Uncertainty and Sensitivity Analysis: From Regulatory Requirements to Conceptual Structure and Computational Implementation

Jon C. Helton[1] and Cédric J. Sallaberry[2]

[1] Department of Mathematics and Statistics, Arizona State University, Tempe, AZ 85287 USA
[2] Sandia National Laboratories, Albuquerque, NM 87185 USA
{jchelto,cnsalla}@sandia.gov

Abstract. An approach to the conversion of regulatory requirements into a conceptual and computational structure that permits meaningful uncertainty and sensitivity analyses is descibed. This approach is predicated on the description of the desired analysis in terms of three basic entities: (i) a probability space characterizing aleatory uncertainty, (ii) a probability space characterizing epistemic uncertainty, and (iii) a model that predicts system behavior. The presented approach is illustrated with results from the 2008 performance assessment for the proposed repository for high-level radioactive waste at Yucca Mountain, Nevada.

1   Introduction

An approach to the conversion of regulatory requirements into a conceptual and computational structure that permits meaningful uncertainty and sensitivity analyses is descibed. This approach is predicated on the description of the desired analysis in terms of three basic entities: (i) a probability space characterizing aleatory uncertainty, (ii) a probability space characterizing epistemic uncertainty, and (iii) a model that predicts system behavior. The presented approach is illustrated with results from the 2008 performance assessment (PA) for the proposed repository for high-level radioactive waste at Yucca Mountain (YM), Nevada, carried out by the U.S. Department of Energy (DOE) to assess compliance with regulations promulgated by the U.S. Nuclear Regulatory Commission (NRC) [1-3].

2   Example: DOE's Licensing Requirements for YM Repository

The NRC's licensing requirements for the YM repository provide a good example of the challenges that are present in the conversion of regulatory requirements into the conceptual structure and associated computational implementation of an analysis that establishes compliance (or noncompliance) with those requirements [4; 5].

The following two radiation protection requirements for a reasonably maximally exposed individual (RMEI) are at the core of the NRC's requirements for the YM repository ([6], p. 10829): "(a) DOE must demonstrate, using performance assessment, that there is a reasonable expectation that the reasonably maximally exposed individual receives no more than the following annual dose from releases from the undisturbed Yucca Mountain disposal system: (1) 0.15 mSv (15 mrem) for 10,000 years following disposal; and (2) 1.0 mSv (100 mrem) after 10,000 years, but within the period of geologic stability. (b) DOE's performance assessment must include all potential environmental pathways of radionuclide transport and exposure." In addition, the following elaboration on the preceding dose requirements for the RMEI is also given ([6], p. 10829): "Compliance is based upon the arithmetic mean of the projected doses from DOE's performance assessments for the period within 1 million years after disposal".

The preceding dose requirements indicate (i) that dose results must be determined for long time periods into the future and also for many different potential modes of exposure and (ii) that some type of averaging process is to be used to determine the dose values to which the regulatory requirements apply. The indicated averaging process (i.e., "arithmetic mean of projected doses") is vague and thus particularly challenging to the design of an analysis to assess compliance with the indicated bounds on (mean) dose. However, of necessity, implementation of this averaging process requires some form of a probabilistic representation of uncertainty.

Additional detail on what is desired in assessing compliance with the indicated dose requirements is provided by the NRC in the following definition for PA ([7], p. 55794): "Performance assessment means an analysis that: (1) Identifies the features, events, processes (except human intrusion), and sequences of events and processes (except human intrusion) that might affect the Yucca Mountain disposal system and their probabilities of occurring during 10,000 years after disposal, (2) Examines the effects of those features, events, processes, and sequences of events and processes upon the performance of the Yucca Mountain disposal system; and (3) Estimates the dose incurred by the reasonably maximally exposed individual, including the associated uncertainties, as a result of releases caused by all significant features, events, processes, and sequences of events and processes, weighted by their probability of occurrence."

The preceding definition makes very clear that a PA used to assess regulatory compliance for the YM repository must (i) consider what could happen in the future, (ii) assign probabilities to what could happen in the future, (iii) model the effects of what could happen in the future, (iv) consider the effects of uncertainties, and (v) weight potential doses by the probability of the occurrence of such doses. Of particular interest and importance to the design of an analysis to assess compliance is the indicated distinction between "uncertainty" and "probability of occurrence". This is a distinction between what is often called epistemic uncertainty and aleatory uncertainty [8; 9]. Specifically, epistemic uncertainty derives from a lack of knowledge about the appropriate value to use for a quantity that is assumed to have a fixed value in the context of a particular analysis, and aleatory uncertainty derives from an inherent randomness in the properties or behavior of the system under study.

The NRC further emphasizes the importance of an appropriate treatment of uncertainty in assessing regulatory compliance for the YM repository in the following definition of reasonable expectation ([7], p. 55813): "Reasonable expectation means that the Commission is satisfied that compliance will be achieved based upon the full record before it. Characteristics of reasonable expectation include that it: (1) Requires less than absolute proof because absolute proof is impossible to attain for disposal due to the uncertainty of projecting long-term performance; (2) Accounts for the inherently greater uncertainties in making long-term projections of the performance of the Yucca Mountain disposal system; (3) Does not exclude important parameters from assessments and analyses simply because they are difficult to precisely quantify to a high degree of confidence; and (4) Focuses performance assessments and analyses on the full range of defensible and reasonable parameter distributions rather than only upon extreme physical situations and parameter values." As the preceding definition makes clear, the NRC intends that a thorough treatment of uncertainty is to be an important part of assessing compliance with licensing requirements for the YM repository.

Similar requirements to the NRC's requirements for the YM repository, either by explicit statement or implication, underlie requirements for analyses of other complex systems, including (i) the NRC's safety goals for nuclear power stations [10], (ii) the U.S. Environmental Protection Agency's certification requirements for the Waste Isolation Pilot Plant [11; 12], and (iii) the National Nuclear Security Administration's mandate for the quantification of margins and uncertainties in assessments of the nation's nuclear stockpile [13-15]. Three recurrent ideas run through all of these examples: (i) the occurrence of future events (i.e., aleatory uncertainty), (ii) prediction of the consequences of future events (i.e., the modeling of physical processes), and (iii) lack of knowledge with respect to appropriate models and associated model parameters (i.e., epistemic uncertainty). The challenge in each case is to define a conceptual model and an associated computational implementation that appropriately incorporates these ideas into analyses supporting compliance determinations.

3 Conceptual Structure and Computational Implementation

The needed conceptual structure and path to computational implementation is provided by viewing the analysis of a complex system as being composed of three basic entities: (i) a probability space (A, A, p_A) characterizing aleatory uncertainty, (ii) a probability space (E, E, p_E) characterizing epistemic uncertainty, and (iii) a model that predicts system behavior (i.e., a function $f(t|\mathbf{a}, \mathbf{e})$, or more typically a vector function $\mathbf{f}(t|\mathbf{a}, \mathbf{e})$, that defines system behavior at time t conditional on elements $\mathbf{a}$ and $\mathbf{e}$ of the sample spaces A and E for aleatory and epistemic uncertainty). In the context of the three recurrent ideas indicated at the end of the preceding section, the probability space (A, A, p_A) defines future events and their probability of occurrence; the functions $f(t|\mathbf{a}, \mathbf{e})$ and $\mathbf{f}(t|\mathbf{a}, \mathbf{e})$ predict the consequences of future events; and the probability space (E, E, p_E) defines "state of knowledge uncertainty" with respect to the ap-

propriate values to use for analysis inputs and characterizes this uncertainty with probability.

