*Journal of the Korean Nuclear Society*
*Volume 35, Number 1, pp.64~79, February, 2003*

# The Explicit Treatment of Model Uncertainties in the Presence of Aleatory and Epistemic Parameter Uncertainties in Risk and Reliability Analysis

**Kwang-Il Ahn and Joon-Eon Yang**
Korea Atomic Energy Research Institute
150 Dukjin-dong, Yuseung-gu, Daejeon 305-353, Korea
kiahn@kaeri.re.kr

(Received November 5, 2002)

**Abstract**

In the risk and reliability analysis of complex technological systems, the primary concern of formal uncertainty analysis is to understand why uncertainties arise, and to evaluate how they impact the results of the analysis. In recent times, many of the uncertainty analyses have focused on parameters of the risk and reliability analysis models, whose values are uncertain in an aleatory or an epistemic way. As the field of parametric uncertainty analysis matures, however, more attention is being paid to the explicit treatment of uncertainties that are addressed in the predictive model itself as well as the accuracy of the predictive model. The essential steps for evaluating impacts of these model uncertainties in the presence of parameter uncertainties are to determine rigorously various sources of uncertainties to be addressed in an underlying model itself and in turn model parameters, based on our state-of-knowledge and relevant evidence. Answering clearly the question of how to characterize and treat explicitly the forgoing different sources of uncertainty is particularly important for practical aspects such as risk and reliability optimization of systems as well as more transparent risk information and decision-making under various uncertainties. The main purpose of this paper is to provide practical guidance for quantitatively treating various model uncertainties that would often be encountered in the risk and reliability modeling process of complex technological systems.

**Key Words** : uncertainty analysis, aleatory and epistemic, parameter and model, formal treatment

## I. Introduction

In the risk and reliability analysis of complex technological systems like nuclear power plants, quantitative uncertainty analysis is an essential part for formal decision-making, and a primary step for the analysis is to understand why uncertainties arise, to identify what types of uncertainty are addressed in the modeling process, and to evaluate how they impact on the results of the analysis.

The risk and reliability modeling process for the potential behavior of a given system requires an appropriate decomposition of the behavior so that available historical and experimental data can be used to help determine the more understandable and manageable inputs. As a trade-off, however, the decomposition leads to a variety of events (or variables) about which we are uncertain in an aleatory or an epistemic way even for a best estimate model. The former type of uncertainty is that addressed when the events or phenomena being taken into account in the predictive model are characterized as occurring in an aleatory manner. Whereas, the latter type of uncertainty is referred as state-of-knowledge uncertainty and it is closely associated with the analyst' s confidence of the actual situation being predicted through the model and the accuracy of the predictive model itself. Because uncertainties addressed in the foregoing aleatory and epistemic events (or variables) are generally characterized and treated differently when creating models of complex systems, it is useful to identify the two classes of uncertainty that are addressed in and impact the results of model predictions.

In a formal analysis of uncertainty, the foregoing rigorous classification of uncertainty sources [1-6] is first related to important practical aspects of modeling for complex technological systems and the definition of analysis model, rather than difference in the concept of uncertainty. The definition or target of analysis model determines the relevancy of available data and whether or not it is necessary to distinguish between uncertainty statements about aleatory variability among the occurrences of individuals in the population (i.e., what we know) and uncertainty due to lack of knowledge about fixed but unknown quantities (i.e., how much we know about it). Another important aspect for exploring different types of uncertainty allows for a proper propagation of different uncertainties in the evaluation process so that consistent decision-making is made for the resulting quantitative uncertainties. When both aleatory and epistemic uncertainties are already mixed up in the course of the analysis without a clear separation, it would not be possible to identify the resulting combined effect of the uncertainties of either type. The last aspect of the formal separation is that the approach is very helpful in understanding the nature of the uncertainties and for estimation of uncertainty measures in practical situations. Many risk and reliability analysis practitioners would often overlook these various facets of uncertainty in their analysis and as a result they would be confused in interpreting the results of uncertainty analysis. Thus, answering clearly the question of how to characterize and treat explicitly the forgoing different types of uncertainty is particularly important for practical aspects such as risk and reliability optimization of systems as well as more transparent risk information and decision-making under various uncertainties.

As for whether we can ever gain precise, accurate and complete knowledge about physical problems, on the other hand, one may eliminate all the uncertainties associated with modeling that problem. In a real situation, however, this is neither possible nor practical since we do not know all complete conditions. Associating with modeling uncertainties particularly, many analysts have taken into account them implicitly through an aggregation of expert opinions, and the uncertainty distributions obtained in such a way, however, would not give a full spectrum of uncertainties explicitly. As the field of parametric uncertainty analysis matures, more attention is being paid to an explicit treatment of uncertainties that are addressed in the predictive model itself as well as the accuracy of the predictive model because it is important to develop an understanding of the impact of different modeling assumptions on the quantitative results of the predictive model. The explicit incorporation of model uncertainty into the framework of uncertainty [7-14] does add another layer of uncertainty to the problem of interest. The main purpose of this paper is to provide formal characterization for various types of uncertainties that would often be encountered in the risk and reliability modeling process of complex technological systems and practical guidance for quantitatively treating model uncertainties that are expressed in various forms.

## 2. Charcterization of Uncertainty Sources

Before tackling the explicit treatment of model uncertainties in the presence of aleatory or epistemic parameter uncertainties that would be addressed in the underlying risk analysis model (characterized as either a logical or physical model), it is an essential step to characterize these various types of uncertainties, based on the states of knowledge involved in the problem of interest and available evidence.

