Terrisk: Battling Uncertainty in Bioterrorism Models

Ross Gore, Barry Ezell
Virginia Modeling, Analysis and Simulation Center
Suffolk, VA
rgore@odu.edu, bezell@odu.edu

ABSTRACT

Modeling under uncertainty has been of paramount importance, as quantitative methods of risk analysis have been developed to evaluate, prevent and respond to bioterror threats. Our simulation platform, Terrisk, improves risk assessment capabilities by providing modelers with novel means to: (1) capture incomplete and conflicting information about bioterror threats and (2) propagate that captured uncertainty to simulation outputs. Terrisk achieves these capabilities by supporting computation through three mathematical frameworks: probability theory, probability boxes, and Dempster-Shafer theory of evidence. The inclusion of the latter two frameworks (probability boxes and Dempster-Shafer theory) is novel. They offer alternatives to traditional probability theory, which enables incomplete and conflicting information about model inputs to be captured and propagated to outputs. This capability of Terrisk facilitates decision-making because it provides more insight into possible model outcomes by more accurately reflecting what is known about model inputs.

ABOUT THE AUTHORS

Ross Gore holds a Doctorate of Philosophy and a Master's degree in Computer Science from the University of Virginia and a Bachelor's degree in Computer Science from the University of Richmond. Dr. Gore has ten years of research experience in problems that lie at the intersection of computer science and modeling and simulation. His work has yielded authorships on more than 20 conference and journal publications and has been recognized by the ARCS (Achievement Rewards for College Scientists) Society as an impactful and novel research avenue.

Barry Ezell is a research associate professor and Chief Scientist for Virginia Modeling Analysis and Simulation Center, Old Dominion University. Barry has 26 years of experience in decision and risk analysis in the U.S. Department of Defense (DOD) and U.S. Department of Homeland Security (DHS), with senior risk analysis expertise in terrorism risk, critical infrastructure, and industrial control systems. His recent projects include panelist to support review for Quadrennial Homeland Security Review-Bio; panelist to advise DHS S&T on future of Bioterrorism Risk Assessment (BTRA); and principal investigator in the development of a decision model to inform Commonwealth of Virginia DHS Grant allocation. Barry is currently the president of the Security Analysis and Risk Management Association.

Terrisk: Battling Uncertainty in Bioterrorism Models

Ross Gore, Barry Ezell
Virginia Modeling, Analysis and Simulation Center
Suffolk, VA
rgore@odu.edu, bezell@email.com

INTRODUCTION

Signed in 2004, Homeland Security Presidential Directive (HSPD) 10 focused on improving the nation's ability to prevent, prepare for, respond to, and recover from terrorism attacks that employed biological agents as their means. An important component of HSPD-10 was the President's requirement for DHS to develop "periodic assessments of the evolving biological weapons threat," explaining that "the United States requires a continuous, formal process for conducting routine capabilities assessments to guide prioritization of ... on-going investments in biodefense-related research, development, planning, and preparedness" (Bush, 2004). Modeling and simulation has gained prominence as the proper tool to study the threat of bioterrorism since it is infeasible and impractical to study such systems directly.

When modeling bioterror representing important quantities such as: (1) adversary intent, (2) adversary capabilities, (3) target vulnerabilities, and (4) countermeasure costs is difficult since historical data is sparse or nonexistent. Frequently expert opinions are used. While it is generally believed that combining information from multiple sources will generate better judgments. It has been shown that aggregating conflicting information from experts with polarizing opinions can result in non-representative quantities (Clemen and Winkler, 2007). Furthermore, studies show that experts tend to give overconfident judgments that may be politicized. For example the 2002 NIE on weapons of mass destruction in Iraq was perceived as "hawkish," while the 2007 NIE on nuclear weapons in Iran was perceived as "dovish" (Rovner, 2011).

Ideally, modelers and decision makers would have a platform to capture the uncertainty associated with probabilistic judgments of quantities from multiple sources and propagate that uncertainty to reveal a range of possible outcomes. The modeling and simulation platform proposed in this paper, Terrisk, represents a step toward achieving this goal.

