Sensitivity Analysis in Conjunction with Evidence Theory Representations of Epistemic Uncertainty

J.C. Helton, J.D. Johnson, W.L. Oberkampf

Sandia National Laboratories, Albuquerque, NM 87185-0779
E-mail: jchelto@sandia.gov, jdjohns@sandia.gov, wloberk@sandia.gov

Abstract: Three applications of sampling-based sensitivity analysis in conjunction with evidence theory representations for epistemic uncertainty in model inputs are described: (i) an initial exploratory analysis to assess model behavior and provide insights for additional analysis, (ii) a stepwise analysis showing the incremental effects of uncertain variables on complementary cumulative belief functions and complementary cumulative plausibility functions, and (iii) a summary analysis showing a spectrum of variance-based sensitivity analysis results that derive from probability spaces that are consistent with the evidence space under consideration.

Keywords: Epistemic uncertainty, evidence theory, sensitivity analysis, uncertainty analysis

1. INTRODUCTION

Uncertainty analysis and sensitivity analysis should be important components of any analysis of a complex system, with (i) uncertainty analysis providing a representation of the uncertainty present in the estimates of analysis outcomes and (ii) sensitivity analysis identifying the contributions of individual analysis inputs to the uncertainty in analysis outcomes[1]. Probability theory provides the mathematical structure traditionally used in the representation of epistemic (i.e., state of knowledge) uncertainty, with the uncertainty in analysis outcomes represented with probability distributions and typically summarized as cumulative distribution functions (CDFs) or complementary cumulative distribution functions (CCDFs) [2-4]. A variety of sensitivity analysis procedures have been developed for use in conjunction with probabilistic representations of uncertainty, including differential analysis [5, 6], the Fourier amplitude sensitivity test (FAST) and related variance decomposition procedures[7-11], regression-based techniques [12, 13], and searches for nonrandom patterns [14].

Although probabilistic representations of uncertainty have been successfully employed in many analyses, such representations have been criticized for inducing an appearance of more refined knowledge with respect to the existing uncertainty than is really present [15, 16]. Much of this criticism derives from the use of uniform distributions to characterize uncertainty in the presence of little or no knowledge with respect to where the appropriate value to use for a parameter is located within a set of possible values. As a result, a number of alternative mathematical structures for the representation of epistemic uncertainty have been proposed, including evidence theory, possibility theory, and fuzzy set theory [17].

Evidence theory provides a promising alternative to probability theory that allows for a fuller representation of the implications of uncertainty than is the case in a probabilistic representation of uncertainty. In particular, evidence theory involves two representations of the uncertainty associated with a set of possible analysis inputs or results: (i) a belief, which provides a measure of the extent to which the available information implies that the true value is contained in the set under consideration, and (ii) a plausibility, which provides a measure of the extent to which the available information implies that the true value might be contained in the set under consideration. One interpretation of the belief and plausibility associated with a set is that (i) the belief is the smallest possible probability for the set that is consistent with all available information and (ii) the plausibility is the largest possible probability for the set that is consistent with all available information. An alternative interpretation is that evidence theory is an internally consistent mathematical structure for the representation of uncertainty without any explicit conceptual link to probability theory. The mathematical operations associated with evidence theory are the same for both interpretations. Just as probability theory uses CDFs and CCDFs to summarize uncertainty, evidence theory uses cumulative belief functions (CBFs), cumulative plausibility functions (CPFs), complementary cumulative belief functions (CCBFs), and complementary cumulative plausibility functions (CCPFs) to summarize uncertainty.

Although evidence theory is beginning to be used in the representation of uncertainty in applied analyses, the authors are unaware of any attempts to develop sensitivity analysis procedures for use in conjunction with evidence theory. Due to the importance of sensitivity analysis in any decision-aiding analysis, the potential usefulness of evidence theory will be enhanced if meaningful and practicable sensitivity analysis procedures are available for use in analyses that employ evidence theory in the representation of uncertainty. As a result, the focus of this presentation is on the development of sensitivity analysis procedures for use in conjunction with evidence theory representations of uncertainty.

After a brief overview of evidence theory (Sect. 2), the following topics are considered: (i) exploratory sensitivity analysis (Sect. 3), (ii) use of sensitivity analysis results in the stepwise construction of CCBFs and CCPFs (Sect. 4), (iii) analysis of evidence theory representations of uncertainty (Sect. 5), and (iv) concluding summary (Sect. 6).

