often not really possible because the dependency is nonlinear. We know that such complex interactions can and do exist among many variables, but human intuition is not well adapted to consider them, especially when nonlinearities, trade-offs, feedbacks, problem cascades, and other complexities complicate the story.

Poor models of dependences are not the only important reason that naïve Monte Carlo methods understate uncertainty. Space limits preclude a full discussion of each of the other causes mentioned above. Ferson and Ginzburg (1996) review the problems that arise for risk analyses from using assumptions of equiprobability or uniform distributions to model incertitude. Henrion and Fischhoff (1986) review the apparently pervasive and systemic underestimation of measurement uncertainties for physical constants. The situation is probably even worse for routine test-site and field measurements (Morgan and Henrion 1990; Youden 1972). On account of these issues, exclusively using Monte Carlo or similar traditional probability techniques in risk, reliability, or uncertainty assessments for critical infrastructures is unsound when extreme events are the focus. This crosses over to the ridiculous for modeling behavior when malevolent human actors are involved such as in warfare, terrorism, arson, crime, and competition. Modern approaches (Beer et al. 2013; Ferson and Ginzburg 1996; Walley 1991) to handling epistemic uncertainty that treat it differently from aleatory uncertainty and respect epistemological limits on inferences are designed to remedy these problems and provide reliable characterizations and conclusions that do not depend on making assumptions that analysts cannot justify or do not even believe.

Considerable progress has been made recently in developing algorithms for numerical calculations that respect and distinguish epistemic and aleatory uncertainty (Beer et al. 2013; Ferson and Siegrist 2012; Sentz and Ferson 2011). Figure 2 below shows several examples of 'uncertain numbers', where the abscissas are arbitrary axes of real values. A *probability box* or p-box (depicted at the far right) is the most general form because it has both aleatory and epistemic uncertainty. In Bayesian analyses, this structure is also called a distribution band (Berger 1994; Basu and DasGupta 1995). They are closely related to what frequentists call meta-distributions or second-order distributions (Vose 2008), and to the collections of probability distribution functions known in the theory of imprecise probabilities as credal sets (Walley 1991). P-boxes arise in robust Bayes analyses in which there is doubt about the prior that should not be neglected or condensed into a single distribution, or when there is non-negligible measurement error about input data that induce a family of likelihoods rather than a unique one. P-boxes can also arise from partial constraint information such as might be available from an incomplete or faulty set of sensors. A p-box is defined by left and right bounds, which delimit the class of probability distributions denoted by the p-box as in the rightmost graph in Figure 2. When those edges coincide, the class has only a single member, which is a precise cumulative probability distribution such as the second graph from left in Figure 2. This distribution represents aleatory uncertainty but not epistemic uncertainty. In contrast, the edges of the p-box could be rectangular in which case the p-box degenerates to an interval (third graph in Figure 2), which might consist entirely of epistemic uncertainty. Finally, the leftmost graph in Figure 2 depicts a known, fixed quantity, which we might call a scalar. It is a completely degenerate kind of uncertain number in that it has neither aleatory nor epistemic uncertainty.

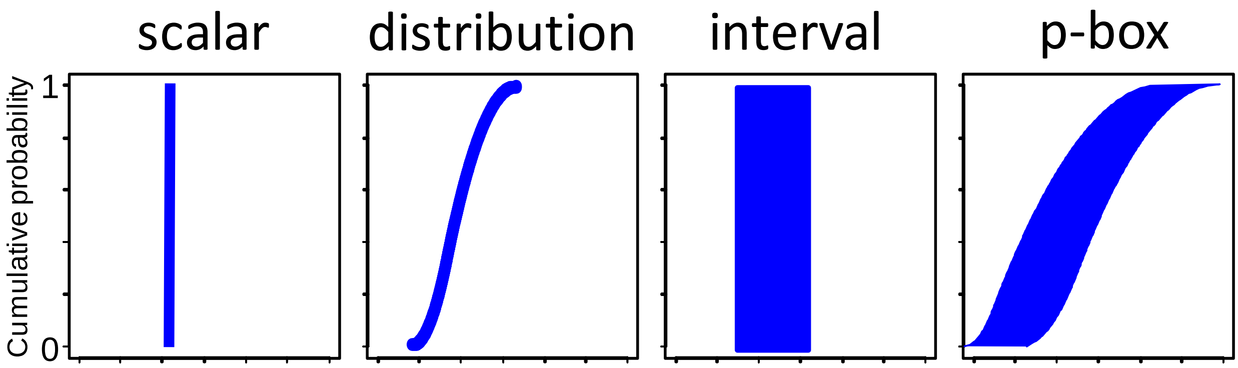

Figure 2. Bestiary of uncertain numbers.

We know how to do math on these uncertain numbers. Mathematical operations such as addition, subtraction, multiplication, division, minimum, maximum, powers, logs, roots, etc. have been worked out and implemented in stable software (Ferson et al. 2003; Ferson 2002). These operations are closed in the set of uncertain numbers, except in obvious cases that involve division by zero or taking the log or root of negative values. The results can be proven to be rigorous in that the bounds are guaranteed to surely contain the actual value or distribution given the uncertainty. Moreover, in many cases defined by well understood conditions, the results can also be proven to be best possible, which is to say that the bounds cannot be made any tighter without reducing the uncertainty about the inputs. P-boxes are a relatively simple way to express both aleatory and epistemic uncertainty about quantities; they are much simpler than credal sets from the theory of imprecise probabilities (Walley 1991), fuzzy numbers (Kaufmann and Gupta 1985) or more general info-gap structures (Ben-Haim 2001), or even Dempster–Shafer structures (Ferson et al. 2003). This simplicity gives analysts great power to do practical computations at moderately large scales that are simply not achievable under the more delicate theories. The application of p-boxes as models of uncertainty in practical assessment problems is called probability bounds analysis. The Wikipedia article on the topic (http://en.wikipedia.org/wiki/Probability_box) lists dozens of applications published over the last decade. The p-box approach can be used to improve uncertainty analysis and the analysis of risk assessments in general by providing automatic sensitivity analysis for a probabilistic assessment.