In turn, this conceptual structure leads to an analysis in which (i) uncertainty in analysis results is defined by integrals involving the function $f(t|\mathbf{a}, \mathbf{e})$ and the two indicated probability spaces and (ii) sensitivity analysis results are defined by the relationships between epistemically uncertain analysis inputs (i.e., elements e_j of $\mathbf{e}$) and analysis results defined by the function $f(t|\mathbf{a}, \mathbf{e})$ and also by various integrals of this function. Computationally, this leads to an analysis in which (i) high-dimensional integrals must be evaluated to obtain uncertainty analysis results and (ii) mappings between high-dimensional spaces must be generated and explored to obtain sensitivity analysis results. In general, $f(t|\mathbf{a}, \mathbf{e})$ is just one component of a high dimensional function $\mathbf{f}(t|\mathbf{a}, \mathbf{e})$. It is also possible for $f(t|\mathbf{a}, \mathbf{e})$ and $\mathbf{f}(t|\mathbf{a}, \mathbf{e})$ to be functions of spatial coordinates as well as time.

In general, the elements $\mathbf{a}$ of A are vectors

$$\mathbf{a} = \left[a_1, a_2, ..., a_m \right] \tag{1}$$

that define one possible occurrence in the universe under consideration. In practice, the uncertainty structure formally associated with the set A and the probability measure p_A is defined by defining probability distributions for the individual elements a_i of $\mathbf{a}$. Formally, this corresponds to defining a density function $d_{Ai}(a_i)$ on a set A_i characterizing aleatory for each element a_i of $\mathbf{a}$ (or some other uncertainty structure such as a cumulative distribution function (CDF) or a complementary CDF (CCDF) when convenient). Collectively, the sets A_i and density functions $d_{Ai}(a_i)$, or other appropriate uncertainty characterizations, define the set A and a density function $d_A(\mathbf{a})$ for $\mathbf{a}$ on A, and thus, in effect, define the probability space (A, A, p_A).

Similarly, the elements $\mathbf{e}$ of E are vectors

$$\mathbf{e} = \left[\mathbf{e}_A, \mathbf{e}_M \right] = \left[e_1, e_2, ..., e_n \right] \tag{2}$$

that define one possible set of epistemically uncertainty analysis inputs, where the vector $\mathbf{e}_A$ contains uncertain quantities used in the characterization of aleatory uncertainty and the vector $\mathbf{e}_M$ contains uncertain quantities used in the modeling of physical processes. As in the characterization of aleatory uncertainty, the uncertainty structure formally associated with the set E and the probability measure p_E is defined by defining probability distributions for the individual elements e_i of $\mathbf{e}$. Formally, this corresponds to defining a density function $d_{Ei}(e_i)$ (or some other uncertainty structure such as a CDF or CCDF when convenient) on a set E_i characterizing epistemic uncertainty for each element e_i of $\mathbf{e}$. Collectively, the sets E_i and density functions $d_{Ei}(e_i)$, or other appropriate uncertainty characterizations, define the set E and a density function $d_E(\mathbf{e})$ for $\mathbf{e}$ on E, and thus, in effect, define the probability space (E, E, p_E). In practice, the distributions for the individual elements of $\mathbf{e}$ are often obtained through an extensive expert review process (e.g., [16]).

The model, or system of models, that predict analysis results can be represented by

$$y\left(t\mid\mathbf{a},\mathbf{e}_M\right)=f\left(t\mid\mathbf{a},\mathbf{e}_M\right) \text{ for a single result} \tag{3}$$

and

$$\mathbf{y}\left(t\mid\mathbf{a},\mathbf{e}_M\right)=\left[y_1\left(t\mid\mathbf{a},\mathbf{e}_M\right),y_2\left(t\mid\mathbf{a},\mathbf{e}_M\right),...\right]=\mathbf{f}\left(t\mid\mathbf{a},\mathbf{e}_M\right) \text{ for multiple results,} \tag{4}$$

where t represents time. In practice, $f(t\mid\mathbf{a},\mathbf{e})$ and $\mathbf{f}(t\mid\mathbf{a},\mathbf{e})$ are very complex computer models and may produce results with a spatial as well as a temporal dependency.

In concept, the probability space (A, A, p_A) and the function $y(t\mid\mathbf{a},\mathbf{e})=f(t\mid\mathbf{a},\mathbf{e})$ are sufficient to determine the expected value $E_A[y(t\mid\mathbf{a},\mathbf{e})]$ of $y(t\mid\mathbf{a},\mathbf{e})$ over aleatory uncertainty conditional on the values for uncertain analysis inputs defined by an element $\mathbf{e}=[\mathbf{e}_A,\mathbf{e}_M]$ of E (i.e., risk in the terminology of many analyses and expected dose in the terminology of the NRC's regulations for the YM repository). Specifically,

$$E_A\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\mid\mathbf{e}_A\right]=\int_A y\left(t\mid\mathbf{a},\mathbf{e}_M\right)d_A\left(\mathbf{a}\mid\mathbf{e}_A\right)\mathrm{dA}$$

$$= \quad \cong \begin{cases} \sum_{j=1}^{nS} y\left(t\mid\mathbf{a}_j,\mathbf{e}_M\right)/nS & \begin{cases} \text{for random sampling with } \mathbf{a}_j, j=1,2,...,nS, \\ \text{sampled from A consistent with } d_A(\mathbf{a}\mid\mathbf{e}_A) \end{cases} \\ \sum_{j=1}^{nS} y\left(t\mid\mathbf{a}_j,\mathbf{e}_M\right)p_A\left(A_j\mid\mathbf{e}_A\right) & \begin{cases} \text{for stratified sampling with } \mathbf{a}_j\in A_j \\ \text{and } A_j, j=1,2,...,nS, \text{ partitioning A.} \end{cases} \end{cases} \tag{5}$$

with the inclusion of "$\mid\mathbf{e}_A$" in $d_A(\mathbf{a}\mid\mathbf{e}_A)$ and $p_A(A_j\mid\mathbf{e}_A)$ indicating that the distribution (i.e., probability space) for $\mathbf{a}$ is dependent on epistemically uncertain quantities that are elements of $\mathbf{e}_A$. Similarly, the probabilities that define CDFs and CCDFs that show the effects of aleatory uncertainty conditional on a specific element $\mathbf{e}=[\mathbf{e}_A,\mathbf{e}_M]$ of E are defined by

$$p_A\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\le y\mid\mathbf{e}_A\right]=\int_A \underline{\delta}_y\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\right]d_A\left(\mathbf{a}\mid\mathbf{e}_A\right)\mathrm{dA} \tag{6}$$

and

$$p_A\left[y<y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\mid\mathbf{e}_A\right]=\int_A \overline{\delta}_y\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\right]d_A\left(\mathbf{a}\mid\mathbf{e}_A\right)\mathrm{dA}, \tag{7}$$

respectively, where

$$\underline{\delta}_y\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\right]=\begin{cases} 1 \text{ for } y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\le y \\ 0 \text{ for } y<y\left(t\mid\mathbf{a},\mathbf{e}_M\right) \end{cases} \tag{8}$$

and

$$\overline{\delta}_y\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\right]=1-\underline{\delta}_y\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\right]=\begin{cases}0 \text{ for } y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\le y\\ 1 \text{ for } y<y\left(t\mid\mathbf{a},\mathbf{e}_M\right).\end{cases}\qquad(9)$$

The integrals in Eqs. (6) and (7) can be approximated with procedures analogous to the sampling-based procedures indicated in Eq. (5).