2. EVIDENCE THEORY

Evidence theory is based on the specification of a triple $(\mathcal{S}, \mathbb{S}, m)$, where (i) $\mathcal{S}$ is the set that contains everything that could occur in the particular universe under consideration, (ii) $\mathbb{S}$ is a countable collection of subsets of $\mathcal{S}$, and (iii) m is a function defined on subsets of $\mathcal{S}$ such that $m(\mathcal{E}) > 0$ if $\mathcal{E} \in \mathbb{S}$, $m(\mathcal{E}) = 0$ if $\mathcal{E} \subset \mathcal{S}$ and $\mathcal{S} \notin \mathbb{S}$, and $\Sigma_{\mathcal{E} \in \mathbb{S}}\, m(\mathcal{E}) = 1$. For a subset $\mathcal{E}$ of $\mathcal{S}$, $m(\mathcal{E})$ characterizes the amount of "likelihood" that can be assigned to $\mathcal{E}$ but to no proper subset of $\mathcal{E}$. In the terminology of evidence theory, (i) $\mathcal{S}$ is the sample space or universal set, (ii) $\mathbb{S}$ is the set of focal elements for $\mathcal{S}$ and m, and (iii) $m(\mathcal{E})$ is the basic probability assignment (BPA) associated with a subset $\mathcal{E}$ of $\mathcal{S}$. The elements of $\mathcal{S}$ are often vectors $\mathbf{x} = [x_1, x_2, \ldots, x_n]$, where each element x_i of $\mathbf{x}$ is a variable with its own evidence space $(\mathcal{S}_i, \mathbb{S}_i, m_i)$. When the x_i's are assumed to be independent, (i) $m(\mathcal{E})$

$= \Pi_i\, m_i(\mathcal{E}_i)$ if $\mathcal{E} = \mathcal{E}_1 \times \mathcal{E}_2 \times \ldots \times \mathcal{E}_n$ and $\mathcal{E}_i \in \mathbb{S}_i$ for $i = 1, 2, \ldots, n$ and (ii) $m(\mathcal{E}) = 0$ otherwise. An evidence space ($\mathcal{S}$, $\mathbb{S}$, m) plays the same role in evidence theory that a probability space ($\mathcal{P}$, $\mathbb{P}$, p) plays in probability theory, where $\mathcal{P}$ is the sample space, $\mathbb{P}$ is a suitably restricted set of subsets of $\mathcal{P}$ (i.e., a σ-algebra), and p is the function (i.e., probability measure) that assigns probabilities to elements of $\mathbb{P}$.

The belief, $Bel(\mathcal{E})$, and plausibility, $Pl(\mathcal{E})$, for a subset $\mathcal{E}$ of $\mathcal{S}$ are defined by

$$Bel(\mathcal{E}) = \sum_{\mathcal{U} \subset \mathcal{E}} m(\mathcal{U}) \text{ and } Pl(\mathcal{E}) = \sum_{\mathcal{U} \cap \mathcal{E} \neq \phi} m(\mathcal{U}). \tag{2.1}$$

In concept, $Bel(\mathcal{E})$ is the amount of "likelihood" that must be assigned to $\mathcal{E}$, and $Pl(\mathcal{E})$ is the maximum amount of "likelihood" that could possibly be assigned to $\mathcal{E}$. When the elements of $\mathcal{S}$ are real valued, a CCBF and a CCPF provide a convenient summary of an evidence space ($\mathcal{S}$, $\mathbb{S}$, m) and correspond to plots of the points

$$\mathcal{CCBF} = \left\{\left[v, Bel(\mathcal{S}_v)\right], v \in \mathcal{S}\right\} \text{ and } \mathcal{CCPF} = \left\{\left[v, Pl(\mathcal{S}_v)\right], v \in \mathcal{S}\right\}, \tag{2.2}$$

where $\mathcal{S}_v = \{x\!: x \in \mathcal{S} \text{ and } x > v\}$.