5. Generalizd Agent-Based Modeling

There has been considerable research in the interface between uncertainty analysis and agent-based modeling (Faanes and Skogestad 2004; Zellner 2008; Bobashev and Morris 2010; Hancock et al. 2010; Hassan et al. 2010; Harp and Vesselinov 2011; Fonoberova et al. 2013; Magliocca and Ellis 2013), but it has focused exclusively on using agent-based models to analyze aleatory uncertainty, rather than what is proposed here, which is to project both forms of uncertainty through agent-based models.

Agent-based models represent and follow agents or meta-agents, which are composed of multiple, partially independent agents. Every agent consists of a name or identifier and a list of attributes, each of which is a vector of real or categorical values that characterize its status. The models also include rules governing the temporal evolution of agents and interactions among agents within the simulation. There may also be environmental variables such as weather shared by all agents that are determined by exogenous stochastic drivers that introduce randomness into the system. Traditionally, the behaviors and interactions of agents are simulated using randomly sampled parameters that probabilistically determine whether and how a rule is realized by an individual agent. A simulation thus represents a single realization from among many possible. Each simulation is a plausible scenario without information about the likelihood of that scenario. Monte Carlo replications of such simulations are typically employed to show differential probabilities and the likelihoods of extreme results.

As explained above, however, Monte Carlo is a very poor way to project epistemic uncertainty through a complex model, and an even worse way to represent the uncertainty about the volitional choices of friends or especially foes. Adding epistemic uncertainty to agent-based models in a proper way will require the ability to

1) Characterize stochastic drivers imprecisely,
2) Specify agent attributes and other quantities as uncertain numbers,
3) Execute rules in a way that respects uncertainty in their conditional clauses, and
4) Accept user specifications of uncertain numbers.

We address each of these facilities below. Through their development, agent-based modeling based on Monte Carlo methods can be generalized and expanded to uncover a more comprehensive picture of the effects of epistemic uncertainty.

6. Uncertainty about Stochastic Drivers

The traditional way to characterize a modeled variable that is influenced by stochastic drivers is to specify a precise probability distribution for the variable. Such distributions are usually derived in one of three ways: statistical analysis of available sample data collected about the variable, a forecast about it from related information, or informed judgment produced by the modeler or elicited from experts. Yet these distributions are rarely known with the confidence that is implied by employing a precise specification. For instance, there can be substantial epistemic uncertainty about diurnal, seasonal and longer-term patterns of weather and space weather that may affect threat detectability and response speeds of movable assets. We know future environmental conditions will be variable, but planners cannot forecast their precise probability distributions even in the short term (cf. Kharin 2009). The further into the future the prediction must be made, the broader the uncertainty will likely be. Conscientious assessments today characterize the epistemic uncertainty about probabilistic forecasts derived from quantitative models. One of the simplest ways to do this is with a p-box (Ferson et al. 2003). Even when the distribution is thought to be stationary and there are actually random sample data available to estimate its shape and parameters, the sample size may be limited, which means there is inferential uncertainty about any empirically estimated distribution. Balch (2012; Ferson et al. 2013) described ways to derive p-boxes from random sample data that respect and propagate this empirical uncertainty. P-boxes can also be obtained directly in expert elicitations, either from introspection by a single expert, or by forming an envelope of the distributional predictions from several experts (Linkov and Burmistrov 2003; Sentz and Ferson 2002; Ferson et al. 2003).

If we use a p-box rather than a precise probability distribution to characterize a stochastic driver, how would the agent-based model use the p-box in simulations? There are essentially two ways, depending on the nature of the uncertainty projection that is intended. The left graph in Figure 3 below depicts a precise cumulative probability distribution for a random variable that ranges between 3 and 6. Also depicted in the graph is a random deviate being generated from the distribution via inverse transform sampling (a uniform random number between zero and one on the probability axis is transformed into a random deviate from the distribution by mapping it to the abscissa through the cumulative distribution). The middle graph in the figure depicts a p-box over the same range that might express the epistemic uncertainty about the distribution in question. The first way to take into account the epistemic uncertainty of the p-box is to generalize the inverse-transform sampling approach that generates random deviates. Instead of producing a single scalar number for each uniform probability value, the p-box yields an *interval* of possible values for each probability value. The interval deviates can then be used in agent-based modeling software augmented with the ability to compute with intervals just as well as regular scalar numbers.

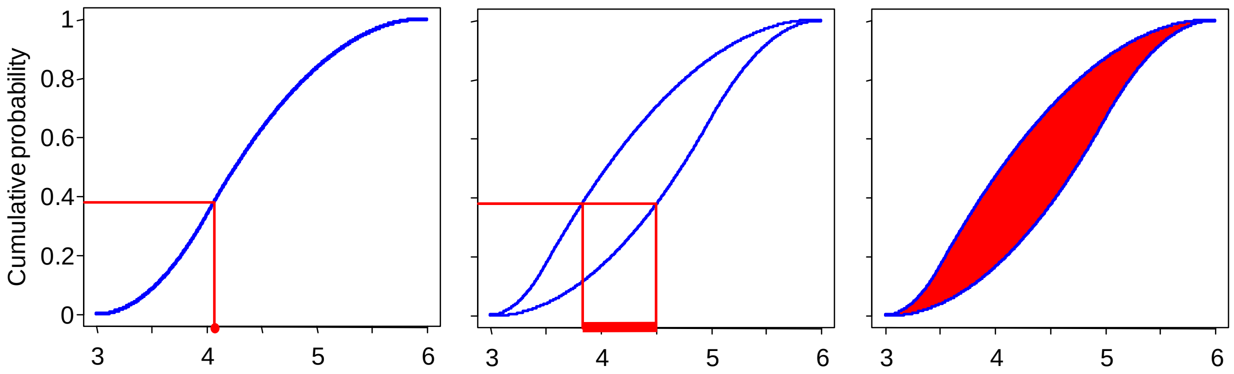

Figure 3. Three depictions of sampling aleatory uncertainty about a stochastic driver.