### 2.1. Aleatory and Epistemic Uncertainty Sources

#### 2.1.1. Clarification of Uncertainty Sources

Many risk and reliability analyses would often include both aleatory events and epistemic variables in their predictive models. However, discussions in the literatures have often been unclear with respect to this distinction of underlying uncertainty sources and types. While the aleatory uncertainty has been conventionally regarded as a property of the system or activity being studied, the epistemic uncertainty takes into account subjective and perceptional aspects. According to Hofer' s definition [3], the aleatory uncertainty arises from the fact that one cannot give a single value for an event, but rather give a population of values with chance. Thus, the value of the event can be thought of as randomly selected from the single true probability distribution that summarizes the variability within the population. Whereas, the epistemic uncertainty is characterized as uncertainty due to knowledge of the single true values of an event and lack of knowledge of a single exact probability distribution summarizing the variability within a population. Since the aforementioned types of uncertainties are inputs to different decisions, then their propagation through the model needs to happen separately and presentation has to cater for two uncertainty dimensions. Helton [4] also claims: 'When a distinction between stochastic and subjective uncertainty is not maintained, the likelihood of the deleterious events associated with a system and the confidence with which both likelihood and consequences can be estimated become mingled in a way that makes it difficult to draw useful insights.' Whereas, Winkler [15] takes another viewpoint about those uncertainties: 'If the problem is not decomposed in a reasonable way, various sources of uncertainties can be commingled in a way that makes it difficult to draw useful insights.' While the first two standpoints focus on the need of separation for 'different types of uncertainty' regardless of the decomposition level, the last one relates it to 'different sources of uncertain information' due to the decomposition as a key motivating factor behind the desire to distinguish among types of uncertainty, i.e., the motivation for attempts to make rigorous distinctions between both uncertainties that is related to important modeling concerns.

Even though all of the above viewpoints about uncertainty focus their own aspects in the modeling process of complex technological systems, it should be noted that the foregoing rigorous classification between aleatory and epistemic uncertainties is basically for our convenience, rather than for conceptual difference. At a fundamental level of detail, uncertainty is just uncertainty and there is only one kind of uncertainty stemming from our lack of knowledge concerning the problem of interest [5,15-17]. For example, all events (effects) follow physical laws and they are completely determined once the initial and boundary conditions (causes) are specified. This means that in a situation where all complete conditions are known, a random event is no more defined. As for whether we can ever gain precise, accurate and complete knowledge about physical problems one may eliminate all the uncertainties associated with modeling that problem. In a real situation, however, this is neither possible nor practical since we do not know all complete conditions. According to their sources, the separation of uncertainties into different types is very helpful in investigating the behavior of such systems, selecting appropriate measure of uncertainty, and interpreting consistently the results of uncertainty analysis made under different sources. In light of uncertainty management, the aleatory portion of uncertainty is not practically reducible under a given proposition of event since we don' t know and understand the underlying reasons and behaviors governing its randomness. On the contrary, as we know more about the underlying problem, the epistemic portion of uncertainty can be effectively reduced. From this point of view, the analysis of aleatory uncertainty is to answer the question on 'what might actually happen and with what probability' . Whereas, the analysis of epistemic uncertainty is to answer the question on 'how well we know a given problem and how much our knowledge about it might change with additional information' .

#### 2.1.2. Guidance for Estimation of Both Uncertainties

A decision for taking either aleatory or epistemic variables depends on the states of knowledge involved in the problem of interest. According to Winkler' s viewpoint [5], some uncertainties are clearly easier to assess than others although information comes in varying forms and from many sources, involving historical and experimental data, models, or experts. In practical situations, however, it is not easy to dichotomize which event is an aleatory event or deterministic one. The difference between aleatory and epistemic treatments depends on two things: (a) our knowledge of the laws governing the occurrence of the event and (b) the sensitivity of the event to small changes in initial conditions. The occurrence of a random event is so sensitive to small changes in initial conditions that it is practically impossible to predict it by means of a deterministic formula. If a statistical regularity is expected, the event can be treated as a random event. Though the outcome of a deterministic event may also be sensitive to initial conditions and in many cases the outcome obtained from predictive formulas is by no means precisely equal to the actual outcome. However, we are willing to tolerate the variations, due to deterministic regularity that a given set of circumstances should lead to the same outcome. As mentioned previously, a decision whether an event is treated to be either aleatory or epistemic is subjectively made by evaluating our state of knowledge, and creating appropriate model characterizing a statistical or deterministic regularity. The epistemic uncertainty requires specification of inaccuracy in the initial and boundary conditions and the