Terrisk builds on the platform Modelica by incorporating novel methods for quantifying uncertainty formally and robustly through the use of imprecise probabilities. Imprecise probability reflects any alternative framework, which measures uncertainty without the crisp numerical probabilities used in traditional probability theory. To our knowledge no one has attempted to use imprecise probabilities to characterize terrorist threats since Los Alamos National Labs (LANL) in 2005 (Singpurwalla and Booker, 2004). While, the LANL approach used one mathematical framework, possibility theory, to characterize expert elicitations of uncertainty, Terrisk will employ three: (1) traditional probability theory, (2) upper and lower probability bounds and (3) Dempster-Schafer theory (a generalization of possibility theory). This is significant; upper and lower probability bounds and Dempster-Schafer theory provide means to address the following weaknesses of traditional probability theory (Sentz and Ferson, 2002):

- Probability theory requires a precise value to be assigned to each element in the set of possible outcomes. However, in the face of terrorist threats it may not be possible to assign exact values or even assign reasonable approximations when little information is available.
- Probability theory imposes Laplace's principle of insufficient reason when no information is available. When n mutually exclusive possible outcomes are indistinguishable except for their names, they must each be assigned a probability of $1/n$.
- Conflicting evidence cannot be represented in traditional probability theory. By assigning probabilities to individual elements, we cannot express incompatibility between conflicting sources of information or cooperative effect between sources of information in agreement.

The result of addressing these weaknesses in a flexible risk assessment platform will: (1) enable modelers to more accurately express their uncertainty about domain knowledge and as a result (2) provide actionable and better-informed analysis to decision-makers. Furthermore, Terrisk provides the ability to calibrate models employing any of the three uncertainty frameworks with a Bayesian inference engine. This is significant. When addressing concerns related to terrorist threats, modeling under uncertainty implies the absence of perfect information (Enders, Sandler and Gaibulloev, 2011). However, often some partial information exists in the form of observations on the model's expected behavior (Relyea and Seifert, 2005). Modelers expect the capability to make the best possible use of the information available to them. Bayesian inference is the appropriate mechanism to meet this expectation.

The remainder of the paper is structured as follows. Next we review related work. Then we present our design and implementation of Terrisk. Following that two case studies from the unclassified domain are presented to demonstrate Terrisk's capabilities. Finally, we summarize our results and provide direction for future work.

RELATED WORK

Several different mathematical systems can be used to perform uncertainty analysis. Terrisk focuses on probability theory, probability boxes, and the Dempster-Shafer theory of evidence. Probability theory is the most traditional representation of uncertainty and the one most familiar to non-mathematicians. The use of probability theory attempts to provide a quantitative analysis to answer the following three questions: (1) what can go wrong, (2) how likely is it that will happen, and (3) if it does happen, what are the consequences (Kaplan and Garrick, 1981)? Probability used as a representation of subjective belief is common in quantitative risk analysis. Safety assessments must deal with rare events and thus it is difficult to assess the relative frequencies of these events (Apostolakis, 1990). The Bayesian approach to uncertainty analysis is to specify a coherent probability measure as the current state of available knowledge, and use Bayes' theorem to adjust probabilities as new evidence is unveiled.

Imprecise probability is a generic term for any mathematical model, which measures chance or uncertainty without crisp numerical probabilities. Two types of imprecise probability, probability boxes and Dempster-Shafer belief structures, offer a more flexible representation of uncertainty over the crisp probabilistic approach. These two mathematical techniques provide an approach to several of the most serious problems that can arise during risk analysis, including: (1) imprecisely specified distributions; (2) poorly known or even unknown dependencies; (3) non-negligible measurement uncertainty; (4) non-detects or other censoring in measurements; (4) small sample size; (5) inconsistency in the quality of input data; (6) model uncertainty; and (7) non-stationary (non-constant distributions) (Kreinovich et al, 2002).

Probability Boxes

Probability boxes (or p-boxes) define upper and lower boundaries for the probabilities of a set of events (Walley, 1991). These upper and lower boundaries (represented by $\overline{P}(X)$ and $\underline{P}(X)$, respectively) can provide additional information that is not available by means of traditional probability theory. The following example illustrates the potential benefits of probability boxes. Imagine you have found a magical lamp with a genie inside. The genie offers you to play one of two games. In the first game you are given a precision casino six-sided die and ask to roll the die. If you roll a one then the genie will give you $10,000. In the second game the genie has a stack of six cards that consists of the four aces and two jokers. He presents the stack of cards (face down) and asks you to pick the top card. If the card you select is the ace of hearts, you will win $10,000. The genie is offering you the choice of playing either one of the two games. Your knowledge of probability informs you that in either case the probability of winning, $P(win)$, is 1/6.