The second way to account for epistemic uncertainty in the p-box is to project the *entire p-box* through the calculations in the model. This idea is depicted in the rightmost graph in Figure 3 above. Doing this characterizes all possible scenarios at the same time, each weighted by their respective probabilities as well as can be specified under the acknowledged epistemic uncertainty of the p-box. This way of handling the p-box produces no additional complexity. In the exceptional case lacking epistemic uncertainty, where the stochastic driver is characterized by a precise distribution, the first way produces a single scalar number and the second way produces the whole probability distribution. In the other special case lacking aleatory uncertainty, where the driver is characterized by a simple interval, the first and second ways both produce the same deviate, which is the whole interval itself. Whatever is generated, whether it is a point value, an interval, a distribution, or a p-box, the structure can be projected through the calculations that depend on this stochastic driver.

7. Uncertainty in Agent Attributes

Attributes of the various agents are updated during the model simulation. In general, attributes are updated in one of two ways. Either the new value of the attribute is computed and assigned to that attribute replacing any value already there, or a modification term or factor is computed and then combined with the current state of the attribute. The new value can be any uncertain number, i.e., a scalar, probability distribution, interval, or p-box. For instance, the mass of a missile may be one of its attributes which should be modified if the missile releases decoys or chaff. The amount of the modification depends on an estimation of the mass of the released material. Note that, when an attribute's value is to be replaced, the uncertainty in the current value is irrelevant, but if the current value is to be modified, its uncertainty cannot thereby decrease. When an attribute has more than a single dimension, such as the current position of an agent specified by a vector consisting perhaps of its longitude, latitude and elevation, issues of dependency in the uncertainties of those components can arise (Miranda et al. 2013). In many cases, modelers will account for these dependencies in the rules they specify that govern agents' evolution. For instance, when a sea-surface asset is attacked, its sensors may perform differently or go off line. There may be epistemic uncertainty about these dependencies, and in principle these dependencies can be handled in the same way as other uncertainties in dependencies between variables, as modeled in probability bounds analysis (Ferson et al. 2004).

There may be other variables besides the attributes of agents and stochastic drivers that the software must update and which may therefore also have uncertainty. For instance, tallies, traces and synopses used to report the results and predictions of the simulation must be computed. Also, if many rules use the same calculation, the modeler may chose for efficiency's sake to define a parameter within the simulation to the hold the result of the calculation so that it can be used by all the rules without recomputing.

8. Uncertainty in Rule Conditions

Agent-based modeling consists of a system of rules specified by the modeler that govern the behavior and evolution of the agents. A simple rule typically has two parts, the condition and its consequence. The condition is the *if* part that determines whether the consequence is to occur. The consequence is the *then* part that specifies what should happen if it does. Compound rules can have an *else* part that specifies what should happen otherwise. The previous section described how agent attributes and other parameters can be updated with uncertainties when rules are applied that replace or modify their values. This section will review how uncertainty can be propagated through the *if* parts of the rules.

Consider a situation in which there is uncertainty about the inputs used in a binary rule, say, to fire an asset or not to fire the asset. Suppose that the rule is stated as

If $A < \theta$, then fire, else don't fire,

where A is the value of some environmental parameter or an attribute of some agent and θ is a threshold set by the modeler who specified the rule. When the uncertainties about A and θ are both characterized as intervals, there are only three cases for the condition $A < \theta$. The first case is when the uncertainty about A and θ are such that it is perfectly clear that $A < \theta$ is true. This case is depicted in the left graph in Figure 4 below, where the horizontal axis gives the magnitudes of the uncertain numbers. The value of the rule's condition is one, representing true, so we fire. The second case, depicted in the middle graph of Figure 4, is when the value of A is surely larger than the threshold θ despite their respective uncertainties. The value of the condition $A < \theta$ in this case is zero, for false, so we don't fire. The third case is when the two uncertainties overlap so that we cannot be sure whether A is larger or smaller or equal to the threshold θ. In this case, we say the value of the condition is the interval [0,1], which we call the 'dunno interval'. Because of the epistemic uncertainty about the terms, we do not know whether we should fire or not.

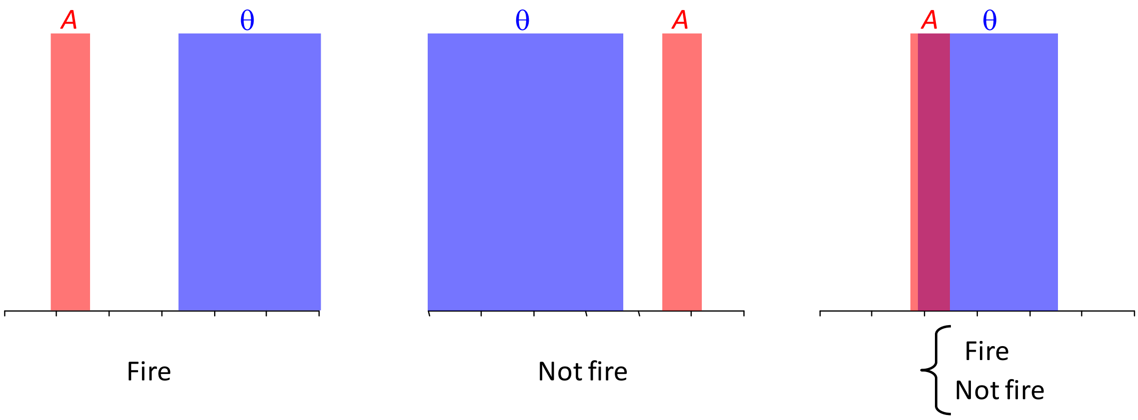

Figure 4. Three possibilities about whether A is less than θ given their respective uncertainties: yes, no and not sure.