The integrals in Eqs. (5)-(7) must be evaluated with multiple values of $\mathbf{e} = [\mathbf{e}_A, \mathbf{e}_M]$ in order to determine the effects of epistemic uncertainty. As illustrated in Sect. 4, the indicated multiple values for $\mathbf{e} = [\mathbf{e}_A, \mathbf{e}_M]$ are often obtained with a Latin hypercube sample (LHS)

$$\mathbf{e}_k =\left[\mathbf{e}_{Ak},\mathbf{e}_{Mk}\right]=\left[e_{1k},e_{2k},...,e_{nk}\right], k=1,2,...,nLHS,\qquad(10)$$

of size $nLHS$ from the sample space E for epistemic uncertainty due to the efficient stratification properties of Latin hypercube sampling [17; 18]. This sample provides the basis for both (i) the numerical estimation of the effects of epistemic uncertainty and (ii) the implementation of a variety sensitivity analysis procedures [19-21].

Just as expected values, CDFs and CCDFs related to aleatory can be defined as indicated in Eqs. (5)-(7), similar quantities can be defined that summarize the effects of epistemic uncertainty. Several possibilities exist: (i) epistemic uncertainty in a result $y(t\mid\mathbf{a}, \mathbf{e}_M)$ conditional on a specific realization $\mathbf{a}$ of aleatory uncertainty, (ii) epistemic uncertainty in an expected value over aleatory uncertainy, and (iii) epistemic uncertainty in the cumulative probability $p_A[y(t\mid\mathbf{a}, \mathbf{e}_M)\le y|\mathbf{e}_A]$ or exceedance (i.e., complementary cumulative) probability $p_A[y<y(t\mid\mathbf{a}, \mathbf{e}_M) \mid\mathbf{e}_A]$ for a specific value y of an analysis result.

For a result $y(t\mid\mathbf{a}, \mathbf{e}_M)$ conditional on a specific realization $\mathbf{a}$ of aleatory uncertainty, the expected value, cumulative probability and exceedance probability over epistemic uncertainty are given by

$$E_E\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\right]=\int_{EM} y\left(t\mid\mathbf{a},\mathbf{e}_M\right)d_E\left(\mathbf{e}_M\right)\mathrm{dEM} \cong \sum_{k=1}^{nLHS} y\left(t\mid\mathbf{a},\mathbf{e}_{Mk}\right)/nLHS, \quad(11)$$

$$p_E\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\le y\right]=\int_{EM}\underline{\delta}_y\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\right]d_E\left(\mathbf{e}_M\right)\mathrm{dEM}$$
$$\cong \sum_{k=1}^{nLHS}\underline{\delta}_y\left[y\left(t\mid\mathbf{a},\mathbf{e}_{Mk}\right)\right]/nLHS,\qquad(12)$$

and

$$p_E\left[y<y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\right]=\int_{EM}\overline{\delta}_y\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\right]d_E\left(\mathbf{e}_M\right)\mathrm{dEM}$$
$$\cong \sum_{k=1}^{nLHS}\overline{\delta}_y\left[y\left(t\mid\mathbf{a},\mathbf{e}_{Mk}\right)\right]/nLHS,\qquad(13)$$

respectively, where (i) EM corresponds to the subspace of E that contains only the vectors $\mathbf{e}_M$ and (ii) the vectors $\mathbf{e}_{Mk}$ are part of the LHS in Eq. (10).

For an expected result $E_A[y(t|\mathbf{a},\mathbf{e}_M)|\mathbf{e}_A]$ over aleatory uncertainty, the expected value, cumulative probability and exceedance probability over epistemic uncertainty are defined analogously to the corresponding results in Eqs. (11)-(13). Specifically,

$$
\begin{aligned}
E_E\left\{E_A\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\mid\mathbf{e}_A\right]\right\} &= \int_E\left[\int_A y\left(t\mid\mathbf{a},\mathbf{e}_M\right)d_A\left(\mathbf{a}\mid\mathbf{e}_A\right)\mathrm{dA}\right]d_E\left(\mathbf{e}\right)\mathrm{dE} \\
&\cong \sum_{k=1}^{nLHS}\int_A y\left(t\mid\mathbf{a},\mathbf{e}_{Mk}\right)d_A\left(\mathbf{a}\mid\mathbf{e}_{Ak}\right)\mathrm{dA}\,/\,nLHS,
\end{aligned}
\tag{14}
$$

$$
\begin{aligned}
p_E\left\{E_A\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\mid\mathbf{e}_A\right]\le\overline{y}\right\} &= \int_E\underline{\delta}_{\overline{y}}\left[\int_A y\left(t\mid\mathbf{a},\mathbf{e}_M\right)d_A\left(\mathbf{a}\mid\mathbf{e}_A\right)\mathrm{dA}\right]d_E\left(\mathbf{e}\right)\mathrm{dE} \\
&\cong \sum_{k=1}^{nLHS}\underline{\delta}_{\overline{y}}\left[\int_A y\left(t\mid\mathbf{a},\mathbf{e}_{Mk}\right)d_A\left(\mathbf{a}\mid\mathbf{e}_{Ak}\right)\mathrm{dA}\right]/\,nLHS,
\end{aligned}
\tag{15}
$$

and

$$
\begin{aligned}
p_E\left\{\overline{y}<E_A\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\mid\mathbf{e}_A\right]\right\} &= \int_E\overline{\delta}_{\overline{y}}\left[\int_A y\left(t\mid\mathbf{a},\mathbf{e}_M\right)d_A\left(\mathbf{a}\mid\mathbf{e}_A\right)\mathrm{dA}\right]d_E\left(\mathbf{e}\right)\mathrm{dE} \\
&\cong \sum_{k=1}^{nLHS}\overline{\delta}_{\overline{y}}\left[\int_A y\left(t\mid\mathbf{a},\mathbf{e}_{Mk}\right)d_A\left(\mathbf{a}\mid\mathbf{e}_{Ak}\right)\mathrm{dA}\right]/\,nLHS,
\end{aligned}
\tag{16}
$$

where, in general, the inner integrals over A will have to be evaluated with some appropriate integration procedure as indicated in Eq. (5).