**Table 1. Characterization of Aleatory and Epistemic Uncertainties**

| Items | Aleatory Portion | Epistemic Portion |
|---|---|---|
| Classification crieria | Level of modeling details<br>Knowledge about the laws governing the occurrence of an event<br>Degree of the sensitivity to the initial conditions or environment | |
| Terminologies | Random/Stochastic, Irreducible, Observable, Inherent Uncertainty | State-of-knowledge, Reducible, Unobservable, Cognitive Uncertainty |
| Uncertainty sources | Randomness/variability of an event, Circumstance variability | Inaccurate knowledge of a fixed quantity Alternative representations of a true but unknown value |
| Measure | Probability model of random variability or circumstance variability among unspecified values in the population | Subjective probability model of a fixed but unknown quantity or different distribution or model assumptions |
| Uncertainty management | We don' t know and understand the underlying reasons and behaviors governing randomness, so it is not practically reducible | As we know more about the underlying problem, it can be effectively reduced |
| Analysis purpose | Answer the question on what might actually happen and with what probability | Answer the question on how well we know and how much our knowledge about it might change with additional information |

pertinent physical laws. In general, the aleatory uncertainty requires a description of the random events and specification of the pertinent probabilistic model (i.e., probability density function). However, the epistemic portion of uncertainty is also present over the distributional parameters if the true distributional parameters (e.g., mean and variance) and shape of the distribution are unknown. Then, each value of the distributional parameters (mean and variance) specifies one aleatory distribution for the random event. A typical example where both uncertainties are specified even for an event, is probability distributions for failure rates of components as performed in the analysis of nuclear power plant systems. For practical convenience in estimating both aleatory and epistemic parameter uncertainties, Table 1 gives their summarized feature.

There are well-known examples where an event is treated as either an aleatory or epistemic event. As one example, let' s consider coin tossing: one is a two-sided coin and another is a one-sided coin hidden inside a box. In the former case, two possible events (i.e., face-up or face-down) exist and the occurrence of each event is so sensitive to initial conditions that are imposed to the coin-tossing event. In that case, we normally treat it as an aleatory event. In the latter case, only one fixed-event (i.e., face-up or face-down) exists, but we have no information about whether or not the coin is face-up. In that case, each of these events is treated as an epistemic event. As another example, let' s consider a severe accident sequence with limited resolution or unspecified in many ways. Then, various phenomenological variables contributing to the containment peak pressure (e.g., in-vessel steam explosion, core melt temperature at the occurrence of all event sequences, etc.) are regarded as aleatory variables whose uncertainties are quantified by a probability distribution summarizing the variability within the potential population of relevant values. In that case, there may be two situations by which the aleatory variables may be treated by epistemic uncertainties: The first is when we redefine the above sequences with much more resolution so that the aforementioned aleatory variables can be subjected to epistemic uncertainties with a fixed, specified accident sequence. Then, this situation explains that possible stochastic variation for all accident sequences is considered to be comparatively negligible in specific aspects of the phenomenological assessment when the specified accident event is applied. In that case, the aforementioned variables can be treated as deterministic quantities with inaccurately known epistemic uncertainties.

### 2.2. Model Uncertainty Sources

In the modeling process of behavior of a given system, we are often faced with two different situations: one is that model input values can be deterministically described and another we are simply unable to predict their values in a deterministic way. In the former case, the model output is characterized by magnitudes of the deterministic inputs and in the probabilistic sense the model outcome becomes always one. Like this, a model whose outcome is determined just by the deterministic inputs is characterized as 'deterministic model'. The deterministic model can be expressed as a simple functional relationship whose output depends on a *deterministic manner on various input parameters.* In many cases, however, we do not have a complete knowledge about the exact values of input parameters, and consequently, some imprecision attaches to the estimate of the model output. Uncertainty about the correct values of input parameters can be quantified by treating subjective probability distributions. On the other hand, there is a situation where the model output is determined by both the occurrence of the event (i.e., aleatory uncertainties) and its magnitude (i.e., epistemic uncertainties). While the model itself is given to be deterministic, different model outcomes occur at random. Since this type of model contains probabilities on the occurrence of the model outcome whose values are evaluated by an aleatory model, it is referred to as an 'aleatory model'. In order to characterize the aleatory model output we would often employ a probabilistic model for the occurrence of the event, and epistemic uncertainty about the magnitude of the model parameters once the model output occurs given that the event has occurred. In the deterministic model, our concern is to determine a specific criterion of the model outcome and in turn to evaluate the magnitude of the model parameter. Whereas, our primary concern in the aleatory model is to estimate an occurrence probability of random event addressed in the aleatory model, with/without magnitude of the parameter value characterized by the event.

When we develop a model describing the behavior of s given system, on the other hand, we often have an additional question on the possibility of different structures as well as the accuracy of the model itself. This question is closely related to the fact that our model itself is an approximation about the true behavior of the system and its accuracy depends on our state-of-knowledge about the system behavior. While our modeling of the system behavior is based on currently available information, and the modeling process may be different from expert to expert. This fact leads us to conclude that our model itself is always subjected to some degree of uncertainty, due to (a) expert-to-expert different assumptions utilized when creating the model structure (whose uncertainty is given in the form of weighting factors, or upper and lower limits), (b) analyst' s subjective confidence attached to the model prediction (whose uncertainty is given in the form of epistemic multiplier), and (c) circumstance-to-circumstance variability attached to the model prediction (whose uncertainty is given in the form of aleatory multiplier). While the former two cases can be treated as a type of epistemic uncertainties on the predictive model, the last case is normally treated a type of aleatory uncertainties. If there are no uncertainties in our model itself and relevant model parameter values, the model will exactly predict the real problem with specification of model parameter values. Due to the model and parameter uncertainty, however, our model predictions are subjected to over- or under-estimation for the true behavior of the system in either an aleatory or epistemic way. Finally, while the essential portion of model uncertainty is basically originated from such different modeling process, it may have different forms depending upon the level of analysis for the underlying system: high-level modeling uncertainties (different structures in a predictive model) or low-level modeling uncertainties (different sub-models characterizing parameters of the predictive model). From the aspect of viewpoint, the probability distribution model for an aleatory event can be regarded as a special form of low-level model because the distribution is subject to model parameters such as mean and variance.