This epistemic uncertainty about the rule's condition should be fully propagated through the simulation. Doing so means that we fire and not fire at the same time. This contradiction can be resolved in software by modeling two universes, one where the action of firing the asset occurs and one where it does not. But instead of following two diverging timelines requiring separate instantiations, the software will translate the consequences of the ambiguity of firing/not firing *back into intervals* in the attributes of the agents involved. This might mean, for instance, that the agent which was the target has status of both destroyed and intact. And, unlike the proverbial cake, we both consumed the asset by firing it, and yet still have it to fire later. Any behavior or result that depends on, or could be produced by, an agent is multiplied by its existence attribute so as to project the uncertainty about whether the agent exists to any outcomes that depend on the agent. If the agent surely exists, its existence attribute is 1, and outcomes are unaffected. If the agent surely does not exist, the attribute is 0, and the outcomes are nullified. If the existence attribute is dunno, the outcome is the interval between the full-strength outcome and a nullified outcome.

Whether the resulting ambiguity in attributes of agents matters or not depends on the dynamics of the system and the other uncertainties that may be present. This approach to agent-based modeling requires a new constitutive attribute 'exists', which, like an agent's name or identifier, must always be present in order to handle such cases. Aside from all the agents that surely exist, the software can also keep track of some agents that may or may not exist. We do not expect this approach to epistemic uncertainty in agent-based models to predict the future perfectly. What it should do is capture all possible scenarios given the available information and model structure and compute best-possible bounds on the probabilities associated with those scenarios. This harkens to the film *Men in Black 3* and its character Griffin who is a "fifth-dimensional" alien who sees all possible futures at the same time. He cannot tell what will happen but only what the possibilities and probabilities are, which is still very useful. It seems to us that Griffin's information is much more valuable than predictions from an oracle that are precise but often wrong, as are the results of traditional agent-based models. Predicting the future is dangerous because we sometimes believe our predictions, so the predictions ought to be guaranteed in some sense and not just merely possible.

Of course, the A and θ parameters may not always be intervals. They could also be probability distributions or p-boxes, in which case the resulting value of the rule's condition may not be one of the three simple logical values (zero for false, one for true, or the interval [0,1] for dunno). If aleatory uncertainty is present, issues of dependence between the terms of the comparison arise. Modelers should specify whether and how the parameters are dependent if they have this knowledge. The default behavior of the software is to presume their dependence is unknown and to use the Fréchet (1935; Hailperin 1986) inequalities to compute the logical value of the condition. The requisite mathematical framework for defining the logical values of magnitude comparisons between arbitrary p-boxes has been worked out, and stable software exists to compute them (Ferson 2002).

It turns out that magnitude comparisons among uncertain numbers of any kind produce at worst only intervals. Comparing uncertain numbers with the $<$, $>$, $\leq$, and $\geq$ operators in general yields interval logical values (Hailperin 1986). The exceptions are when (*i*) both operands are scalars, in which case the result is a Boolean value, (*ii*) one operand is a scalar and one is a probability distribution, in which case the result is a frequency, and (*iii*) both operands are probability distributions and they are independent or have another precisely specified dependence function, in which cases the comparison yields a frequency. For example, the comparison 3 < uniform(2,6) gives the scalar 0.75, but the comparison [1,3] < uniform(2,6) gives the interval [0.75, 1]. Assuming the operands are independent, the comparison uniform(2,4) < uniform(1,3) gives the scalar 0.125, but *without* an assumption about their dependence, the comparison instead gives the big interval [0, 0.5]. As we saw above, comparing intervals that overlap at more than a single point always produces the dunno interval [0,1]. No matter what kind of uncertain numbers are compared, the result is always a scalar or an interval and never anything more complex. Rule conditions that produce these scalars and intervals are also translated to existence attributes as appropriate. Existence attributes are uncertain logical values and may take the form of any p-box on the unit interval, including scalars and arbitrary subintervals of [0,1]. The same multiplication convention used to project dunno uncertainty can be used to propagate the epistemic and aleatory uncertainty encoded as interval or scalar frequencies and probabilities. Despite the simplicity of magnitude comparisons always producing interval logical values, the value in the condition of a rule may not be an interval. If the rule is expressed in reference to an event which is described by its probability, the value of the condition is in general an uncertain logical value. These can include probability distributions and p-boxes on the unit interval. Our strategy for handling uncertainty in the rule condition for such cases is the same as was used before for magnitude comparisons. The uncertainty, which can include both epistemic and aleatory features, implies an *ensemble of ensembles* of possible universes expressing a class of different frequencies. We do not have to instantiate or even envision the entire population of possible universes. Indeed, we could not do so. Instead, we translate the uncertainty from those possible universes back into the existence attributes. For magnitude comparisons, these attributes were only intervals, but in the general case with arbitrary conditions, they will be uncertain logical values characterized by p-boxes. The multiplication convention can again be used to project the uncertainty whenever an agent acts or produces something that affects other agents or parameters.