An important situation in the application of evidence theory is the consideration of a variable $y = f(\mathbf{x})$, where f is a function defined for elements $\mathbf{x}$ of the sample space $\mathcal{X}$ associated with an evidence space ($\mathcal{X}$, $\mathbb{X}$, m_X) and $\mathbf{x}$ is represented as a vector because this is the case in most real analyses. The properties of f and ($\mathcal{X}$, $\mathbb{X}$, m_X) induce an evidence space ($\mathcal{Y}$, $\mathbb{Y}$, m_Y) on y, which provides a characterization of the uncertainty associated with y. In turn, this uncertainty can be summarized with a CCBF and a CCPF defined by

$$\mathcal{CCBF} = \left\{\left[v, Bel_X\left\{f^{-1}(\mathcal{Y}_v)\right\}\right], v \in \mathcal{Y}\right\} \text{ and } \mathcal{CCPF} = \left\{\left[v, Pl_X\left\{f^{-1}(\mathcal{Y}_v)\right\}\right], v \in \mathcal{Y}\right\}, \tag{2.3}$$

where Bel_X and Pl_X denote belief and plausibility defined with respect to ($\mathcal{X}$, $\mathbb{X}$, m_X) and $\mathcal{Y}_v = \{y\!: y \in \mathcal{Y} \text{ and } y > v\}$. The generation and analysis of CCBFs and CCPFs of the preceding form are fundamental parts of the use of evidence theory to characterize the uncertainty in model predictions.

3. EXPLORATORY SENSITIVITY ANALYSIS

An initial exploratory sensitivity analysis plays an important role in helping to guide any study that involves uncertain inputs. This is particularly true in uncertainty analyses based on evidence theory as the uncertainties are likely to be large and an appropriate understanding of these uncertainties and their implications can provide insights that facilitate the computational estimation of beliefs and plausibilities.

Given that large uncertainties in many variables are likely to be present, a sampling-based approach to sensitivity analysis with Latin hypercube sampling [18, 19] is a broadly applicable procedure for an exploratory analysis in conjunction with an evidence theory representation for uncertainty. Use of this approach requires the specification of distributions for the uncertain variables for sampling purposes. This specification should provide for an adequate exploration of the range of each uncertain variable and be consistent, in some sense, with the evidence theory specification of the uncertainty associated with individual analysis inputs.

A distribution that meets the preceding criteria can be obtained by sampling each focal element associated with a variable in consistency with its BPA and then sampling uniformly within that focal element. With the assumption that each focal element for a variable x_i with an evidence space ($\mathcal{X}_i$, $\mathbb{X}_i$, m_i) is an interval, this corresponds to defining a sampling distribution with a density function d_i given by

$$d_i(v) = \sum_{j=1}^{C(\mathbb{X}_i)} \delta_{ij}(v)\, m_i(\mathcal{E}_{ij}) / (b_{ij} - a_{ij}), \tag{3.1}$$

where (i) $v \in \mathcal{X}_i$, (ii) $C(\mathbb{X}_i)$ is the cardinality of $\mathbb{X}_i$, (iii) $\mathcal{E}_{ij} = [a_{ij}, b_{ij}]$, $j = 1, 2, \ldots, C(\mathbb{X}_i)$, are the focal elements associated with x_i (i.e., the elements of $\mathbb{X}_i$), and (iv) $\delta_{ij}(v) = 1$ if $v \in \mathcal{E}_{ij}$ and 0 otherwise. Appropriate modifications can be made to the preceding definition to handle focal elements with a finite number of elements and focal elements that are unions of disjoint intervals.

Given that a relationship of the form $y = f(\mathbf{x})$, $\mathbf{x} = [x_1, x_2, \ldots, x_n]$, is under consideration, sampling according to the distributions indicated in Eq. (3.1) generates a mapping $y_k = f(\mathbf{x}_k)$ from uncertain analysis inputs to uncertain analysis results, where $\mathbf{x}_k$, $k = 1, 2, \ldots, nS$, are the sampled values for $\mathbf{x}$. As previously indicated, Latin hypercube sampling is a likely candidate for the sampling procedure because of its efficient stratification properties. Once this mapping is generated, it can be explored with various sensitivity analysis procedures to develop an understanding of the relationship between y and the individual elements of $\mathbf{x}$.

A variety of techniques are available for use in sampling-based sensitivity analyses [13, 20]. However, given that the analysis problem is based on evidence theory, sensitivity analysis procedures that do not place excessive reliance on the sampling distributions indicated in Eq. (3.1) are desirable. Of course, no approach can fully divorce itself from these distributions because they ultimately give rise to the raw material of the sensitivity analysis (i.e., the mapping $[\mathbf{x}_k, y_k]$, $k = 1, 2, \ldots, nS$); however, this is an unavoidable situation when the sample space associated with $\mathbf{x}$ is infinite as no approach can consider all values of $\mathbf{x}$ and so a subset of the values for $\mathbf{x}$ must be selected in some manner. The examination of scatterplots is a natural initial procedure. Then, rank-based procedures (e.g., rank repression, partial rank correlation, squared rank differences) are natural techniques to employ because they reduce the effects of both nonlinearities and the original sampling distributions [13, 21, 22].