For a cumulative probability $p_A[y(t|\mathbf{a},\mathbf{e}_M)\le y|\mathbf{e}_A]$ or an exceedance (i.e., complementary cumulative) probability $p_A[y<y(t|\mathbf{a},\mathbf{e}_M)|\mathbf{e}_A]$ over aleatory uncertainty for a specific value y of an analysis result, the expected value, cumulative probability and exceedance probability over epistemic uncertainty are defined analogously to the corresponding results in Eqs.(14)-(16). For example, the expected value and cumulative probability for $p_A[y<y(t|\mathbf{a},\mathbf{e}_M)|\mathbf{e}_A]$ that derive from epistemic uncertainty are defined by

$$
\begin{aligned}
E_E\left\{p_A\left[y<y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\mid\mathbf{e}_A\right]\right\} &= \int_E\left\{\int_A\overline{\delta}_y\left[y\left(t\mid\mathbf{a},\mathbf{e}_M\right)\right]d_A\left(\mathbf{a}\mid\mathbf{e}_A\right)\mathrm{dA}\right\}d_E\left(\mathbf{e}\right)\mathrm{dE} \\
&\cong \sum_{k=1}^{nLHS}\int_A\overline{\delta}_y\left[y\left(t\mid\mathbf{a},\mathbf{e}_{Mk}\right)\right]d_A\left(\mathbf{a}\mid\mathbf{e}_{Ak}\right)\mathrm{dA}\,/\,nLHS
\end{aligned}
\tag{17}
$$

and

$$p_E\left\{p_A\left[y<y(t\mid\mathbf{a},\mathbf{e}_M)\mid\mathbf{e}_A\right]\le p\right\}$$

$$=\int_E\underline{\delta}_P\left\{\int_A\overline{\delta}_y\left[y(t\mid\mathbf{a},\mathbf{e}_M)\right]d_A(\mathbf{a}\mid\mathbf{e}_A)dA\right\}d_E(\mathbf{e})dE \qquad(18)$$

$$\cong\sum_{k=1}^{nLHS}\underline{\delta}_P\left\{\int_A\overline{\delta}_y\left[y(t\mid\mathbf{a},\mathbf{e}_{Mk})\right]d_A(\mathbf{a}\mid\mathbf{e}_{Ak})dA\right\}/nLHS,$$

respectively.

The conceptual structure and computational procedures described in this section are illustrated in the next section with results from the 2008 YM PA [1].

4   Example: 2008 PA for YM Repository

The individual elements of the sample space A for aleatory uncertainty in the 2008 YM PA are vectors of the form

$$\mathbf{a}=\left[nEW,nED,nII,nIE,nSG,nSF,\mathbf{a}_{EW},\mathbf{a}_{ED},\mathbf{a}_{II},\mathbf{a}_{IE},\mathbf{a}_{SG},\mathbf{a}_{SF}\right], \qquad(19)$$

where, for a time interval $[a, b]$ (e.g., $[0, 10^4$ yr] or $[0, 10^6$ yr]), nEW = number of early waste package (WP) failures, nED = number of early drip shield (DS) failures, nII = number of igneous intrusive (II) events, nIE = number of igneous eruptive (IE) events, nSG = number of seismic ground (SG) motion events, nSF = number of seismic fault (SF) displacement events, $\mathbf{a}_{EW}$ = vector defining the nEW early WP failures, $\mathbf{a}_{ED}$ = vector defining the nED early DS failures, $\mathbf{a}_{II}$ = vector defining the nII igneous intrusive events, $\mathbf{a}_{IE}$ = vector defining the nIE igneous eruptive events, $\mathbf{a}_{SG}$ = vector defining the nSG seismic ground motion events, and $\mathbf{a}_{SF}$ = vector defining the nSF seismic fault displacement events. The definition of the probability space (A, A, p_A) for aleatory uncertainty was completed by defining probability distributions for the individual elements of $\mathbf{a}$ (see [1], App. J). In the 2008 YM PA, elements of the sample space A are referred to as scenarios, and elements of the set A are referred to as scenario classes. With this usage, scenarios and scenario classes correspond to what are called elementary events and events, respectively, in the usual terminology of probability theory.

The individual elements of the sample space E for epistemic uncertainty in the 2008 YM PA are vectors of the form

$$\mathbf{e}=\left[\mathbf{e}_A,\mathbf{e}_M\right]=\left[e_1,e_2,...,e_{392}\right], \qquad(20)$$

where, as examples, the following quantities are elements of $\mathbf{e}$: $DSNFMASS$ = scale factor used to characterize uncertainty in radionuclide content of defense spent nuclear fuel; $IGRATE$ = frequency of intersection of the repository footprint by a volcanic event (yr^{-1}); $MICTC99$ = groundwater biosphere dose conversion factor (BDCF) for ^{99}Tc in modern interglacial climate; $SCCTHRP$ = residual stress threshold for stress

8

corrosion cracking nucleation of Alloy 22 (as a percentage of yield strength in MPa); *SZGWSPDM* = logarithm of scale factor used to characterize uncertainty in ground-water specific discharge (dimensionless); and *WDGCA22* = temperature dependent slope term of Alloy 22 general corrosion rate (K). Distributions characterizing epistemic uncertainty were assigned to the individual elements of **e** and, in effect, defined the probability space (E, E, p_E) for epistemic uncertainty. A complete listing of the 392 elements of **e** and sources of additional information on these variables and the development of their distributions are given in Table K3-3 of Ref. [1].
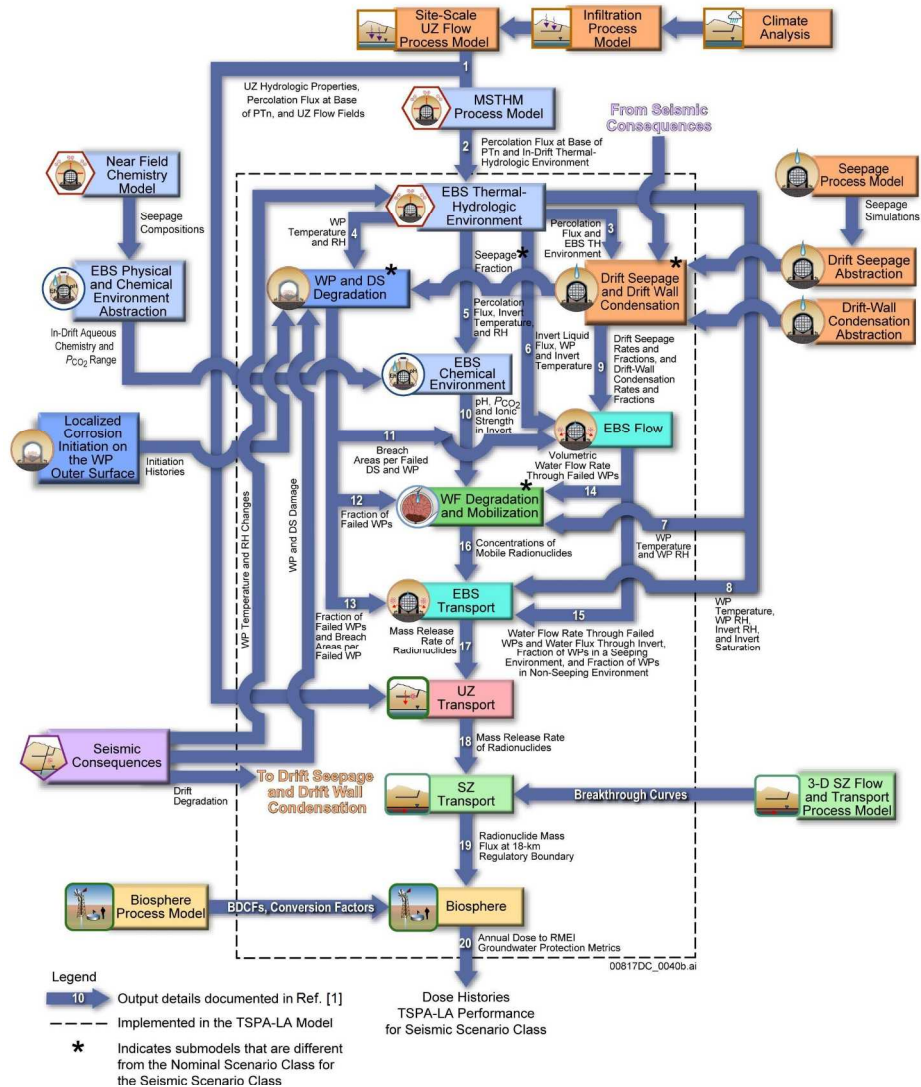
A very complex system of models was used to predict a large number of time-dependent results related to evolution of the repository, including (i) the release of radionuclides from WPs, (ii) the transport of radionuclides away from the engineered component of the repository, and (iii) human exposure to released radionuclides (see [1], Table K3-4, for a listing of the analysis results selected for study). An overview description of these models and extensive sources of additional information are available in Ref. [1]. As an example, a high-level overview of the models used in the analysis of seismic ground motion events is given in Fig. 1. The models indicated in Fig. 1 correspond to part of what is very simplistically represented by **f**(t|**a**, **e**) in Eq. (4).