The general issue of dependency in agent-based modeling is addressed in various ways. It is usually the job of the modeler to understand and represent in the model the basic functional dependencies that tether agents together and link stochastic processes, including relevant command-control chains, cooperations and collaborations, agent hierarchies, learning processes, cross-referential decision making, and other nonlinear interactions and mission dependencies. The software will have a sufficiently flexible language for specifying rules to develop these dependencies whenever they are understood by the modeler. However, we believe that modelers often cannot be aware of all the dependencies in a system that may become significant in some context. For this reason, the software will have default features that produce conservative treatments of uncertainties even when the modeler lacks details about the interdependencies among the components of the model. When a modeler is confident that two processes cannot interact, because they are distant in space or time or causation, this fact can be captured to improve (tighten) the propagation of uncertainty by using independence assumptions. But, when the modeler does not make a conscious decision to declare the processes independent, they are treated by the software as potentially interacting. The software does this by employing mathematical operations that generalize the classical Fréchet inequalities which make no assumptions whatsoever about inter-variable dependencies (Frank et al. 1987; Williamson and Downs 1990; Ferson et al. 2003).

Compound conditions consist of functions of logical quantities combined with AND, OR and NOT operations. The mathematical definitions and software implementations needed to handle compound conditions composed from elements with epistemic and aleatory uncertainty are exactly the same as are used in the analysis of fault trees, and present no special problems in the context of agent-based models. Whenever those elements—which are the operands of the logical operations—are uncertain logical values, whether coming as magnitude comparisons (which are always intervals) or arbitrary p-boxes on the unit interval, the logical functions built from conjunctions (AND), disjunctions (OR) and negations (NOT) are always still uncertain logical values. At worst, these are p-boxes on the unit interval which we have already described how to handle. The issue of stochastic dependence arises whenever multiple logical elements are combined, and the software supports a fairly rich array of possible assumptions about these dependencies, including independence and the Fréchet case making no assumption at all, perfect and opposite dependence, positive- and negative-sign dependence, and dependence parameterized by various indices including Pearson, Kendall and Spearman correlations coefficients (Ferson et al. 2004; Nelsen 1986; 1987).

9. Can Uncertainty Grow Too Wide?

This paper has argued that the uncertainty associated with results from probabilistic modeling and simulation is often underestimated, particularly with agent-based models. Is it possible that, in our zeal to not underestimate uncertainty, we might overestimate it? Can the uncertainty grow wider than in mathematically justified? In fact, there are reasons one might worry about this.

Although it can always be guaranteed to rigorously enclose the actual uncertainty, naïve application of bounding methods of uncertainty propagation can sometimes yield results with unnecessarily inflated uncertainties (Beer et al. 2013). In such cases, the uncertainty is not characterized in a best-possible way and is overly conservative. This inflation usually arises either from repeated variables (Manes 1982), which the interval analysis community knows as the "dependency problem", or from the "wrapping effect" (Moore 1966) which can occur in multivariate problems when there is inter-variable dependence does not fit well into a rectangular multivariate box. Uncertainty inflation problems may be common and dominating in some complex calculations, and, if they are, special methods must be employed to keep the artifactual inflation of uncertainty to a minimum. Practical approaches to handling the problems have been worked out in several fields where the repeated variable problem has been encountered, including finite-element models (Zhang et al. 2012), differential equation models (Enszer et al. 2011), statistical analyses (Ferson et al. 2007; Nguyen et al. 2011).

We have already mentioned that the evaluation of epistemic uncertainty about logical quantities can be improved by reducing the replications of uncertain values in the expressions. The same is true for the bounding calculations used in the generalized agent-based modeling described in this paper. The idea is that, whatever universe the simulation is talking about, an event either happens or doesn't happen in that particular universe and both of these events do not occur together, even when we may not be sure which happens. In a mathematical expression with multiple instances of an uncertain logical quantity, if we do not account for this self-dependency (that the event either happens or doesn't) and instead used naïve interval or other bounding methods to evaluate the logical expression, the uncertainty about the logical value would be entered into the evaluation multiple times and the uncertainty of the result could be larger than necessary. In practice, software will be able, in any case, to detect when such uncertainty inflation is possible, and can report a warning to users if appropriate.

Beyond a possible inflation of uncertainty, some analysts incorrectly believe that bounding methods are not useful in practice because their uncertainties often quickly explode to become vacuous. In order to assess whether there is an intrinsic problem with bounding approaches to uncertainty propagation in complex problems, we conducted numerical experiments with randomly constructed mathematical expressions involving addition, subtraction, multiplication, division, minimum and maximum operations on randomly constructed interval ranges centered between zero and one hundred with widths defined by significant digits randomly chosen between one and five. Such intervals vary widely in magnitude and width, as illustrated by these examples:

[95, 105]	[60.5, 61.5]	[43.9995, 44.0005]	[46.35, 46.45]	[6.35, 6.45]
[25, 35]	[18.5, 19.5]	[34.135, 34.145]	[99.15, 99.25]	[54.5, 55.5]