If carried out successfully, an initial exploratory sensitivity analysis should provide important insights with respect to the relationship between y and the elements of $\mathbf{x}$. Often, only a few of the elements of $\mathbf{x}$ will have significant effects on y. This is information that can be productively used in the estimation of the evidence theory structure associated with y.

4. STEPWISE CONSTRUCTION OF CCBFs AND CCPFs

For most models, the determination of beliefs and plausibilities for model predictions in general, and CCBFs and CCPFs in particular, is a demanding numerical challenge due to the need to determine the inverse of the model (i.e., function) involved. Sampling-based (i.e., Monte Carlo) procedures provide one way to carry out such determinations. With this approach, a sample $\mathbf{x}_k$, $k = 1, 2, \ldots, nS$, is generated from $\mathcal{X}$ (e.g., with distributions for the elements of $\mathbf{x}$ of the form indicated in Eq. (3.1)), and y is evaluated for each $\mathbf{x}_k$ to create the mapping $[\mathbf{x}_k, y_k]$, $k = 1, 2, \ldots, nS$, from $\mathcal{X}$ to $\mathcal{Y}$. Then, the CCBF and CCPF for y can be estimated by

$$\mathcal{CCBF} \cong \left\{\left[y, 1 - Pl_X\left(\{\mathbf{x}_k : y_k \le y\}\right)\right], y \in \mathcal{Y}\right\} \tag{4.1}$$

and

$$\mathcal{CCPF} \cong \left\{\left[y, Pl_X\left(\{\mathbf{x}_k : y_k > y\}\right)\right], y \in \mathcal{Y}\right\}, \tag{4.2}$$

respectively. The approximation to $\mathcal{CCBF}$ for y in Eq. (4.1) is based on the equality $Bel(\mathcal{E}) + Pl(\mathcal{E}^c) = 1$ and the fact that the subset criterion in the definition of belief (see Eq. (2.1)) does not allow for the direct estimate of belief with a finite sample when sets with infinite numbers of elements are under consideration. In general, the same approach can be used to estimate the belief $Bel_Y(\mathcal{E})$ and plausibility $Pl_Y(\mathcal{E})$ for any subset $\mathcal{E}$ of $\mathcal{Y}$.

The problem with the preceding approach is that it can be prohibitively expensive computationally when the cardinality $C(\mathbb{X})$ of $\mathbb{X}$ is high, which is usually the case in real analyses. Specifically, $C(\mathbb{X}) = \Pi_i\, C(\mathbb{X}_i)$, where $C(\mathbb{X}_i)$ is the cardinality of $\mathbb{X}_i$. For example, if $n = 8$ and $C(\mathbb{X}_i) = 10$, then $C(\mathbb{X}) = 10^8$; and as a result, a very large sample would be required to converge the approximations to the CCBF and CCPF in Eqs. (4.1) and (4.2).

The results of the exploratory sensitivity analysis described in Sect. 3 provide a basis for a potential path forward in developing the CCBF and CCPF approximations in Eqs. (4.1) and (4.2). The uncertainty in most analysis outcomes is significantly affected by the uncertainty in only a small number of analysis inputs (e.g., 3-5). Of course, this does not have to be the case but it does seem usually to be the case. In this situation, the approximations in Eqs. (4.1) and (4.2) can be determined by only considering the uncertainty (i.e., the evidence spaces $(\mathcal{X}_i, \mathbb{X}_i, m_i)$) associated with the x_i that significantly affect y. The remaining x_i (i.e., those that do not have a significant effect on y) can be assigned degenerate evidence spaces (i.e., spaces $(\mathcal{X}_i, \mathbb{X}_i, m_i)$ for which $m_i(\mathcal{X}_i) = 1$) for use in evaluating the approximations in Eqs. (4.1) and (4.2).

Increasing the resolution in the evidence spaces assigned to individual x_i (i.e., by subdividing elements of $\mathbb{X}_i$ and then apportioning the BPA for an original element of $\mathbb{X}_i$ over the subsets into which it is subdivided) tends to decrease, and can never increase, the uncertainty associated with evidence space for y. Specifically, beliefs tend to increase (and can never decrease) and plausibilities tend to decrease (and can never increase); or put another way, beliefs and plausibilities for subsets of $\mathcal{Y}$ move closer together as the resolution in the characterization of the uncertainties associated with the x_i is increased.