## 3. Explict Treatment of Model Uncertainties

As given in Figure 1, the explicit treatment of modeling uncertainties does add another layer of uncertainty to the problem of interest. In real applications, two different situations are often faced: one situation is when there is only a single deterministic model but a variety of relevant actual data is available so that the accuracy of the model can be assessed, and another situation is when various alternative models are available, regardless of evidence available. The former case is useful to some situations that actual data of interest is obtained from various circumstances. The latter case analyzes the impact of using alternative modeling assumptions by performing an appropriate statistical propagation of subjective probabilities assigned to different models or model sensitivity studies.

### 3.1. When There is a Single Deterministic Model, But Relevant Actual Data is Available

A population of different circumstances that lead to different values for each output calculated by a single deterministic model of interest can be characterized by aleatory uncertainty. Also, a particular circumstance of interest where the same model parameter values can be applied to the single deterministic model will be one of many potential circumstances. Let' s consider a situation that actual data of a quantity of interest can be obtained from different circumstances and there are several actual values that corresponds to a given model prediction (or the same model parameter values), but we do not know the actual circumstances. From another point of view, some variability in the actual data that corresponds to the same parameter values of the single deterministic model can be accounted for as (epistemic) model uncertainty in the predictions of the model. The uncertainty is due to the approximation made to develop the model. If there are no model uncertainties, the model will exactly predict actual data. In order to account for such uncertainties in the single model predictions of interest, Siu and/or Apostolakis [5, 7], employ a concept of 'a multiplicative factor (E)' that modifies the predictions of the model with actual data obtained from different circumstances.

$$\varphi_r = E \cdot \varphi_{drm}(\vec{\pi}) \quad (1)$$

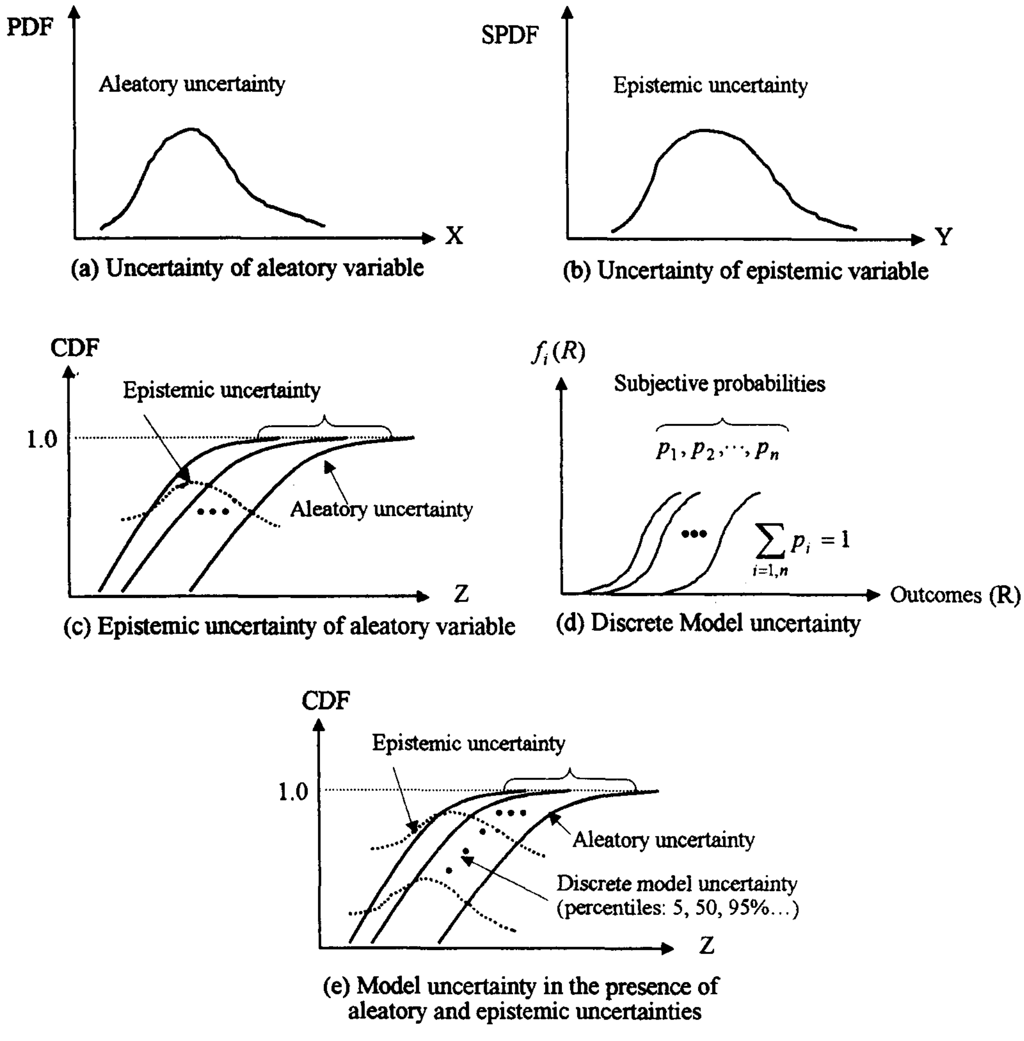
where

$\varphi_{drm}$ = single deterministic model (prediction model),

$\varphi_r$ = unknown real model predicting various circumstances,

$\vec{\pi}$ = model parameter vector characterizing $\varphi_{drm}$,

$E$ = multiplicative factor modifying the predictions of $\varphi_{drm}$ with actual data.