However you feel that the games are not identical to each other. By applying probability boxes, we can deduce that $\underline{P}(win) = 0$ and $\overline{P}(win) = 1$ in the case of the card game. This is because the genie can decide to stack the deck either in your favor or against your favor. Therefore before you measure the event, you can deduce that the genie might know whether or not you will select the ace of hearts. On the other hand in the case of the dice roll game, $\underline{P}(win) = \overline{P}(win) = 1/6$, since there is no additional information available that can be used to categorize the potential states of the system.

A gambler's interpretation of $\underline{P}(X)$ is that it represents the highest price he is willing to pay in order to receive one dollar if X occurs, or receive nothing if X does not occur. Similarly, $\overline{P}(X)$ represents the infimum selling price of an event, which is the lowest price that he is willing to receive in order to sell one dollar if X occurs. Probability boxes are the upper and lower distribution functions ($\underline{F}$ and $\overline{F}$) of an event X where $\underline{F}(x) = \underline{P}(X \leq x)$ and $\overline{F}(x) = \overline{P}(X \leq x)$. Upper and lower distribution functions allow an analyst to make no assumptions about the shape of the true probability distribution function. A series of coherency axioms ensure that $\underline{F}(x) \leq F(x) \leq \overline{F}(x)$ for all real numbers x. Probability boxes enable some separation of epistemic uncertainty and aleatory uncertainty (Ferson and Ginzburg, 1996; Ferson and Hajagos, 2004). Under classical probability theory, the principle of indifference dictates one should select a uniform distribution when presented with a lack of information concerning the shape of that distribution. Traditional probabilistic analysis ignores the epistemic uncertainty of the model and thus can result in misleading risk assessment calculations.

Dempster-Schafer Theory of Evidence

In the Dempster-Shafer theory of evidence, the concept of imprecise probabilities is extended to account for both nonspecificity and discord of available evidence (Shafer, 1976; Yager, Fedrizzi and Kacprzyk, 1994). In comparison, probability boxes account for nonspecificity by propagating lower and upper bounds without specifying the shape of the distribution. However, probability boxes require that all the available evidence concludes in one non-overlapping interval. Dempster-Shafer theory allows a decision-maker to reason about several candidate probability intervals for a random process, even when the candidate intervals conflict with one other. Dempster-Shafer theory is formulated in terms of a function known as the basic probability assignment. If Ω is the set of all possible outcomes under consideration and 2^Ω is the power set of Ω (the set of all subsets of Ω), then a basic probability assignment is defined as $m(A): 2^\Omega \to [0,1]$ such that: $m(\emptyset) = 0$ and $\sum_{A \in 2^\Omega} m(A) = 1$.

Evidence theory is a generalization of traditional probability theory where probabilities are assigned to sets of events instead of individual events. A degree of belief, *Bel*, given to the set *A,* is defined as the sum of all masses that support *A:* $Bel(A) = \sum_{B \subset A} m(B)$. The degree of plausibility, *Pl*, quantifies the total amount of belief that might support *A:* $Pl(A) = Bel(\Omega) - Bel(A) = \sum_{B \cap A \neq \emptyset} m(B)$ (Smets, 1994). The belief and plausibility functions can be viewed as lower and upper bounds of probability in the sense that $Bel(A) \leq P(A) \leq Pl(A)$. It is often the case that some subsets of Ω will have basic probability assignments of zero. Those sets with nonzero assignments are called the focal elements of *m*. The set of all focal elements is known as a body of evidence.