The preceding observations provide a basis for the use of sensitivity analysis results to guide a stepwise procedure for the construction of the CCBF and CCPF approximations in Eqs. (4.1) and (4.2). At Step 1, the approximations in Eqs. (4.1) and (4.2) are determined with the most important variable affecting the uncertainty in y assigned its original evidence space and all other variables assigned evidence spaces in which their original sample spaces are assigned a BPA of 1. At Step 2, the approximations in Eqs. (4.1) and (4.2) are determined with the two most important variables affecting the uncertainty in y assigned their original evidence spaces and all other variables assigned evidence spaces in which their original sample spaces are assigned a BPA of 1. Analogous steps follow for additional important variables determined in the sensitivity analysis until substantive changes in the CCBF and CCPF approximations in Eqs. (4.1) and (4.2) no longer occur, at which point the approximation procedure stops. This approach can produce substantial computational savings over what would be incurred if the approximations in Eqs. (4.1) and (4.2) were evaluated with the original evidence spaces assigned to all the x_i.

The construction procedure just outlined can also can be viewed as a sensitivity analysis in the context of evidence theory. The changes in the location of the CCBF and CCPF as additional variables are added in the preceding procedure provides an indication of the importance of individual variables with respect to the uncertainty in y characterized by $(\mathcal{Y}, \mathscr{Y}, m_Y)$. At an intuitive level, this approach is analogous to the use of stepwise regression analysis in traditional sensitivity analyses.

5. SUMMARY SENSITIVITY ANALYSIS

Together, a CCBF and CCPF for y provide bounds on all possible CCDFs for y that could derive from different distributions for the x_i that are consistent with their specified evidence spaces $(\mathcal{X}_i, \mathscr{X}_i, m_i)$. In particular, if $(\mathcal{P}_i, \mathscr{P}_i, p_i)$ is a probability space for x_i that is consistent with the evidence space $(\mathcal{X}_i, \mathscr{X}_i, m_i)$ for $i = 1, 2, \ldots, n$, then these probability spaces give rise to corresponding probability spaces $(\mathcal{P}_X, \mathscr{P}_X, p_X)$ and $(\mathcal{P}_Y, \mathscr{P}_Y, p_Y)$ for $\mathbf{x}$ and y with the CCDF associated with $(\mathcal{P}_Y, \mathscr{P}_Y, p_Y)$ falling somewhere between the CCBF and CCPF for y. Traditional sensitivity analysis methods can be used to investigate the relationships between the uncertainty in the x_i characterized by the probability spaces $(\mathcal{P}_i, \mathscr{P}_i, p_i)$ and the uncertainty in y characterized by the probability space $(\mathcal{P}_Y, \mathscr{P}_Y, p_Y)$. A possible approach is a variance decomposition for y that partitions the variance for y into the contributions to this variance from the individual x_i [8-10]. However, unlike a traditional sensitivity analysis in which the probability spaces $(\mathcal{P}_i, \mathscr{P}_i, p_i)$ are uniquely specified, there are many possibilities for the spaces $(\mathcal{P}_i, \mathscr{P}_i, p_i)$ in an evidence theory context and thus many possible variance decompositions for y. In variance-based sensitivity analysis, the variance $V(y)$ of y is expressed as

$$V(y) = \sum_{i=1}^{n} V_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} V_{ij} + \cdots + V_{12\cdots n}, \tag{5.1}$$

where V_i is the contribution of x_i to $V(y)$, V_{ij} is the contribution of the interaction of x_i and x_j to $V(y)$, and so on up to $V_{12\ldots n}$ which is the contribution of the interaction of $x_1, x_2, \ldots, x_n$ to $V(y)$. Possible sensitivity measures are provided by

$$s_i = V_i / V(y) \text{ and } s_{iT} = \left(V_i + \sum_{j \neq i} V_{ij} + \cdots + V_{12\cdots n} \right) \Big/ V(y), \tag{5.2}$$

where s_i the fraction of $V(y)$ contributed by x_i alone and s_{iT} is the fraction of $V(y)$ contributed by x_i and interactions of x_i with other variables. The term V_i is defined by iterated integrals involving the probability spaces for the individual variables. For example, when $n = 3$,

$$V_1 = \int_{\mathcal{P}_1} \left[\int_{\mathcal{P}_2} \int_{\mathcal{P}_3} f(x_1, x_2, x_3)\, d_3(x_3)\, d_2(x_2)\, dx_3 dx_2 \right]^2 d_1(x_1)\, dx_1 - E^2(y), \tag{5.3}$$