Note: (S)PDF: (subjective) probability density function, CDF: cumulative distribution function
$p_i$ : subjective probability for $i$ -th model, R: output variable, $f_i(R)$ : model prediction for R

**Fig. 1. Impacts of Model Uncertainties in the Presence of Various Parametric Uncertainties**

According to their definition the factor is given as the ratio of the actual value over the calculated value by a deterministic single model for the specific quantity. The recognition of the aforementioned circumstance variability leads us naturally to the modeling of the aleatory and epistemic uncertainties. Then, the relative frequency of the actual data values with the same calculated values of a specific quantity determines an aleatory distribution that the specific quantity will have the corresponding values. The lack of information about a true form of aleatory distribution leads to the presence of epistemic uncertainty over the distributional parameters (e.g., mean and variance).

### 3.2. When Various Alternative Models are Available, Regardless of Available Evidence

For a given situation, two different approaches are possible for quantifying the impacts of alternative models to the prediction of a predictive model of interest (characterized as a high level or target model): one approach is to utilize a statistical integration of competing models of interest through the predictive model, and another approach is to perform a sensitivity analysis for each of the competing models. If model uncertainties are properly defined, both approaches may be implemented in a similar manner with a parametric approach

#### 3.2.1. Integration of Alternative Models

If there exist various competing models (given in the form of either different model structure or different model predictions) and they are assumed to be independent of each another (e.g., mutually exclusive and independent models), the model uncertainty can be characterized by assigning subjective probabilities (or relative weights) to each of these alternative models [9-11, 13-14]. Then, the subjective probability accounts for the relative importance of alternative models so that models subject to the relatively high probability gives a greater impact when they are propagated through the target model. There are two approaches for propagating model uncertainties: one is to utilize Bayesian aggregation model and another is to use a statistical propagation method like Monte Carlo. The Bayesian approach [8, 18] can be applied to the situations where the model uncertainties are given in the level of model predictions rather than the model structure itself. On the contrary, the statistical propagation approach makes it possible to directly quantify uncertainties for alternative models themselves in the sampling stage as well as the values of uncertain model parameters. Typically three practices where model uncertainties can be explicitly quantified using the statistical integration would be encountered as given as follows:

Case 1: Deterministic models for phenomenological assessment, subject to epistemic model and parameter uncertainties

A deterministic model for the assessment of phenomenological behavior itself that are typically formulated by epistemic model parameters whose uncertainties is characterized as epistemic ones. In that case, all uncertainties we have to handle are epistemic uncertainties. Model uncertainties may stem from high-level models differently formulated by analysts (e.g., structural models) and low-level models resulting in different predictions of particular model parameters (e.g., physic-numerical correlations). Since in many cases model parameters of the high-level model determine the predictions of low-level models, these low-level models cannot be treated independently with the high-level model parameters. Once a set of the high-level parameter values are specified, the underlying different low-level models produce the corresponding uncertainty distributions. This fact indicates that the low-level models must be treated as models themselves rather than as the level of model predictions. With a situation that a high level model contains other lower level models, the mathematical formulation for handling those uncertainties can be expressed as follows,

$$Y(t) = w_m G_m(\vec{\pi}_m, \vec{\varphi}_m \mid t, D_m), \ \sum_m w_m = 1 \quad (2a)$$

$$\vec{\pi}_m \equiv \{\pi_{m,1}, \pi_{m,2}, \cdots, \pi_{m,N_m^\pi}\} \quad (2b)$$

$$\vec{\varphi}_m \equiv \{\varphi_{m,1}, \varphi_{m,2}, \cdots, \varphi_{m,N_m^\varphi}\} \quad (2c)$$

$$\varphi_{m,j} \equiv \sum_k \{\nu_{m,jk}\varphi_{m,jk}\}, \ \sum_k \nu_{m,jk} = 1 \quad (2d)$$

where

- $Y(t)$ = time-dependent target variable (weighted over different model $G_m$),
- $G_m$, $w_m$ = m-*th* high-level model and the corresponding weighting factors,
- $\vec{\pi}_m$, $N_m^\pi$ = uncertain parameter vector for $G_m$ and number of relevant parameters,
- $\vec{\varphi}_m$, $N_m^\varphi$ = low-level uncertain model vector for $G_m$ and number of relevant models,
- $\varphi_{m,jk}$, $\upsilon_{m,jk}$ = k-*th* element for the j-*th* uncertain low-level model $\varphi_{m,j}$ and weighting factor,
- $D_m$ = entire body of data information characterizing the model $G_m$.

In real applications, the synthesis of epistemic modeling and parameter uncertainties can be made by either a single-step or two-step sampling procedure, and its choice depends on uncertainties of basic variables employed in the low-level models, or whether or not any differentiation between modeling and parameter sources of uncertainty is required in the final outcome of the target variable.