We measured the relative uncertainty of the outcomes, from many sequences of mathematical operations, as the widths of the final intervals divided by their midpoints. We found that most such calculations had uncertainties smaller than about 10%, and more than 95% yielded uncertainties smaller than unity. These results were typical for mathematical expressions with lengths ranging from 10 to 250 sequential operations, and, importantly, do not seem to vary with the number of operations. It is possible for relative uncertainties to be very large (perhaps 10 or larger), but these seems to be rather rare in random simulations. And when the output uncertainties are very large, it is fair to say that the uncertainty is authentically large. The outputs are as tight as possible given their respective inputs. Artifactual inflation of uncertainty which is not an authentic reflection of the uncertainty in the inputs can only occur in interval analysis when there are repeated variables in the mathematical expressions. When we broadened the uncertainties of the inputs still further by allowing there to be possibly zero significant digits, then input intervals such as [50,150] and [–50,50] were entered into the calculation streams. Even when about 17% of the inputs were these very broad intervals, still more than 50% of the resulting uncertainties were less than one, and about 80% were less than 5. About 15% of the final results were actually infinite, resulting from a division by zero (that was not later erased by subsequent minimum and maximum operations that return the result to a finite range). These results also seem to be invariant to the number of sequential operations ranging from 10 to 250. These simulations reveal plainly that uncertainty need not, and probably doesn't, grow to become vacuously wide even when there are many broadly uncertainty inputs involved in a calculation. This conclusion extends immediately to probability bounds analysis as well because intervals give outer bounds on the uncertainties that it would yield.

10. Conclusions

There are twin myths in modeling and simulation that prevent our reaping the benefits from broader application of uncertainty quantification and analysis. The first myth is that the results of traditional calculations are reliable and precise merely because they are expressed with 7 or 14 decimal places. It is unlikely that many actually believe this myth, but they *behave* as though they do. Few modelers construct comprehensive uncertainty or sensitivity analyses that might explore the full implications of the limits of knowledge about model structure and parameter values. Without such analyses, we cannot really say whether or to what extent our model outputs should be trusted. The second myth is that a full and proper accounting of the uncertainty inherent in the simulations and models would blow up to a vacuous conclusion that says nothing because any signal would be lost in the noise of uncertainty. Many people actually seem to believe this pernicious myth although simple simulations described above show it to be as misguided as the first myth.

We have developed alpha-version software implementing the features described above in the R programming language (R Core Team 2015). Preliminary applications on hypothetical model systems characterizing population growth patterns in biological species have corroborated several expectations, including the broad discrepancy between Monte Carlo results and the true range of possible outcomes as illustrated in Figure 1. We expect to encounter computational challenges from circular calculations (one result depends on other results that depend on the first), and repeated variable problems which may arise when the interaction rules in our agent-based models become more complex. Special strategies may be needed to tackle these computational challenges.

Acknowledgements

We thank Nick Friedenberg who does not for a second believe that it is possible to project epistemic uncertainty through agent-based models in the way we have described. We thank him as well as Robert Garrett, Jon Helton, and Lev Ginzburg for helpful discussions. This paper was developed under funding from a Small Business Technology Transfer grant (award number HQ0147323251) from the Missile Defense Agency to Applied Biomathematics, and under a Cooperative Research and Development Agreement (CRADA No. LA13C10694) between Los Alamos National Laboratory and Applied Biomathematics.

References

Alderson, D. L., G. G. Brown, W. M. Carlyle and L. A. Cox, Jr. Sometimes there is no “most vital” arc: assessing and improving the operational resilience of systems. Society for Risk Analysis, San Francisco, CA, 2012.

Atwood, J. A. *A Monte Carlo Study of the Incidence of Probability Underestimation Versus Probability Overestimation Errors With Alternative Safety First Models (Risk)*. ETD collection for University of Nebraska–Lincoln. AAI8602105, 1985. http://digitalcommons.unl.edu/dissertations/AAI8602105

Basu, S., and A. DasGupta. Robust Bayesian analysis with distribution bands. *Statistics and Decisions* 13:333–349, 1995.

Beer, M., S. Ferson, and V. Kreinovich. Imprecise probabilities in engineering analyses. *Mechanical Systems and Signal Processing* 37:4–29, 2013.

Berleant, D. Automaticaly verified reasoning with both intervals and probability density functions. *Interval Computations* 1993 (2):48–70, 1993.

Berleant, D. and C. Goodman-Strauss. Bounding the results of arithmetic operations on random variables of unknown dependency using intervals. *Reliable Computing* 4:147–165, 1998.

Berleant, D., and J. Zhang. Using Pearson correlation to improve envelopes around distributions of functions. *Reliable Computing* 10:139–161, 2004. http://class.ee.iastate.edu/berleant/home/me/cv/papers/correlationpaper.pdf

Berleant, D., G. Anderson and C. Goodman-Strauss. Arithmetic on bounded families of distributions: a DEnv algorithm tutorial. Pages 183–210 in *Knowledge Processing with Interval and Soft Computing*, edited by C. Hu, R.B. Kearfott, A. de Korvin and V. Kreinovich, Springer (ISBN 978-1-84800-325-5), 2008.

Ben-Haim, Y. *Information-gap Theory: Decisions Under Severe Uncertainty*. Academic Press, 2001.

Berger, J. O. An overview of robust Bayesian analysis. *Test* 3:5-124, 1994.

Bobashev, G. V. and R. J. Morris. Uncertainty and Inference in Agent-Based Models. *Second Int. Conference on Advances in System Simulation (SIMUL2010)*, 22–27 August 2010, Nice, France, pages 67–71. IEEE Xplore Digital Library, 2010.

Cossette, H., M. Denuit and É. Marceau. Distributional bounds for functions of dependent risks. Discussion paper 0128, Institut de Statistique, Université Catholique de Louvain, 2001.

Enszer, J. A., Y. Lin, S. Ferson, G. F. Corliss and M. A. Stadtherr. Probability bounds analysis for nonlinear dynamic process models. *AIChE [American Institute of Chemical Engineers] Journal* 57:404–422, 2011.

Faanes, A. and S. Skogestad. Feedforward control under the presence of uncertainty. *European Journal of Control* 10:30–46, 2004.