Owing to its central role in the NRC's regulatory requirements for the YM repository,

$$D\left(t \mid \mathbf{a}, \mathbf{e}_M\right) = \text{dose (mrem/yr) to the RMEI at time } t \text{ (yr) conditional on } \mathbf{a} \in \mathrm{A}$$
$$\text{and } \mathbf{e} = \left[\mathbf{e}_A, \mathbf{e}_M\right] \in \mathrm{E} \tag{21}$$

will be used as an example to illustrate results of the form indicated in Sect. 3. It is in the generation of such results where the challenge of bringing conceptual structure and computational implementation together arises. Bluntly put, it is not possible to evaluate integrals of the form indicated in Sect. 3 for $D(t|\mathbf{a}, \mathbf{e}_M)$ defined by modeling systems of the complexity shown in Fig. 1 without a carefully designed computational strategy that makes efficient use of what will almost always be a limited number of

Fig. 1. Information transfer between the model components and submodels for the seismic scenario class in the 2008 YM PA ([1], Fig. 6.1.4-6).

detailed, mechanistic calculations. Such a strategy will be analysis specific and designed to take advantage of particular properties of the models in use.

A core quantity in the NRC's regulatory requirements for the YM repository is the expected value $E_A[D(t|\mathbf{a}, \mathbf{e}_M)|\mathbf{e}_A]$ of $D(t|\mathbf{a}, \mathbf{e}_M)$ over aleatory uncertainty, with $E_A[D(t|\mathbf{a}, \mathbf{e}_M)|\mathbf{e}_A]$ being an example of the expected value formally defined in Eq. (5). Specifically,

$$E_A\left[D\left(t\mid \mathbf{a},\mathbf{e}_M\right)\mid \mathbf{e}_A\right] = \int_A D\left(t\mid \mathbf{a},\mathbf{e}_M\right)d_A\left(\mathbf{a}\mid \mathbf{e}_A\right)\mathrm{dA}$$
$$\cong \int_A\left[\sum_{C\in\mathrm{C}}D_C\left(t\mid \mathbf{a},\mathbf{e}_M\right)\right]d_A\left(\mathbf{a}\mid \mathbf{e}_A\right)\mathrm{dA},\ \mathrm{C}=\{N,EW,ED,II,IE,SG,SF\}$$
$$=\sum_{C\in\mathrm{C}}\int_A D_C\left(t\mid \mathbf{a},\mathbf{e}_M\right)d_A\left(\mathbf{a}\mid \mathbf{e}_A\right)\mathrm{dA}==$$
$$=\sum_{C\in\mathrm{C}}E_A\left[D_C\left(t\mid \mathbf{a},\mathbf{e}_M\right)\mid \mathbf{e}_A\right],\tag{22}$$

where the subscripts contained in the set X are used to denote doses to the RMEI from individual scenario classes with N designating the scenario class in which no disruptions occur and the remaining subscripts designating scenario classes associated with the correspondingly subscripted aleatory occurrences in Eq. (19). Results associated with dose $D_C(t|\mathbf{a}, \mathbf{e}_M)$ for scenario class C in Eq. (22) are assumed to be calculated with only those elements of $\mathbf{a}$ and $\mathbf{e} = [\mathbf{e}_A, \mathbf{e}_M]$ that are related to scenario class C. The approximation to $E_A[D(t|\mathbf{a}, \mathbf{e}_M)|\mathbf{e}_A]$ in Eq. (22) can be justified on the basis of tradeoffs between the effects of high probability-low consequence scenario classes and low probability-high consequence scenario classes.

Epistemic uncertainty is propagated in the 2008 YM PA with an LHS

$$\mathbf{e}_k=\left[\mathbf{e}_{Ak},\mathbf{e}_{Mk}\right]=\left[e_{1k},e_{2k},...,e_{392k}\right], k=1,2,...,nLHS=300,\tag{23}$$

from the sample space E for epistemic uncertainty. In turn, the approximation

$$\hat{E}_A\left[D\left(t\mid \mathbf{a},\mathbf{e}_{Mk}\right)\mid \mathbf{e}_{Ak}\right]\cong \sum_{C\in\mathrm{C}}\hat{E}_A\left[D_C\left(t\mid \mathbf{a},\mathbf{e}_{Mk}\right)\mid \mathbf{e}_{Ak}\right]\tag{24}$$

results for each element $\mathbf{e}_k$ of the LHS in Eq. (23), where $\hat{E}_A[D(t\mid \mathbf{a},\mathbf{e}_{Mk})\mid \mathbf{e}_{Ak}]$ and $\hat{E}_A[D_C(t\mid \mathbf{a},\mathbf{e}_{Mk})\mid \mathbf{e}_{Ak}]$ denote approximations to $E_A[D(t\mid \mathbf{a},\mathbf{e}_{Mk})\mid \mathbf{e}_{Ak}]$ and $E_A[D_C(t\mid \mathbf{a},\mathbf{e}_{Mk})\mid \mathbf{e}_{Ak}]$, respectively. Because of the occurrence of the same elements of $\mathbf{e}_M$ in the evaluation of $D_C(t|\mathbf{a}, \mathbf{e}_M)$ for different values of C (i.e., for different scenario classes), it is essential that the doses $D_C(t|\mathbf{a}, \mathbf{e}_M)$ in Eq. (24) be evaluated for the same elements of the LHS in Eq. (23) for the indicated approximation to $E_A[D(t\mid \mathbf{a},\mathbf{e}_{Mk})\mid \mathbf{e}_{Ak}]$ to be valid.

As an example, analysis results are presented for seismic ground motion events occurring in the time interval $[0, 2\times10^4$ yr] (i.e., for what is called the seismic ground motion scenario class in the 2008 YM PA). This restriction reduces the elements of the sample space A for aleatory uncertainty to

$$\mathbf{a}=\left[nSG,t_1,v_1,A_1,t_2,v_2,A_2,...,t_{nSG},v_{nSG},A_{nSG}\right],\tag{25}$$

where (i) nSG = number of seismic ground motion events in 20,000 yr, (ii) t_i = time (yr) of event i, (iii) v_i = peak ground velocity (m/s) for event i, (iv) A_i = damaged area (m^2) on individual WPs for peak ground velocity v_i, (v) the occurrence of seismic

ground motion events is characterized by a hazard curve for peak ground velocity, and (vi) damaged area is characterized by distributions conditional on peak ground velocity.