The single-step sampling and propagation can be enough when all low-level models are given as functions of fixed variables rather than uncertain variables so that every model provides a single prediction for given variable values or when epistemic parameter and model uncertainties are not required to be treated separately. The latter case basically follows the fact that both uncertainties are basically subject to epistemic uncertainties although they stem from different sources. According to a single-step sampling process, then a random sample of size n of sets of uncertain model parameters ($\vec{\pi}_m$) and uncertain low-level models ($\vec{\varphi}_m$) for simultaneous variations is selected according to the specified epistemic probability distributions. As the result of statistical propagation, each sample combination of parameter inputs and model inputs results in the corresponding single prediction. The final aggregation of the resulting n predictions for n sample vectors gives a single distribution for the target values.

When we consider a situation that uncertainties addressed in the high-level model parameters are separately expressed with the modeling sources of uncertainty, the statistical approach must be implemented with a two-stage propagation [5,18,19]: the first stage is taken for model uncertainties themselves ($\vec{\varphi}_m$) and the second stage is taken for uncertain parameters contained a target model ($\vec{\pi}_m$). By this approach, the two different sources of uncertainties are propagated separately through the target model. The resulting uncertainty distribution in the target model prediction becomes conditional on each of the first sample vectors, which is expressed as a family of uncertainty distributions that folds portions of modeling uncertainties and relevant parameter uncertainties. Of course, the aforementioned family of uncertainty distributions can be averaged over modeling sources of uncertainties so that a single epistemic uncertainty distribution is obtained at the stage of decision-making. In the presence of both the epistemic model and parameter uncertainties, practical examples using the statistical integration approach can be found in relevant literatures [11, 13].

Case 2: System reliability models, subject to aleatory, epistemic model and parameter uncertainties

A typical example for this case [12] is a fault tree analysis of system reliability that has just considered aleatory uncertainties of component failures. While the fault tree logic model itself is a deterministic model, its analysis allows for epistemic model and parameter sources of uncertainties. In the fault tree analysis, the epistemic portions of the model uncertainty may arise when modeling assumptions are made under our lack of precise knowledge about the system functionality, e.g., success criteria for the system performance that is generally determined by thermal-hydraulic analysis. A few applicable success criteria can be considered for more realistic evaluation of the system reliability and each of them generates the corresponding fault tree model. This fact leads naturally to a possibility of handling of the uncertainty in different possible success criteria and thus we can assign weighting factors to each alternative success criteria as done in the case 1. Then the weighting factors assigned to each success criterion are characterized as the epistemic uncertainty of the success criteria (i.e., uncertainty for high-level structural model). The aleatory uncertainty is accounted for in the probability models that are assigned to each of component-level basic events (such as binomial, Poisson, and exponential models). This can be classified into one type of low-level model uncertainty. Then, an additional type of epistemic model uncertainty is given as weighting factors to each of these competing probability models. Finally, epistemic parameter uncertainties are assigned to parameters of the underlying probability model (such as failure rate or demand probability). The most typical form of failure rate probability distribution is a lognormal distribution. Finally, a probabilistic combination of the failure rate distribution through the underlying probability model generates the corresponding lognormal distribution for each basic event. Now, the synthesis of the aforementioned aleatory (i.e., probability models) and epistemic modeling (i.e., alternative success criteria) and parameter (i.e., failure rate or demand probability distributions) sources of uncertainties can be made through minimal cutsets of the fault tree model. The mathematical formulation for handling those uncertainties can be expressed as follows,

$$Y = w_m G_m(\vec{\pi}_m \mid D_m), \ \sum_m w_m = 1 \qquad (3a)$$

$$\vec{\pi}_m \equiv \{\pi_{m,1}, \pi_{m,2}, \cdots, \pi_{m,N_m^\pi}\} \qquad (3b)$$

$$\pi_{m,j} \equiv \sum_k v_{m,jk}\varphi_{m,jk}, \ \sum_k v_{m,jk} = 1 \qquad (3c)$$

where

$Y$ = target variable (system reliability weighed over $G_m$),

$G_m$, $w_m$ = minimal cutset models for m-*th* success criteria and weighting factors,

$\vec{\pi}_m$, $N_m^\pi$ = basic event vector for $G_m$ (subject to aleatory probability models) and number,

$\varphi_{m,jk}, \upsilon_{m,jk}$ = k-*th* aleatory model for the j-*th* basic event of $\pi_m$ and weighting factor,

$D_m$ = entire body of generic or plant

specific information characterizing the model $G_m$.

In real applications, the underlying uncertainties given in the above formulation can be statistically integrated by either a single-step or two-step sampling procedure, and its choice depends on whether or not any differentiation between aleatory and epistemic sources of uncertainty is required in the final outcome of the target variable. The single-step sampling and propagation can be enough when we would like to evaluate both uncertainties in one, integrated manner. Then, all portions of aleatory uncertainties (i.e., failure probabilities for $\varphi_{m,jk}$) are sampled and they are propagated through the above weights-averaged minimal cutsets. As a result, we can obtain a single aleatory probability distribution for the system reliability. A practical example using the single-step sampling procedure can be found in the literature [14]. In order to assess explicitly the impacts of modeling source of uncertainty to the system reliability, on the other hand, the synthesis of uncertainties must be implemented with a two-stage propagation: the first step is taken for uncertainties addressed in the model themselves (structural model $G_m$ and probability model $\varphi_{m,jk}$) and the second step is taken for all portions of aleatory uncertainties (i.e., failure probabilities for $\varphi_{m,jk}$). The resulting uncertainty distribution in the target model prediction becomes conditional on each of the first sample vectors, which is expressed as a family of aleatory uncertainty distributions that folds both portions of modeling uncertainties ($G_m$, $\varphi_{m,jk}$) and relevant aleatory uncertainties (failure probabilities for $\varphi_{m,jk}$).