Ferson, S. *RAMAS Risk Calc 4.0 Software: Risk Assessment with Uncertain Numbers*. Lewis Publishers, Boca Raton, Florida, 2002.

Ferson, S. and L. R. Ginzburg. Different methods are needed to propagate ignorance and variability. *Reliability Engineering and Systems Safety* 54:133–144, 1996.

Ferson, S. and J. Siegrist. Verified computation with probabilities. Pages 95–122 in *Uncertainty Quantification in Scientific Computing*, edited by A. Dienstfrey and R.F. Boisvert, Springer, New York, 2012.

Ferson, S., V. Kreinovich, L. Ginzburg, K. Sentz and D. S. Myers. *Constructing Probability Boxes and Dempster–Shafer Structures*. Sandia National Laboratories, SAND2002-4015, Albuquerque, New Mexico, 2003. http://www.ramas.com/ unabridged.zip

Ferson, S., R. Nelsen, J. Hajagos, D. Berleant, J. Zhang, W. T. Tucker, L. Ginzburg and W. L. Oberkampf. Dependence in Probabilistic Modeling, Dempster–Shafer Theory, and Probability Bounds Analysis. Sandia National Laboratories, SAND2004-3072, Albuquerque, New Mexico, 2004. http://www.ramas.com/depend.pdf

Ferson, S., V. Kreinovich, J. Hajagos, W.L. Oberkampf and L. Ginzburg. *Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty*. SAND2007-0939, Sandia National Laboratories, Albuquerque, New Mexico, 2007. http://www.ramas.com/intstats.pdf

Ferson, S., M. Balch, K. Sentz, and J. Siegrist. Computing with confidence. *Proceedings of the 8th International Symposium on Imprecise Probability: Theories and Applications*, edited by F. Cozman, T. Denœux, S. Destercke and T. Seidenfeld. SIPTA, Compiègne, France, 2013.

Fetz, T. and M. Oberguggenberger. Propagation of uncertainty through multivariate functions in the framework of sets of probability measures. *Reliability Engineering and System Safety* 85:73–88, 2004.

Fonoberova, M., V. A. Fonoberov and I. Mezić. Global sensitivity/uncertainty analysis for agent-based models. *Reliability Engineering and System Safety* 118:8–17, 2013.

Frank, M. J., R. B. Nelsen and B. Schweizer. Best-possible bounds for the distribution of a sum—a problem of Kolmogorov. *Probability Theory and Related Fields* 74:199–211, 1987.

Fréchet, M. Généralisations du théorème des probabilités totales. *Fundamenta Mathematica* 25:379–387, 1935.

Gardner, M. Mathematical games: the fantastic combinations of John Conway's new solitaire game "life". *Scientific American* 223:120–123, 1970.

Goovaerts, M. J., J. Dhaene and A. De Schepper. Stochastic upper bounds for present value functions. *Journal of Risk and Insurance* 67:1–14, 2000.

Hailperin, T. *Boole's Logic and Probability*. North-Holland, Amsterdam, 1986.

Hamlin, T. L., E. Thigpen, J. Kahn and Y. Lo. Shuttle risk progression: use of the shuttle probabilistic risk assessment (PRA) to show reliability growth. American Institute of Aeronautics and Astronautics, 2011. http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa. gov/20110004917_2011004008.pdf

Hancock, P., N. Simpson and E. Rhee. Modeling uncertainty through agent-based participatory simulation: implication to businesses in China. *International Journal of China Marketing* 1, 2010.

Harp, D. R. and V. V. Vesselinov. An agent-based approach to global uncertainty and sensitivity analysis. *Computers & Geosciences* doi:10.1016/j.cageo.2011.06.025, 2011.

Hartung, J.A., H. W. Ho, P. D. Rutherford, E. U. Vaughan. The statistical correlation model for common cause failures. Pages 337-352 in *Risk Analysis in the Private Sector, Advances in Risk Analysis* Volume 220, edited by C. Whipple and V.T. Covello. Springer-Verlag, 1985.

Hassan, S., L. Garmendia and J. Pavón. Introducing uncertainty into social simulation: using fuzzy logic for agent-based modeling. *International Journal of Reasoning-based Intelligent Systems* 2:118–124, 2010.

Henrion, M. and B. Fischhoff. Assessing uncertainty in physical constants. *American Journal of Physics* 54:791–798, 1986.

Hickman, J. W., et al. *PRA Procedures Guide: A Guide to the Performance of Probabilistic Risk Assessments for Nuclear Power Plants*, in two volumes. NUREGCR-2300-V1 and -V2, National Technical Information Service, Washington, DC, 1983.

Hokstad, P. and M. Rausand. Common cause failure modeling: status and trends. Pages 621-640 in *Handbook of Performability Engineering*, edited by K.B. Misra. Springer, London, 2008.

Ilachinski, A. *Artificial War: Multiagent-Based Simulation of Combat*. World Scientific, Singapore, 2004.

Kaas, R., J. Dhaene and M. J. Goovaerts. Upper and lower bounds for sums of random variables. *Insurance:Mathematics and Economics* 27:151–168, 2000.

Kaufmann, A., and M. M. Gupta. *Introduction to Fuzzy Arithmetic: Theory and Applications*. Van Nostrand Reinhold, 1985.

Kharin, V. V., Q. Teng, F. W. Zwiers, G. J. Boer, J. Derome and J. S. Fontecilla. Skill assessment of seasonal hindcasts from the Canadian Historical Forecast Project. *Atmosphere-Ocean* 47:204–223, 2009. See also http://weather.gc.ca/saisons/info_pc_e.html

Li, W. and J. M. Hyman. Computer arithmetic for probability distribution variables. *Reliability Engineering and System Safety* 85:191–210, 2004.