Computing belief and plausibility measures have been shown to be #P-complete problems (Orponen, 1990). The complexity class #P contains the counting problems which are associated with decision problems that can be solved in polynomial time on a non-deterministic Turing machine (NP problems). It is believed that there does not exist a polynomial-time algorithm for solving #P-complete problems, since this would imply P = NP. As with many suprapolynomial computational problems, a research emphasis has been placed on devising polynomial-time approximation algorithms to computing belief and plausibility measures. Joslyn and Kreinovich describe a Monte Carlo method for approximating belief and plausibility that can produce good estimates when the conflict between evidence is not very high (Joslyn and Kreinovich, 2005). Dempster-Shafer theory provides a method of reasoning under imprecisely defined sources of information. The belief function is a measure of the extent to which available evidence implies that something is true. And plausibility is a measure of the extent to which evidence implies that something might be true.

Bayesian Inference

Bayesian inference is a technique for updating a priori probability distributions based on the observation of data that are relevant to the variables of interest (Lee, 2012). Consider a generic application of Bayesian inference by assuming a model $\overline{y} = f(\overline{x}, \overline{\Theta})$, where $\overline{x}$ represents a vector of input quantities to the model, $\overline{\Theta}$ represents a vector of scalar parameters to the model, and $\overline{y}$ represents a vector of observable output quantities. The *a priori* beliefs $\overline{\Theta}$ are expressed using a probability distribution function $p(\overline{\Theta})$. On the basis of independent observations $X = (X_1, X_2, \ldots, X_n)$, the probability distribution is then updated according to Bayes Law: $p(\overline{\Theta}|\overline{X}) = \frac{p(\overline{\Theta})p(\overline{X}|\overline{\Theta})}{p(\overline{X})} = \frac{p(\overline{\Theta})p(\overline{X}|\overline{\Theta})}{\int p(\overline{\Theta})p(\overline{X}|\overline{\Theta})\,d\overline{\Theta}}$. The posterior distribution can be approximated numerically using Monte Carlo integration

techniques. Traditional Monte Carlo algorithms perform a series of independent samples from a probability distribution to approximate the desired integral. In order to numerically calculate high-dimensional integrals on large parameter spaces, $\Theta = (\Theta_1, \Theta_2, \dots, \Theta_m)$, a Markov chain Monte Carlo (MCMC) integration technique can be used. MCMC integration techniques are often more efficient than traditional Monte Carlo techniques in high-dimensional parameter spaces. For example, the Gibbs sampler is a widely used MCMC method for Bayesian inference (Gill, 2007). In the Gibbs sampler each of the parameters is expressed as a probability conditioned on all of the other parameters. Next, a Markov chain is produced that cycles through all of the conditional probabilities. To perform the Gibbs sampler, start with estimates for $\overline{\Theta}^{[0]} = (\theta_1^{[0]}, \theta_2^{[0]}, \dots, \theta_m^{[0]})$. Next, begin a cycle by drawing values for each parameter Θ_k according to the distribution $\Theta_k \sim p(\theta_k | \theta_{-k}, X)$, where θ_{-k} indicates the set of parameters without the θ_k coefficient, and X is the set of independent observations. Repeat cycling through the parameter distributions until the process converges on a set $\overline{\Theta}^{|n|}$. Any problem that can be expressed in terms of $p(\theta_k | \theta_{-k}, X)$ can be solved using the Gibbs sampler.

TERRISK

Terrisk is designed as a series of extensions to the Modelica modeling platform which support the representation of imprecise probabilities (Fritzon, 2010). It contains new data types and new operators for these types to manipulate random variables in the three probability frameworks we have previously discussed (probability theory, probability boxes, Dempster-Schafer Theory). The underlying support for these operators are probability density functions (Park, 2005). Terrisk implements probability density functions using a calculus of sampling functions. In order to construct sampling functions, we introduce the sampling expression *S*. *S* consumes a random number from an infinite sequence drawn independently from $U[\mathbf{0.0, 1.0}]$, and *S* evaluates to the random number that has been consumed. Whereas a sampling function is sufficient to represent a probability density function, a pair of sampling functions can be used to represent a probability box, such that $\underline{F}(x) \leq F(x) \leq \overline{F}(x)$ for all real numbers *x*. Dempster-Schafer structures are represented by computing the power set of sampling functions that could represent a probability box.