Case 3: Treatment of epistemic uncertainties addressed in an aleatory probability distribution

When the epistemic uncertainties are just addressed in parameters of an aleatory probability distribution such as mean and variance and there is no uncertainty in the high-level model structure, their simultaneous treatment has been widely taken into account in the field of plant system safety analysis. This is a special situation of the above case 2, which is often encountered in the system reliability analysis. In a system fault/event tree model, for example, aleatory uncertainty is accounted for by the probability distribution models that are assigned to each of component-level basic events as mentioned in the above case (2). When the underlying probability distributions are estimated using plant-specific and/or generic data, however, they do not represent an exact population of component failures. This leads naturally to the possibility of epistemic uncertainty (different means and variances, or distribution shape) for a true probability distribution (with true mean and variance). Thus, both aleatory and epistemic uncertainties are addressed in every basic event. In this case, we can view the probability distribution model either as a type of model or a simple parameter.

In order to assess explicitly the impacts of both aleatory and epistemic uncertainties on the system reliability, the synthesis of these uncertainties must be implemented with a two-step propagation: the first step is taken for probability models for basic events (from joint distribution for mean and variance or different shapes) and the second step is taken to sample failure probabilities from fixed probability distributions sampled from the first stage (with fixed mean and variance, shape). The resulting uncertainty distribution becomes conditional on each of the first sample vectors (i.e., statistical combination for sampled means and variances), which is expressed as a family of aleatory uncertainty distributions that folds both portions of epistemic uncertainties and relevant aleatory uncertainties. Typical examples taking into account the above situation can be found in relevant literatures [1, 19-20].

On the other hand, it should be noted that when the analysis of model uncertainties is implemented using the aforementioned statistical approach, two essential difficulties would be encountered especially for physical models. The first is that there is no generally accepted, robust approach for handling quantitatively the impacts of model uncertainties on the final prediction of the target model because models do not always have a simple, intuitively appealing interpretation. The second is that it is necessary to make an appropriate selection of a statistical combination of competing models in a reasonable way and use that combination to get some insight into what the uncertainties are. If all possible combinations are taken, we can get some strange physical situations because some are non-physical, even though the computer codes would allow them. This is, in part, because we do not really understand the processes that are occurring. In addition, the aforementioned integration of model uncertainties often obscures the differences among competing models and does not explain the reasons for the differences.

### 3.2.2. Model Sensitivity Analysis

When an analysis of the relative impacts of individual models to the final results is required, it is more preferable to retain separate models for each model sensitivity analysis. In that case, model sensitivity analysis must be made in a manner that parametric uncertainties are made conditionally upon each alternative model. The second situation that model sensitivity analysis may be particularly helpful is when differences in probability estimates assigned to individual models do not affect the final results very much or give little influence on conclusions. As another situation, it may be important for the decision maker to appreciate the degree of disagreement among the different models and its effect on the results. If the degree of sensitivity is very high, the aforementioned integration of model uncertainties should be avoided since they may tend to obscure critical differences of alternative models. Different assumptions for a probability distribution given in the level of model prediction can be regarded as a kind of model uncertainty (addressed in a true probability model). Then, the impact of each of the different assumptions for various model inputs (e.g., mean, variance, shape, etc.) on the output uncertainty would often be assessed using a distributional sensitivity analysis approach. References [11,13,21] provide typical methods to quantify the sensitivity of deterministic model uncertainties.

## 4. Summary and Conclusions

In this paper, we have provided formal characterization for various types of uncertainties that would often be encountered in the risk and reliability modeling process of complex technological systems (i.e., aleatory, epistemic parameter and model uncertainties), and derived practical guidance for quantitatively treating model uncertainties that are expressed in various forms. Associating with the formal treatment of model uncertainties, the present guidance can be summarized as follows:

### 4.1. Distinction Between Aleatory and Epistemic Uncertainties

Associating with characterization of aleatory and epistemic types of uncertainty, our insights and positions are as follows,

- In connection with risk analysis and communication, it is common to divide uncertainty into at least two dimensions of aleatory and epistemic portions. Since the foregoing two types of uncertainty would be input for different decisions, they must be kept separated in the computational process. If both aleatory and epistemic uncertainties were already mixed up in the course of the analysis without a clear separation, it would not be possible to identify the resulting combined effect of the uncertainties of either type.

- However, it should be noted that the reason why we clarify them into more detailed types is primarily related to important practical aspects of modeling for complex technological systems in real applications and for our convenience rather than for conceptual difference. At a fundamental level of detail, uncertainty is just uncertainty and there is only one kind of uncertainty stemming from our lack of knowledge concerning the problem of interest.