To evaluate results of the form defined in Sect. 3 for the seismic ground motion scenario class, it is necessary to integrate the function $D_{SG}(t|\mathbf{a}, \mathbf{e}_M)$ over the sample space A for aleatory uncertainty with $\mathbf{a}$ defined as indicated in Eq. (25). In full detail, $D_{SG}(t|\mathbf{a}, \mathbf{e}_M)$ is defined by the model system shown in Fig. 1. Evaluation of this system is too computationally demanding to permit its evaluation 1000's of times for each element $\mathbf{e}_k = [\mathbf{e}_{Ak}, \mathbf{e}_{Mk}]$ of the LHS in Eq. (23). This is a common situation in analyses of complex systems, where very detailed physical models are developed which then turn out to be too computationally demanding to be naively used in the propagation of aleatory uncertainty. In such situations, it is necessary to find ways to efficiently use the results of a limited number of model evaluations to predict outcomes for a large number of different possible realizations of aleatory uncertainty.
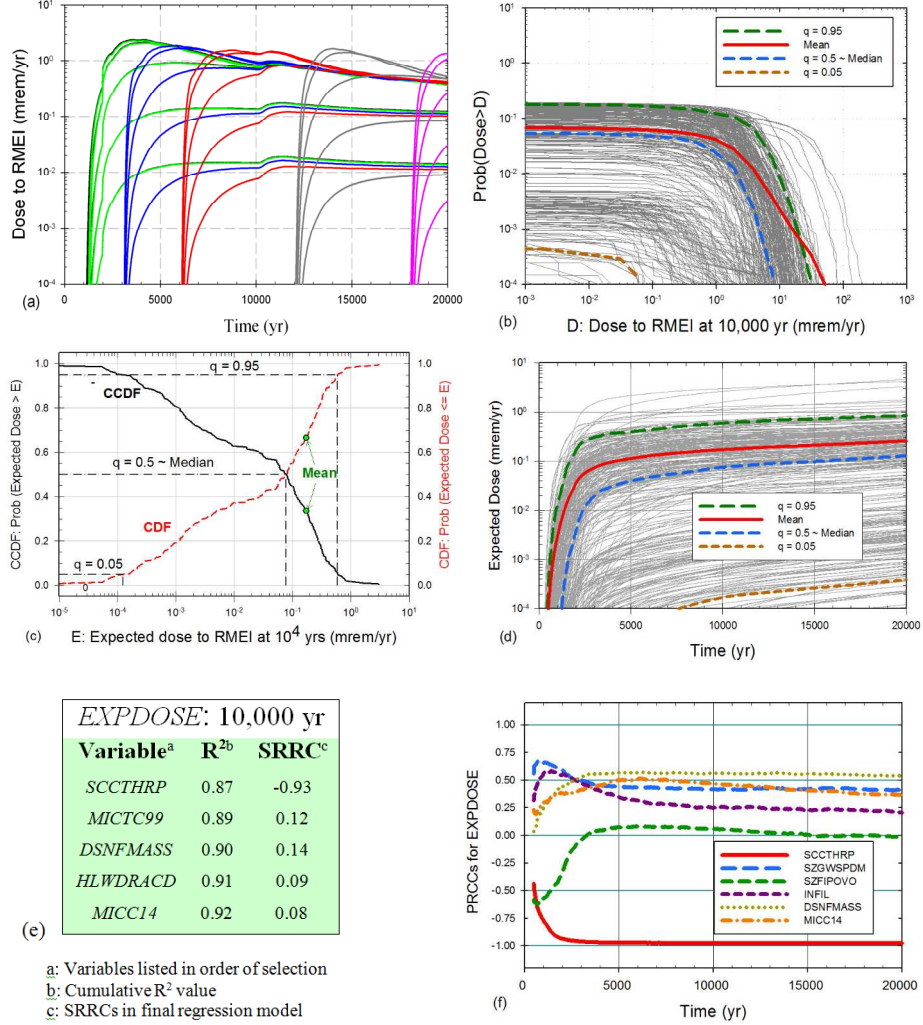
For the seismic ground motion scenario class and the time interval [0, 2×10^4 yr], the needed computational efficiency was achieved by evaluating $D_{SG}(t|\mathbf{a}, \mathbf{e}_{Mk})$ at a sequence of times (i.e., 100, 1000, 3000, 6000, 12000, 18000 yrs) and for a sequence of damaged areas areas (i.e., $10^{-8+s}(32.6 \text{ m}^2)$ for $s = 1, 2, …, 5$ with 32.6 m^2 corresponding to the surface area of a WP) at each time (Fig. 2a). This required $6 \times 5 = 30$ evaluations of the system indicated in Fig. 1 for each LHS element in Eq.(23). Once obtained, these evaluations can be used with appropriate interpolation and additive procedures to evaluate $D_{SG}(t|\mathbf{a}, \mathbf{e}_{Mk})$ for different values of $\mathbf{a}$ for each LHS element $\mathbf{e}_k = [\mathbf{e}_{Ak}, \mathbf{e}_{Mk}]$.

The individual CCDFs in Fig. 2b are defined by probabilities of the form shown in Eq. (7) with (i) $D_{SG}(t|\mathbf{a}, \mathbf{e}_{Mk})$ and $\mathbf{e}_{Ak}$ replacing $y(t|\mathbf{a}, \mathbf{e}_M)$ and $\mathbf{e}_A$ and (ii) $t = 10^4$ yr. Numerically, the integrals that define exceedance probabilities for the individual CCDFs are approximated with (i) random sampling from the possible values for $\mathbf{a}$ as indicated in Eq. (5) and (ii) estimated values $\hat{D}_{SG}(t|\mathbf{a}_j, \mathbf{e}_{Mk})$ for $D_{SG}(t|\mathbf{a}, \mathbf{e}_{Mk})$ constructed from results of the form shown in Fig. 2a. Specifically,

$$\hat{p}_A\left[y < D_{SG}\left(t|\mathbf{a}, \mathbf{e}_{Mk}\right)|\mathbf{e}_{Ak}\right] = \sum_{j=1}^{nS} \overline{\delta}_y\left[\hat{D}_{SG}\left(t|\mathbf{a}_j, \mathbf{e}_{Mk}\right)\right]/nS, \tag{26}$$

with the $\mathbf{a}_j$, $j = 1, 2, …, nS$, sampled in consistency with the density function $d_A(\mathbf{a}|\mathbf{e}_{Ak})$ for vectors of the form shown in Eq. (25). The mean and quantile curves in Fig. 2b are (i) defined and approximated as indicated in Eqs. (17) and (18) and (ii) provide a summary of the epistemic uncertainty present in the estimation of exceedance probabilities (i.e., $p_A[y < D_{SG}(10^4|\mathbf{a}, \mathbf{e}_M)|\mathbf{e}_A]$) for $D_{SG}(10^4|\mathbf{a}, \mathbf{e}_M)$.