Once the user creates a model the sampling expressions are generated to reflect the desired uncertainty in the appropriate framework. Recall, Terrisk is implemented as a superset of Modelica. As a result the architecture includes a Terrisk compiler to translate the Terrisk model (including generated sampling expressions) into an executable Modelica model. The output of the Terrisk compiler is a Modelica model, along with special instructions for the Terrisk execution framework. The Terrisk architecture, shown in Figure 1, consists of the Terrisk Compiler, the Modelica compiler, the Monte Carlo engine, and the Bayesian engine. We review each of these next.
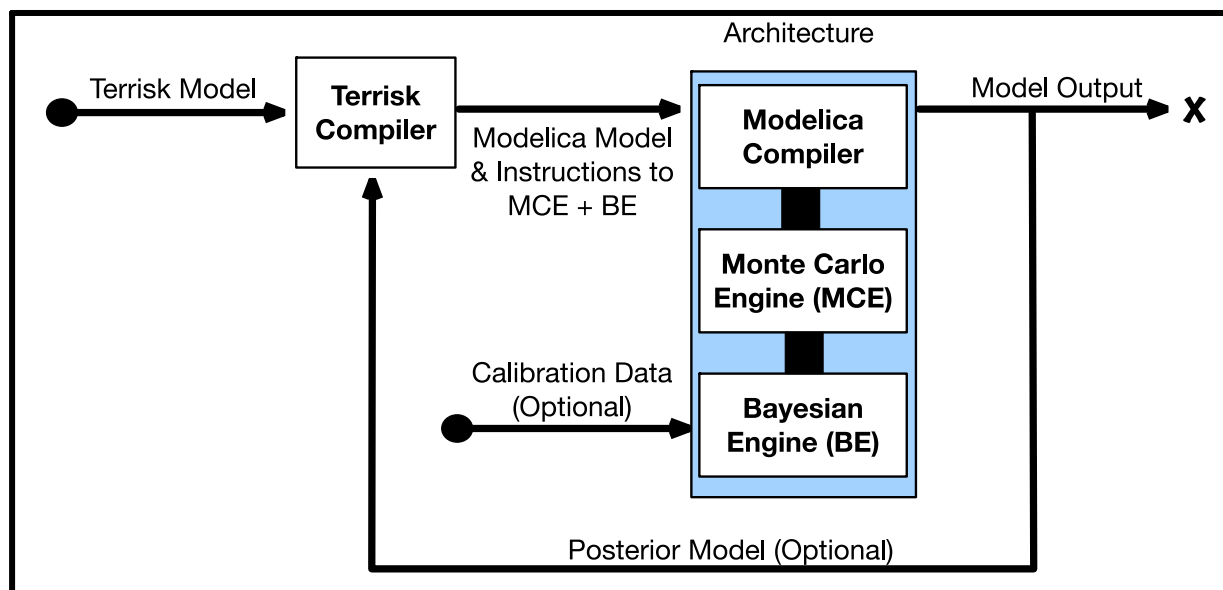

Figure 1. Terrisk Architecture

Terrisk Compiler and Modelica Compiler

It is the function of the Terrisk compiler to translate Terrisk-specific constructs, and pass along the remaining Modelica constructs downstream to the Modelica compiler. The Terrisk compiler extends the grammar specification that is available from the OpenModelica project (Fritzon, 2010), which uses ANTLR (Parr, 2007), a predicated LL(k) parser generator, to construct the Modelica parser. For the Modelica Compiler, we use the open source OpenModelica compiler designed by Programming Environment Laboratory of Peter Fritzson (Fritzon, 2010). The OpenModelica compiler translates the Modelica code into C code that is then compiled and executed to yield results.

Monte Carlo and Bayesian Engine

The remaining units of the Terrisk execution framework, the Monte Carlo engine and the Bayesian engine, are responsible for directing execution flow in order to analyze the uncertainty information that has been encoded in the original Terrisk program. The Monte Carlo engine and the Bayesian engine receive instructions from the Terrisk compiler, and use these directions to control the execution of the compiled Modelica program. The Monte Carlo (MC) engine is responsible for: (1) setting up multiple iterations of the model and (2) managing the sampling expression S and the sampling constructs that are created for imprecise probability structures. The Bayesian engine is responsible for applying Bayes' Law to calibrate Terrisk models. When the Bayesian engine is engaged, the output of the Terrisk execution framework does not consist of model output. Instead the Bayesian engine output is a calibrated Terrisk program that can be subsequently compiled and executed at a later time. The Bayesian engine has only one iteration by default, as the output of the engine is a calibrated program that has turned off Bayesian inference. This inference could be manually turned back on when additional calibration data becomes available to achieve multiple iterations.