where d_i denotes the density function associated with ($\mathcal{P}_i$, $\mathbb{P}_i$, p_i) and $E(y)$ denotes the expected value of y; similar defining integrals hold for V_2 and V_3, and related, but more complicated, integrals define V_{12}, V_{13}, V_{23} and V_{123}. Analogous relationships hold for $n > 3$. By suitably orchestrating an analysis, V_i and s_i for $i = 1, 2, \ldots, n$ can be estimated with two independent random or Latin hypercube samples; further, s_i and s_{iT} for $i = 1, 2, \ldots, n$ can be estimated with a total of $n + 2$ suitably defined samples.

Three questions arise with respect to the implementation of a variance-based sensitivity analysis in the context of evidence theory: (i) How to select an appropriate spectrum of distributions for each x_i from the infinite number of distributions that are consistent with ($\mathcal{X}_i$, $\mathbb{X}_i$, m_i)?, (ii) How to implement the analysis in a computationally practicable manner for multiple distributions (i.e., multiple probability spaces ($\mathcal{P}_i$, $\mathbb{P}_i$, p_i)) for each x_i?, and (iii) How to display the results of the sensitivity analyses for multiple distributions of the x_i and hence multiple distributions for $\mathbf{x}$ and y?

The first question arises because there is no inherent structure associated with the infinite number of distributions for x_i that are consistent with ($\mathcal{X}_i$, $\mathbb{X}_i$, m_i). The situation is analogous to that encountered in an interval analysis for a real-valued quantity except that the uncertain quantity is now a probability space rather than a number. As there is no way to consider all probability spaces consistent with ($\mathcal{X}_i$, $\mathbb{X}_i$, p_i) and also no specific structure to guide the selection of individual probability spaces, some type of ad hoc procedure is needed to select representative probability spaces that are consistent with ($\mathcal{X}_i$, $\mathbb{X}_i$, p_i). Further, the number of selected distributions for each x_i must be relatively small; otherwise, the total number of combinations of selected distributions for all n variables will be too large to be computationally practicable.

An exploratory approach that should provide valuable information for many situations is to select three distributions for each x_i, with (i) one distribution emphasizing the smaller values associated with each focal element, (ii) one distribution uniform over the range of each focal element, and (iii) one distribution emphasizing the larger values associated with each focal element. The distributions indicated in (i) and (iii) could be left and right triangular or left and right quadratic. Left and right triangular distributions are actually quite similar to uniform distributions and thus may not be good choices. For focal element $\mathcal{E}_{ij} = [a_{ij}, b_{ij}]$ associated with x_i, the corresponding density functions d_{lij}, d_{uij} and d_{rij} for left quadratic, uniform, and right quadratic distributions, respectively, over $\mathcal{E}_{ij}$ are

$$d_{lij}(v) = \frac{3(b_{ij} - v)^2}{(b_{ij} - a_{ij})^3},\ d_{uij}(v) = \frac{1}{(b_{ij} - a_{ij})},\ \text{and } d_{rij}(v) = \frac{3(v - a_{ij})^2}{(b_{ij} - a_{ij})^3} \tag{5.4}$$

if $v \in \mathcal{E}_{ij}$ and $d_{lij}(v) = d_{uij}(v) = d_{rij}(v) = 0$ otherwise. In turn, the left quadratic, uniform and right quadratic distribution functions d_{li}, d_{ui} and d_{ri} for x_i are given by

$$d_{ci}(v) = \sum_{j=1}^{C(\mathcal{X}_i)} m_i(\mathcal{E}_{ij})\, d_{cij}(v) \tag{5.5}$$

for $v \in \mathcal{X}_i$ and $c = l, u, r$.

The second question arises because computational cost can easily become unreasonable unless the analysis is carefully planned. As a first step, only those variables that actually affect y need to be considered. The preliminary sensitivity analysis described in Sect. 3 should, in most analyses, identify the four or five variables that have significant effects on y. It is only those variables that require consideration of their original evidence spaces as indicated in Eq. (5.5); the remaining variables can be assigned a uniform or some other convenient distribution. For example, if four x_i affect y and the three distributions defined in Eq. (5.5) are considered for each of these x_i, then $3^4 = 81$ different probability spaces result for $\mathbf{x}$ and hence for y. As a second step, the analysis can be designed to use the same samples in the evaluation of s_i and s_{iT} for all probability spaces defined for $\mathbf{x}$ (e.g., the 81 spaces indicated above). For example, if Latin hypercube sampling is used, it is necessary to actually evaluate f for samples from only one of the probability spaces for $\mathbf{x}$; after these evaluations for f are performed, results for the other probability spaces for $\mathbf{x}$ under consideration (e.g., the remaining 80 probability spaces in the example above) can be obtained by reweighting the results obtained for the individual sample elements on the basis of the changed distributions for the x_i's [19, 23]. A similar reweighting procedure is also available for random sampling [24].