As indicated in Eq. (5), the expected value $E_A[D_{SG}(t|\mathbf{a}, \mathbf{e}_{Mk})|\mathbf{e}_{Ak}]$ of $D_{SG}(t|\mathbf{a}, \mathbf{e}_{Mk})$ over aleatory uncertainty can also be defined and estimated, with the estimate $\hat{E}_A[D_{SG}(t|\mathbf{a}, \mathbf{e}_{Mk})|\mathbf{e}_{Ak}]$ obtained as shown in Eq. (26) with removal of the indicator function $\overline{\delta}_y$. The expected values $E_A[D_{SG}(t|\mathbf{a}, \mathbf{e}_{Mk})|\mathbf{e}_{Ak}]$ and their corresponding esti-

$EXPDOSE$: 10,000 yr		
Variable[a]	$\mathbf{R^2}$[b]	**SRRC**[c]
SCCTHRP	0.87	-0.93
MICTC99	0.89	0.12
DSNFMASS	0.90	0.14
HLWDRACD	0.91	0.09
MICC14	0.92	0.08

a: Variables listed in order of selection
b: Cumulative R^2 value
c: SRRCs in final regression model

Fig. 2. Example results for dose (mrem/yr) to RMEI for seismic ground motion scenario class: (a) dose for seismic events occurring at different times and causing different damaged areas on WPs ([1], Fig. J8.3-3a), (b) CCDFs for dose at 10,000 yr ([1], Fig. J8.3-10a), (c) CCDF for expected dose at 10,000 yr ([1], Fig. J8.3-5c), (d) time-dependent expected dose ([1], Fig. J8.3-6), (e) stepwise rank regression for expected dose at 10,000 yr ([1], Fig. K7.7.1-2a), and (f) time-dependent PRCCs for expected dose ([1], Fig. K7.7.1-1c).

mates are the result of reducing each CCDF in Fig. 2b to a single number. As indicated in Eqs. (14)-(16) and illustrated in Fig. 2c, epistemic uncertainty associated $E_A[D_{SG}(t|\mathbf{a}, \mathbf{e}_M)|\mathbf{e}_A]$ can be summarized by (i) an expected (mean) value $E_E\{E_A[D_{SG}(t|\mathbf{a}, \mathbf{e}_M)|\mathbf{e}_A]\}$ over epistemic uncertainty as defined in Eq. (14) , (ii) a CDF as defined by the cumulative probabilities in Eq. (15), or (iii) a CCDF as defined by the complementary cumulative probabilities in Eq. (16). The indicated mean defined in Eq.
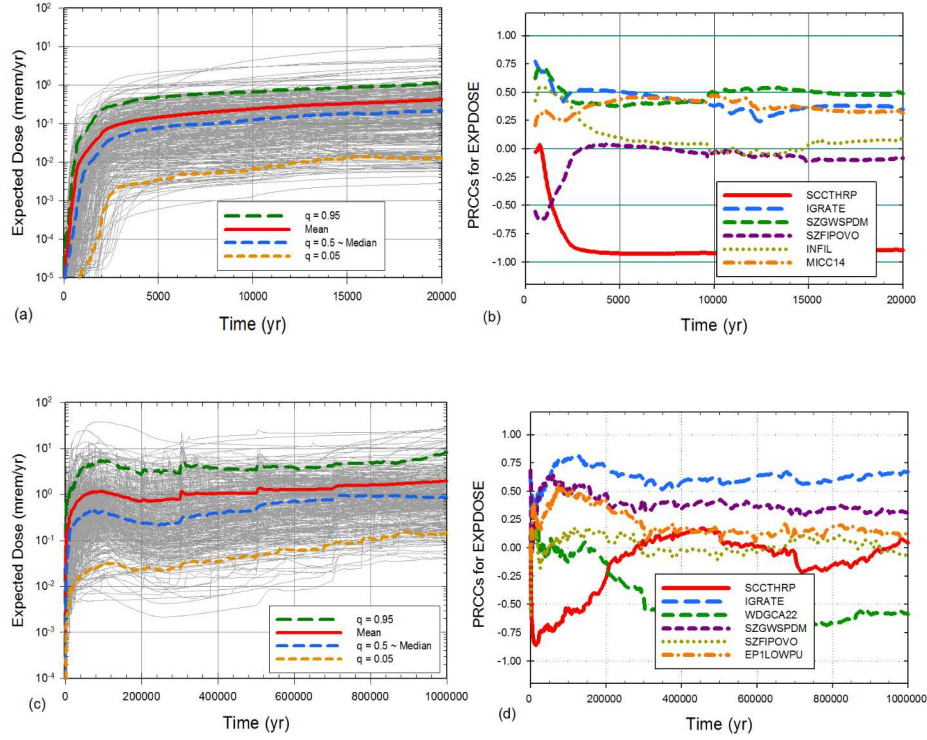
(14) and illustrated in Fig. 2c is the outcome of reducing all the information in Fig. 2b to a single number. The approximation process for a CDF also provides the basis for obtaining specific quantile values (e.g., q = 0.05, 0.5 ~ median, 0.95) as indicated in Fig. 2c.

In the NRC's regulatory requirements for the YM repository, bounds apply over time to expected dose to the RMEI. Thus, the analysis results of greatest interest are expected dose and the uncertainty in expected dose as a function of time (Fig. 2d). Specifically, expected doses for individual LHS elements correspond to the lighter lines in Fig. 2d, and quantile and mean values for expected dose that summarize the effects of epistemic uncertainty correspond to the darker dashed and solid lines. The results on Fig. 2d at 10,000 years correspond to the results shown in more detail in Fig. 2c. For reasons of computational efficiency, the individual expected dose curves in Fig. 2d were estimated with a quadrature procedure as described in Sect. J8.3 of Ref. [1] rather than with a sampling-based procedure as illustrated in Fig. 2b.

Sensitivity analysis is an important component of the 2008 YM PA and contributes to an establishment of "reasonable expectation" by supporting a detailed examination of the operation of the models that predict dose and expected dose to the RMEI and many other analysis outcomes of interest. Specifically, sensitivity analysis in the 2008 YM PA was based on an exploration of the mapping between elements of the LHS indicated in Eq. (23) and analysis results of interest (e.g., dose and expected dose to the RMEI) with a variety techniques including stepwise rank regression (Fig. 2e) and time-dependent partial rank correlation coefficients (PRCCs) (Fig. 2f) [19]. In stepwise rank regression, variable importance is indicated by the order in which variables are selected in the stepwise process, the incremental changes in R^2 values as variables are added to the regression model, and the values of the standardized rank regression coefficients (SRRCs) in the regression model. The indicated results with R^2 values and SRRCs provide a measure of the amount of epistemic uncertainty in the dependent variable under consideration that derives from the epistemic uncertainty in individual analysis inputs (i.e., elements of **e**); in contrast, PRCCs provide a measure of the strength of the monotonic relationship between individual epistemically uncertain analysis inputs and the dependent variable under consideration after removal of the monotonic effects of all other epistemically uncertain analysis inputs. Definitions for selected variables appearing Figs. 2e and 2f are given after Eq. (20).

Additional detailed information on uncertainty and sensitivity analysis for the seismic ground motion scenario class in the 2008 YM PA is available in Sects. J8 and K7 of Ref. [1]. Included in this material are extensive examples of results conditional on individual realizations of aleatory uncertainty as defined in Eqs. (11)-(13).