CASE STUDY

Two case studies elucidate the capabilities Terrisk will provide in practice. First we review an example that demonstrates Terrisk's ability to propagate uncertainty through a system using a combination of the three mathematical frameworks. Then we demonstrate the value in using Terrisk's calibration engine to compensate for this lack of information in the data by taking prior information into account. Both of these case studies are in the unclassified domain.

Specification and Propagation of Uncertainty: Mass-Spring-Damper System

For our case study we chose to work with one of the challenge problems presented in a workshop on uncertainty representation at Sandia National Laboratories (Oberkampf et al, 2004). The workshop created seven time-independent models and discussed various types of generalized uncertainty analyses of these models. The seventh model was inspired by a dynamic system known as the "Mass-spring-damper" (MSD) system shown in Figure 2. In the MSD system, a mass is attached to a cart on wheels, which is attached to a wall with both a spring and a linear damper. An external force is applied to the cart.
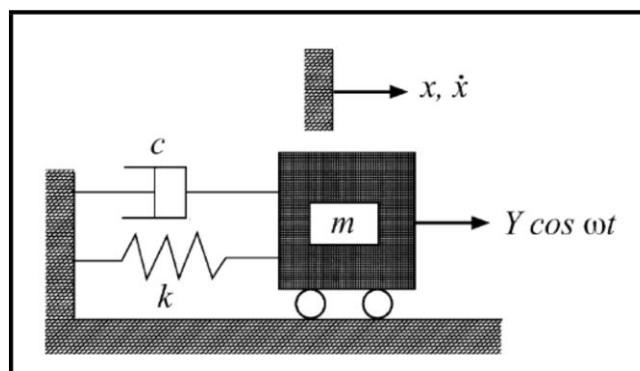

Figure 2: Mass-Spring-Damper (MSD) System.

The challenge problem asks for the position of the mass over time. Terrisk calculates the position distribution functions of the mass using numerical integration techniques but we also derived these quantities analytically in order to verify the numerical integration scheme.

The parameters of the MSD system are represented as a combination of aleatory uncertainty and epistemic uncertainty. The parameters $\{m, k, \omega\}$ are assumed to be of an aleatory character and they are represented by traditional probability distributions. k and ω are represented by triangular distributions, and the input parameters to these triangular distributions are assumed to be of an epistemic nature. The parameters $\{k_{min}, k_{mod}, k_{max}, c, Y, \omega_{min}, \omega_{mod}, \omega_{max}\}$ are represented by Dempster-Shafer belief structures using continuous intervals, where $\alpha_{min}, \alpha_{mod}, \alpha_{max}$ are the three input parameters to a triangular distribution for some parameter α. The parameters k and ω are represented using a two-tier system of uncertainty, using belief-structures as input parameters to the probability distributions of these parameters.

Terrisk calculates the position of the mass, $x(t)$, at time slices of $t = 0.05$ seconds over an interval of 30 seconds. The analysis begins at time t= 0 when the system is at rest. Figure 3 show the movement of the mass over time $t = 3$ and $t = 6$ seconds, respectively. The blue line represents the lower bound of the position and the red line represents the upper bound of the position.
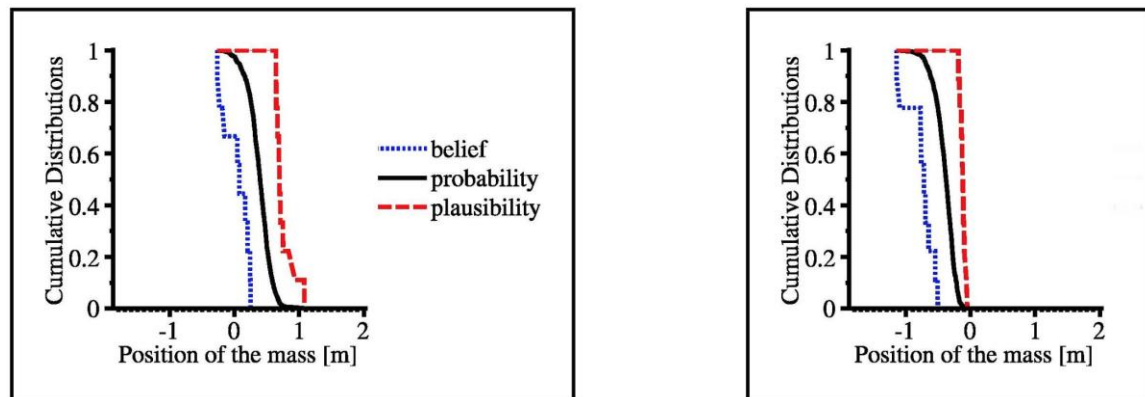