Linkov, I. and D. Burmistrov. Model uncertainty and choices made by modelers: lessons learned from the International Atomic Energy Agency model intercomparisons. *Risk Analysis* 23:1297–1308, 2003.

Lochbaum, D. *The NRC and Nuclear Power Plant Safety in 2010: A Brighter Spotlight Needed*. Union of Concerned Scientists, Cambridge, Massachusetts, 2011. http://www.ucsusa.org/assets/documents/nuclear_power/nrc-2010-full-report.pdf

Magliocca, N. R. and E. C. Ellis. Using pattern-oriented modeling (POM) to cope with uncertainty in multi-scale agent-based models of land change. *Transactions in GIS* [DOI: 10.1111/tgis.12012], 2013. http://onlinelibrary.wiley.com/doi/10.1111/tgis. 12012/pdf

Manes, E. G. A class of fuzzy theories. *Journal of Mathematical Analysis and Its Applications* 85:409–451, 1982.

Miranda, E., I. Montes, R. Pelessoni and P. Vicig. Bivariate p-boxes. *Proceedings of the 8th International Symposium on Imprecise Probability: Theories and Applications,* edited by F. Cozman, T. Denœux, S. Destercke and T. Seidenfeld. SIPTA, Compiègne, France, 2013.

Moore R. E. *Interval Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey, 1966.

Morgan, M. G., and M. Henrion. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, Cambridge, United Kingdom, 1990.

Morss, R. E. and E. Wahl. An ethical analysis of hydrometeorological prediction and decision making: The case of the 1997 Red River flood. *Environmental Hazards* 7:342–352, 2007.

Nelsen, R. B. Properties of a one-parameter family of bivariate distributions with specified marginals. *Communications in Statistics (Theory and Methods)* A15:3277–3285, 1986.

Nelsen, R. B. Discrete bivariate distributions with given marginals and correlation. *Communications in Statistics (Simulation and Computation)* B16:199–208, 1987.

Nguyen, H. T., V. Kreinovich, B. Wu and G. Xiang. *Computing Statistics under Interval and Fuzzy Uncertainty: Applications to Computer Science and Engineering*. Studies in Computational Intelligence, volume 393. Springer-Verlag, Berlin, 2011.

NRC [National Research Council Committee on Estimating and Communicating Uncertainty in Weather and Climate Forecasts]. *Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts*, http://www.nap.edu/openbook.php?record_id=11699, The National Academies Press, Washington, DC, 2006.

NWS [National Weather Service]. Service assessment and hydraulic analysis: Red River of the North 1997 floods, 1998. http://www.nws.noaa.gov/oh/Dis_Svy/RedR_Apr97/TOC.htm

Oberkampf, W. L. and C. J. Roy. *Verification and Validation in Scientific Computing*. Cambridge University Press, New York, 2010.

R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. https://www.R-project.org/

RAWG [Risk-Assessment Working Group]. *Risk-Assessment Methodologies for Use in the Electric Utility Industry*, Critical Infrastructure Protection Committee, North American Electric Reliability Council [North American Electric Reliability Corporation], Atlanta, Georgia, 2005. http://www.ercot.com/content/meetings/cipag/keydocs/2007/0115/Risk_Assessment_ Methodologies_9_9_2005.pdf

Savage, S. L. *The Flaw of Averages: Why We Underestimate Risk in the Face of Uncertainty*. Wiley, 2012.

Sentz, K. and S. Ferson. *Combination of Evidence in Dempster–Shafer Theory*. Sandia National Laboratories, SAND 2002-0835, Albuquerque, New Mexico, 2002.

Sentz, K. and S. Ferson. Probabilistic bounding analysis in the quantification of margins and uncertainties. *Reliability and Engineering System Safety* 96:1126–1136, 2011.

Shoham, Y. and K. Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2009.

USCPSOTF [U.S.-Canada Power System Outage Task Force]. *Final Report on the Implementation of the Task Force Recommendations*. Natural Resources Canada and U.S. Department of Energy., 2006. http://www.nrcan.gc.ca/energy/sources/ electricity/1378, http://energy.gov/oe/downloads/us-canada-power-system-outage-task-force-final-report-implementation-task-force

Vesely, W. E., F. F. Goldberg, N. H. Roberts and D. F. Haasl. *Fault Tree Handbook*. Nuclear Regulatory Commission, Washington, DC, 1981.

Vose, D. *Risk Analysis: A Quantitative Guide*. John Wiley and Sons, West Sussex, England, 2008.

Walley, P. *Statistical Reasoning with Imprecise Probabilities*. Chapman & Hall, London, 1991.

Weiss, G. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. MIT Press, 2001.

Williamson, R. C. and T. Downs. Probabilistic arithmetic I: Numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning* 4:89–158, 1990.

Wooldridge, M. *An Introduction to MultiAgent Systems*. John Wiley and Sons, West Sussex, England, 2009.

Wu, W., E. Ekaette and B. H. Far. Uncertainty management framework for multi-agent system. Pages 122–131 in *Proceedings of the ATS*, 2003. http://www.enel.ucalgary.ca/People/far/pub/papers/2003/ ATS2003-06.pdf

Yager, R. R. Arithmetic and other operations on Dempster–Shafer structures. *International Journal of Man-machine Studies,* 25:357–366, 1986.

Youden, W. J. Enduring values. *Technometrics*, 14:1–11, 1972.

Zhang, H., R. L. Mullen and R. L. Muhanna. Safety structural analysis with probability-boxes. *International Journal of Reliability and Safety*, 6:110–129, 2012.

Zellner, M. L. Embracing complexity and uncertainty: the potential of agent-based modeling for environmental planning and policy. *Planning Theory and Practice* , 9:437–457, 2008.