### 4.2. Guidance For Formal Treatment of Model Uncertainties

Associating with the formal treatment of model uncertainties, the present investigation gives the following guidance:

- The primary step for evaluating impacts of model uncertainties is to determine sources and types of uncertainty to be addressed in an underlying model itself and in turn model parameters. Depending on the underlying evidence, the model uncertainty may be subject to two different forms: one is the high-level model uncertainty (due to different structures in a risk analysis model) and another is the low-level model uncertainty (due to different sub-models contained in the risk analysis model).
- The next step for the analysis of model uncertainties is to provide a proper method for treating them explicitly. Except for Siu and/or Apostolakis' s method using a concept of 'a random multiplicative factor' , most of the existing applications deal with model uncertainty probabilistically by assigning subjective probabilities as a measure of the relative importance of one model over another model (with an assumption of, mutually exclusive and independent models). Then, the underlying methods for model uncertainties can be treated by either a statistical synthesis or a model sensitivity analysis. The selection of approaches basically depends on whether or not an analysis of the relative impacts of individual models to the final results is required or whether or not we would like to evaluate both uncertainties in one integrated manner for summarized information (or in explicit two dimensional assessments for a clear discrimination among uncertainty sources).
- The final step for the analysis of model uncertainties is to make a consistent interpretation of the uncertainty analysis results and to provide a summary of distributions (such 5-, 50-, 95-percentile distributions, mean distribution for a family of uncertainty distributions, and the corresponding statistical parameter values for a single output distribution). For example, a family of aleatory uncertainty distributions can be averaged over epistemic uncertainties to give one-dimensional aleatory distribution to the decision-makers.

### Acknowledgments

This study has been carried out under the Nuclear R&D Program planned by the Korean Ministry of Science and Technology (MOST).

### References

1. F. O. Hoffman and J.S. Hammonds, " Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between

Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability, *Risk Analysis*, 14 (5), pp.707-712 (1993).

2. J.C. Helton, "Treatment of Uncertainty in Performance Assessments for Complex Systems," *Risk Analysis*, 14, pp.483-511 (1994).
3. E. Hofer, " When to separate uncertainties and when not to separate, " *Reliability Engineering and System Safety*, 54, pp.113-118 (1996).
4. J.C. Helton and D.E. Burmaster, Guest Editors, "Treatment of Aleatory and Epistemic Uncertainty," Special Issue of *Reliability Engineering and System Safety*, 54 (2-3) (1996).
5. G. E. Apostolakis, "The Distinction between Aleatory and Epistemic Uncertainties is Important: An Example from the Inclusion of Aging Effects into PSA, *Proceedings of PSA 99*, Washington DC, August 22-26 (1999).
6. S. Apeland, T. Aven, and T. Nilsen, "Quantifying Uncertainty under a Predictive, Epistemic Approach to Risk Analysis," *Reliability Engineering and System Safety, 75*, pp.93-102 (2002).
7. N. Siu and G. Apostolakis, " On the Quantification of Modeling Uncertainties," *8th Intl. Conf. On Structural Mechanics in Reactor Technology*, Brussels, Belgium, August 19-23, Vol.M2, Paper 1/5, pp.375-378 (1985).
8. A. Patwardhan and M. J. Small, "Bayesian Methods for Model Uncertainty Analysis with Application to Future Sea Level Rise," *Risk Analysis*, 12 (4), pp.513-523 (1992).
9. G. E. Apostolakis, "A Commentary on Model Uncertainty," In *Model Uncertainty*: Its Characterization and Quantification, NUREG/CP-0138, (eds A. Mosleh et al.), USNRC (1994).
10. K.B. Laskey, " Implications of Model Uncertainty for the Practice of Risk Assessment," In *Model Uncertainty*: Its Characterization and Quantification, NUREG/CP-0138, (eds A. Mosleh et al.), USNRC (1994).
11. K.I. Ahn and H. Jin, "A Formal Approach for Quantitative Treatment of Modeling Uncertainties in Safety Analysis," *Nuclear Technology*, 116 (2), pp. 146-159 (1996).
12. M.E. Pate-Cornell, "Uncertainties in Risk Analysis: Six levels of treatment," *Reliability Engineering and System Safety*, 54, pp.95-111 (1996).
13. K.I. Ahn and H.D. Kim, "A Formal Procedure for Probabilistic Quantification of Modelling Uncertainties Employed in Phenomenological Transient Models," *Nuclear Technology*, 130, pp.132-144 (2000).
14. J.K. Knudsen and C. L, Smith," Estimation of System Failure Probability Uncertainty Including Model Success Criteria," CD-Rom Proceeding of PSAM 6, San Juan, Puerto Rico, June 24-28 (2002).
15. R.L. Winkler, "Uncertainty in Probabilistic Risk Assessment," *Reliability Engineering and System Safety*," 54, pp.127-132 (1996).
16. B. De Finetti, *Theory of Probability*, Vols 1 and 2, Wiley, New York (1974).
17. G. Apostolakis, "The Concept of Probability in Safety Assessments of Technological Systems," *Science*, 250, pp.1359-1364 (1990).
18. S. Kaplan, "On a "Two-Stage Bayesian Procedure for Determining Failure Rates from Experimental Data," *IEEE Transactions on Power Apparatus and Systems*, PSA102, PP.195-202 (1983).
19. Model Uncertainty: Its Characterization and Quantification, NUREG/CP-0138, (eds A. Mosleh et al.), USNRC (1994).

20. C.J, Lee and K.Y. Sung, "An Uncertainty Analysis for the Plant-specific Reliability Data Application," Proc. Of the Korean Nuclear Society Autumn Meeting, Yongpyong, Korea (October 2002).
21. T. Ishigami, E. Cazzoli, M. Khatib-Rahbar, and S.D. Unwin, " Techniques to Quantify the Sensitivity of Deterministic Model Uncertainties," Nuclear Science and Engineering, 101, pp.371-383 (1989).