Figure 3: Position of mass at time t = 3 (left) and t = 6 (right).

An animated short film showing the dynamic behavior of the probability distribution functions is available online. In this film, we can observe the initial oscillation of the probability measures as the object is placed into motion, followed by a steady-state behavior of the mass following the initial transient period. Combined with Figure 3 these demonstrate that Terrisk is capable of propagating plausability and belief measures through a dynamic simulation over time using the sampling functions for sufficiently complex models.

Model Calibration with a Lack of Data: Annual Crop Growth

Next, we briefly discuss an example of Bayesian calibration within Terrisk. Our case study is the calibration of a corn growth model. The model simulates corn growth during one growing season, with 21 parameters related to weather conditions as input. To construct the model we used a prior distribution for the parameters in the model obtained from an analysis of worldwide weather conditions during 1985 (Metselaar, 1999). This prior distribution may be viewed as a description of the variation of the parameters of different corn crops over the world. In addition to the prior data we also have a harvest measurement of 15 tons with a standard error of 1 ton, for the corn growth of the United States in 1985.

The Bayesian analysis is based on sampling from the posterior distribution of the parameters. A sample of 1000 independent draws is constructed with a MCMC method. Terrisk performs independent draws from the prior parameter distribution and moves them into the posterior by chance with probability proportional to the likelihood.
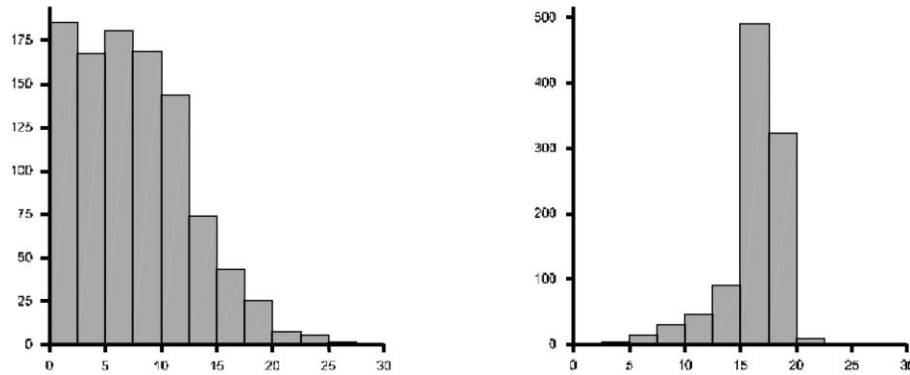
Figure 4: Histogram of: (left) prior predictive sample and (right) posterior predictive sample.

The effect of combining this generic prior information with region-specific data enables the model to make predictions under the weather conditions for a specific region (United States) at a different time (post-1985). In contrast the generic prior information came from 1985 and included weather conditions across the world.

Figure 4 (left) shows a histogram of a size-1000 sample from the prior predictive harvest distribution for 1986 where only the prior information is used (worldwide, 1985). The histogram shows that this prior information is insufficient by itself: the distribution is very vague. The effect of the information about the 1985 United States harvest on the prediction for 1986 is illustrated by Figure 4 (right). The histogram shown in Figure 4 (right) is much sharper, which shows that combining the two kinds of information has a positive effect. Moreover, the mean of the posterior predictive distribution is greater than the prior mean, which might express that corn is thriving in the United States climate. Since the data consist of merely one observation, whereas 21 parameters have to be estimated, these parameters cannot be estimated with a classical analysis, which requires that the prediction of the data to be different for different values of the parameter vector. Terrisk enables us to compensate for this lack of information in the data by taking prior information into account.