The third question arises because of the difficulty of displaying the results of multiple sensitivity analyses for y in a reasonably compact and understandable format. Presenting the sensitivity analyses individually is unlikely to be adequate because of the large number of analyses involved and the resultant difficulty of observing trends in variable importance across analyses. A promising presentation format to employ for this representation is a cobweb plot, which provides a representation for a multidimensional distribution in a two-dimensional plot [25]. For example, if nPS probability spaces ($\mathcal{P}_{Xj}$, $\mathbb{P}_{Xj}$, p_{Xj}) for $\mathbf{x}$ are under consideration and 4 uncertain variables have been identified for analysis, the results of the sensitivity analyses for y might be of the form

$$\mathbf{s}_j = \left[e_j, v_j, s_{1j}, s_{2j}, s_{3j}, s_{4j}\right],\ j = 1, 2, \ldots, nPS, \tag{5.6}$$

where e_j and v_j are the expected value and variance for y that derive from the probability space $(\mathcal{X}_{Xj}, \mathbb{X}_{Xj}, p_{Xj})$ for **x** and s_{ij}, $i = 1, 2, 3, 4$, are the fractional contributions to v_j as defined in the first equality in Eq. (5.2) for the 4 uncertain variables under consideration.

With a cobweb plot, the *nPS* vectors in Eq. (5.6) can be presented in a single plot frame. Specifically, the individual elements of $\mathbf{s}_j$ are designated by locations on the horizontal axis and their values correspond to locations on the vertical axis. In general, it may be necessary to use multiple axis scales for the vertical axis or to plot quantiles for the elements of $\mathbf{s}_j$ rather than their actual values. Each $\mathbf{s}_j$ results in a single point in each of the vertical columns associated with its elements. The identity of $\mathbf{s}_j$ is maintained by a line that connects the values of its elements. As desired, the cobweb plot allows the presentation of all sensitivity analysis results in a single plot frame and also facilitates the recognition of interactions between variables.

In summary, the approach presented in this section to the performance and presentation of a sensitivity analysis for a function defined on an evidence space has three components: (i) Definition of representative probability spaces for the analysis input **x** that are consistent with the evidence space for **x**, (ii) Use of efficient sampling-based numerical procedures to decompose the variance of the analysis outcome y for each probability space for **x**, and (iii) Use of cobweb plots to summarize the results of the sensitivity analyses for y carried out for the individual probability spaces for **x**. Thus, rather than having a single set of sensitivity analysis results for y, a spectrum of sensitivity analysis results for y is obtained that is consistent with the evidence space that characterizes the uncertainty in **x**.

6. SUMMARY

Three applications of sampling-based sensitivity analysis in conjunction with evidence theory representations for epistemic uncertainty in model inputs have been described: (i) an initial exploratory analysis to assess model behavior and provide insights for additional analysis, (ii) a stepwise analysis showing the incremental effects of uncertain variables on CCBFs and CCPFs, and (iii) a summary analysis showing a spectrum of variance-based sensitivity analysis results that derive from probability spaces that are consistence with the evidence space under consideration. It is hoped that the ideas associated with these approaches will provide a start towards the development of effective sensitivity analysis procedures for use in conjunction with evidence theory representations for epistemic uncertainty.

ACKNOWLEDGMENTS

Work performed for Sandia National Laboratories, which is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Security Administration under contract DE-AC04-94AL-85000. Editorial support provided by F. Puffer and J. Ripple of Tech Reps, a division of Ktech Corporation.

REFERENCES

[1] A. Saltelli, K. Chan, and E.M. Scott, *Sensitivity Analysis*. New York: Wiley, 2000.

[2] G. W. Parry and P. W. Winter, "Characterization and Evaluation of Uncertainty in Probabilistic Risk Analysis," *Nuclear Safety*, vol. 22, pp. 28-42, 1981.

[3] G. Apostolakis, "The Concept of Probability in Safety Assessments of Technological Systems," *Science*, vol. 250, pp. 1359-1364, 1990.