As indicated in Eq. (22), expected dose results are obtained for individual scenario classes and then summed to determine expected dose over all scenario classes (i.e., over all futures as described by the vectors defined in Eq. (19)). The results of this summation are shown in Figs. 3a and 3c for the time intervals [0, 2×10^4 yr] and [0, 10^6 yr], respectively. The individual expected dose curves in Fig. 3c are not as smooth as the individual expected dose curves in Fig. 3a. This difference results because the complexity of the calculations for the [0, 10^6 yr] time interval required the

Fig. 3. Expected dose to RMEI: (a, b) Expected dose and associated PRCCs for $[0, 2\times10^4$ yr] ([1], Fig. K8.1-1[a]) and (c, d) Expected dose and associated PRCCs for $[0, 10^6$ yr] ([1], Fig. K8.2-1[a]).

use of a sampling-based procedure to approximate expected dose from seismic ground motion events (see [1], Sect. J8.4); in contrast, the quadrature procedures used to determine expected dose for the individual scenarios for the $[0, 2\times10^4$ yr] time interval resulted in the smoother expected dose curves in Fig. 3a.

Ultimately, the NRC decided that their expected (i.e., "arithmetic mean") dose requirements of 15 mrem/yr and 100 mrem/yr for the time intervals $[0, 10^4$ yr] and $[10^4, 10^6$ yr], respectively, applied to the mean (i.e., solid line) doses in Figs. 3a and 3c (i.e., to results of the form defined in Eq. (14)). However, it is the spread of the individual dose curves in Figs. 3a and 3c that provide the NRC requested uncertainty information and thus a basis for a "reasonable expectation" that the requirements are being met despite the presence of substantial epistemic uncertainty. Further, sensitivity analyses of the form illustrated in Figs. 3b and 3d also contribute to "reasonable expectation" by enhancing understanding of the overall analysis and thus confidence in the numerical implementation of the analysis. Definitions for selected variables appearing Figs. 3b and 3d are given after Eq. (20).

Additional detailed information on uncertainty and sensitivity analysis in the 2008 YM PA is available in Apps. J and K of Ref. [1].

15

5   Concluding Message

Everyone cannot be expected to agree on the details of an analysis of a complex system, but everyone should be able to know what those details are. Without such knowledge, it is not possible to have informed and meaningful discussions involving the views of all parties interested in the analysis.

A necessary starting point for the design, computational implementation and, ultimately, communication of an analysis for a complex system is a clear conceptual structure. As described in this presentation, this structure for many, if not most, analyses, can be based on three basic entities: a probability space (A, A, p_A) characterizing aleatory uncertainty, a probability space (E, E, p_E) characterizing epistemic uncertainty, and a model that predicts system behavior (i.e., a function $f(t|\mathbf{a}, \mathbf{e})$, or more typically a vector function $\mathbf{f}(t|\mathbf{a}, \mathbf{e})$, that defines system behavior at time t conditional on elements $\mathbf{a}$ and $\mathbf{e}$ of the sample spaces A and E for aleatory and epistemic uncertainty). As illustrated with results from the 2008 YM PA, this conceptual view of the analysis of a complex system provides (i) a way to formally describe that analysis and then (ii) a clear path from formal description to computational implementation to written documentation.

References

1. SNL (Sandia National Laboratories). *Total System Performance Assessment Model/Analysis for the License Application.*MDL-WIS-PA-000005 Rev 00, AD 01.Las Vegas, NV: U.S. Department of Energy Office of Civilian Radioactive Waste Management 2008.
2. U.S. DOE (U.S. Department of Energy). *Yucca Mountain Repository Safety Analysis Report.*DOE/RW-0573.Las Vegas, NV: U.S. Department of Energy 2008.
3. U.S. DOE (U. S. Department of Energy). *Yucca Mountain Repository License Application.*DOE/RW-0573, Rev. 0.Las Vegas, NV: U. S. Department of Energy 2008.
4. Helton JC, Sallaberry CJ. Conceptual Basis for the Definition and Calculation of Expected Dose in Performance Assessments for the Proposed High-Level Radioactive Waste Repository at Yucca Mountain, Nevada. *Reliability Engineering and System Safety* 2009;94(3):677-698.
5. Helton JC, Sallaberry CJ. Computational Implementation of Sampling-Based Approaches to the Calculation of Expected Dose in Performance Assessments for the Proposed High-

Level Radioactive Waste Repository at Yucca Mountain, Nevada. *Reliability Engineering and System Safety* 2009;94(3):699-721.

6. U.S. NRC (U.S. Nuclear Regulatory Commission). 10 CFR Part 63: Implementation of a Dose Standard after 10,000 Years. *Federal Register* 2009;74(48):10811-10830.

7. U.S. NRC (U.S. Nuclear Regulatory Commission). 10 CFR Parts 2, 19, 20, etc.: Disposal of High-Level Radioactive Wastes in a Proposed Geologic Repository at Yucca Mountain, Nevada; Final Rule. *Federal Register* 2001;66(213):55732-55816.

8. Helton JC, Burmaster DE. Guest Editorial: Treatment of Aleatory and Epistemic Uncertainty in Performance Assessments for Complex Systems. *Reliability Engineering and System Safety* 1996;54(2-3):91-94.

9. Paté-Cornell ME. Uncertainties in Risk Analysis: Six Levels of Treatment. *Reliability Engineering and System Safety* 1996;54(2-3):95-111.

10. Helton JC, Breeding RJ. Calculation of Reactor Accident Safety Goals. *Reliability Engineering and System Safety* 1993;39(2):129-158.

11. Helton JC, Anderson DR, Marietta MG, Rechard RP. Performance Assessment for the Waste Isolation Pilot Plant: From Regulation to Calculation for 40 CFR 191.13. *Operations Research* 1997;45(2):157-177.

12. Helton JC, Anderson DR, Jow H-N, Marietta MG, Basabilvazo G. Conceptual Structure of the 1996 Performance Assessment for the Waste Isolation Pilot Plant. *Reliability Engineering and System Safety* 2000;69(1-3):151-165.

13. NAS/NRC (National Academy of Science/National Research Council). Evaluation of Quantification of Margins and Uncertainties for Assessing and Certifying the Reliability of the Nuclear Stockpile. Washington, DC: National Academy Press, 2008.

14. Helton JC, Pilch M. Guest Editorial: Quantification of Margins and Uncertainties. *Reliability Engineering and System Safety* 2011;96(9):959-964.

15. Helton JC. Quantification of Margins and Uncertainties: Conceptual and Computational Basis. *Reliability Engineering and System Safety* 2011;96(9):976-1013.

16. Hora SC, Iman RL. Expert Opinion in Risk Analysis: The NUREG-1150 Methodology. *Nuclear Science and Engineering* 1989;102(4):323-331.

17. McKay MD, Beckman RJ, Conover WJ. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* 1979;21(2):239-245.

18. Helton JC, Davis FJ. Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems. *Reliability Engineering and System Safety* 2003;81(1):23-69.

19. Helton JC, Johnson JD, Sallaberry CJ, Storlie CB. Survey of Sampling-Based Methods for Uncertainty and Sensitivity Analysis. *Reliability Engineering and System Safety* 2006;91(10-11):1175-1209.

20. Storlie CB, Helton JC. Multiple Predictor Smoothing Methods for Sensitivity Analysis: Description of Techniques. *Reliability Engineering and System Safety* 2008;93(1):28-54.

21. Storlie CB, Swiler LP, Helton JC, Sallaberry CJ. Implementation and Evaluation of Nonparametric Regression Procedures for Sensitivity Analysis of Computationally Demanding Models. *Reliability Engineering & System Safety* 2009;94(11):1735-1763.