CONCLUSION

Modeling bioterror requires estimates of important quantities such as: (1) adversary intent, (2) adversary capabilities, (3) target vulnerabilities, and (4) countermeasure costs. These estimates are full of uncertainty since historical data is sparse or nonexistent. Frequently expert opinions are used. While it is generally believed that combining information from multiple sources will generate better judgments (Clemen and Winkler 2007) simple approaches aggregating conflicting information from experts can result in non-representative judgments. Ideally, modelers and decision makers would have a platform to capture the uncertainty associated probabilistic judgments of quantities from multiple sources and propagate that uncertainty to reveal a range of possible outcomes. The modeling and simulation platform proposed in this paper, Terrisk, represents a step toward achieving this goal. Terrisk improves the risk management process by offering better capabilities to define, evaluate and solve terrorism-related models under the conditions of limited available information. Ultimately, Terrisk will lead to assessments that better inform terrorism related protective measures and policy decisions.

ACKNOWLEDGEMENTS

REFERENCES

Apostolakis, G. (1990). The concept of probability in safety assessments of technological systems. Science, 250(4986), 1359-1364.

Bush, G. W. (2004). Homeland security presidential directive 5. National Security Presidential Directives.

Clemen, R. T., & Winkler, R. L. (2007). Aggregating probability distributions. Advances in Decision Analysis, 154-176.

Enders, W., Sandler, T., & Gaibulloev, K. (2011). Domestic versus transnational terrorism: Data, decomposition, and dynamics. Journal of Peace Research, 48(3), 319-337.

Ferson, S., & Ginzburg, L. R. (1996). Different methods are needed to propagate ignorance and variability. Reliability Engineering & System Safety, 54(2), 133-144.

Ferson, S., & Hajagos, J. G. (2004). Arithmetic with uncertain numbers: rigorous and (often) best possible answers. Reliability Engineering & System Safety, 85(1), 135-152.

Fritzson, P. (2010). Principles of object-oriented modeling and simulation with Modelica 2.1. John Wiley & Sons.

Gill, J. (2007). Bayesian methods: A social and behavioral sciences approach. CRC press.

Joslyn, C., & Kreinovich, V. (2005). Convergence properties of an interval probabilistic approach to system reliability estimation. International journal of general systems, 34(4), 465-482.

Kaplan, S., & Garrick, B. J. (1981). On the quantitative definition of risk. Risk analysis, 1(1), 11-27.

Kreinovich, V., Ginzburg, L., Myers, D. S., & Sentz, K. (2002). Constructing probability boxes and Dempster-Shafer structures (Vol. 835). Sandia National Laboratories.

Lee, P. M. (2012). Bayesian statistics: an introduction. John Wiley & Sons.

Metselaar, K. (1999). Auditing predictive models: a case study in crop growth. Landbouwuniversiteit Wageningen.

Oberkampf, W. L., Helton, J. C., Joslyn, C. A., Wojtkiewicz, S. F., & Ferson, S. (2004). Challenge problems: uncertainty in system response given uncertain parameters. Reliability Engineering & System Safety, 85(1), 11-19.

Orponen, P. (1990). Dempster's rule of combination is#< i> P</i>-complete. Artificial Intelligence, 44(1), 245-253.

Parr, T. (2007). The definitive ANTLR reference: building domain-specific languages. Pragmatic Bookshelf.

Rovner, J. (2011). Fixing the facts: national security and the politics of intelligence. Cornell University Press.

Smets, P. (1994). What is Dempster-Shafer's model. Advances in the Dempster-Shafer theory of evidence, 5-34.

Shafer, G. (1976). A mathematical theory of evidence (Vol. 1). Princeton: Princeton university press.

Singpurwalla, N. D., & Booker, J. M. (2004). Membership functions and probability measures of fuzzy sets. Journal of the American Statistical Association, 99(467), 867-877.

Walley, P. (1991). Statistical reasoning with imprecise probabilities (pp. 632-638). London: Chapman and Hall.

Yager, R., Fedrizzi, M., & Kacprzyk, J. (1994). Advances in the Dempster-Shafer theory of evidence.