[4] J. C. Helton, "Uncertainty and Sensitivity Analysis in the Presence of Stochastic and Subjective Uncertainty," *Journal of Statistical Computation and Simulation*, vol. 57, pp. 3-76, 1997.

[5] T. Turányi, "Sensitivity Analysis of Complex Kinetic Systems. Tools and Applications," *Journal of Mathematical Chemistry*, vol. 5, pp. 203-248, 1990.

[6] D. G. Cacuci, *Sensitivity and Uncertainty Analysis, Vol. 1: Theory*. Boca Raton, FL: CRC Press, 2003.

[7] R. I. Cukier, H. B. Levine, and K. E. Shuler, "Nonlinear Sensitivity Analysis of Multiparameter Model Systems," *Journal of Computational Physics*, vol. 26, pp. 1-42, 1978.

[8] I. M. Sobol', "Sensitivity Analysis for Nonlinear Mathematical Models," *Mathematical Modeling and Computational Experiment*, vol. 1, pp. 407-414, 1993.

[9] H. Rabitz, O. F. Alis, J. Shorter, and K. Shim, "Efficient Input-Output Model Representations," *Computer Physics Communications*, vol. 117, pp. 11-20, 1999.

[10] A. Saltelli, S. Tarantola, and K. P.-S. Chan, "A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output," *Technometrics*, vol. 41, pp. 39-56, 1999.

[11] M. D. McKay, J. D. Morrison, and S. C. Upton, "Evaluating Prediction Uncertainty in Simulation Models," *Computer Physics Communications*, vol. 117, pp. 44-51, 1999.

[12] R. L. Iman, "Uncertainty and Sensitivity Analysis for Computer Modeling Applications," presented at *Reliability Technology - 1992, The Winter Annual Meeting of the American Society of Mechanical Engineers, Anaheim, California, November 8-13, 1992*, Vol. 28, pp. 153-168. New York, NY, 1992.

[13] J. C. Helton and F. J. Davis, "Sampling-Based Methods," in *Sensitivity Analysis*, A. Saltelli, K. Chan, and E.M. Scott, Eds. New York, NY: Wiley. pp. 101-153, 2000.

[14] J. P. C. Kleijnen and J. C. Helton, "Statistical Analyses of Scatterplots to Identify Important Factors in Large-Scale Simulations, 1: Review and Comparison of Techniques," *Reliability Engineering and System Safety*, vol. 65, pp. 147-185, 1999.

[15] G. J. Klir, "Is there more to uncertainty than some probability theorists might have us believe?," *International Journal of General Systems*, vol. 15, pp. 347-378, 1989.

[16] G. J. Klir, "On the Alleged Superiority of Probability Representation of Uncertainty," *IEEE Transactions on Fuzzy Systems*, vol. 2, pp. 27-31, 1994.

[17] G. J. Klir and M. J. Wierman, "Uncertainty-Based Information," Physica-Verlag, New York, NY 1999.

[18] M. D. McKay, R. J. Beckman, and W. J. Conover, "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, vol. 21, pp. 239-245, 1979.

[19] J. C. Helton and F. J. Davis, "Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems," *Reliability Engineering and System Safety*, vol. 81, pp. 23-69, 2003.

[20] J. C. Helton and F. J. Davis, "Illustration of Sampling-Based Methods for Uncertainty and Sensitivity Analysis," *Risk Analysis*, vol. 22, pp. 691-622, 2002.

[21] R. L. Iman and W. J. Conover, "The Use of the Rank Transform in Regression," *Technometrics*, vol. 21, pp. 499-509, 1979.

[22] S. C. Hora and J. C. Helton, "A Distribution-Free Test for the Relationship Between Model Input and Output when Using Latin Hypercube Sampling," *Reliability Engineering and System Safety*, vol. 79, pp. 333-339, 2003.

[23] R. L. Iman and W. J. Conover, "Small Sample Sensitivity Analysis Techniques for Computer Models, with an Application to Risk Assessment," *Communications in Statistics: Theory and Methods*, vol. A9, pp. 1749-1842, 1980.

[24] R. J. Beckman and M. D. McKay, "Monte-Carlo Estimation Under Different Distributions Using the Same Simulation," *Technometrics*, vol. 29, pp. 153-160, 1987.

[25] R. M. Cooke and J.M. van Noortwijk, "Graphical Methods," in *Sensitivity Analysis*, A. Saltelli, K. Chan, and E.M. Scott, Eds. Wiley: New York, NY, 2000, pp. 245-264.