Transactions of the Korean Nuclear Society Autumn Meeting
Gyeongju, Korea, October 27-28, 2016

Improved Monte Carlo Method for PSA Uncertainty Analysis

Jongsoo Choi
Korea Institute of Nuclear Safety, 62 Gwahak-ro, Yuseong-gu, Daejeon 34142; k209cjs@kins.re.kr

1. Introduction

The use of probabilistic safety assessment (PSA) technology should be increased to the extent supported by the state-of-the-art in PSA methods and data and in a manner that complements deterministic regulatory approach. The treatment of uncertainty is an important issue for regulatory decisions. Uncertainties exist from knowledge limitations. A probabilistic approach has exposed some of these limitations and provided a framework to assess their significance and assist in developing a strategy to accommodate them in the regulatory process. The uncertainty analysis (UA) is usually based on the Monte Carlo method.

This paper proposes a Monte Carlo UA approach to calculate the mean risk metrics accounting for the SOKC between basic events (including CCFs) using efficient random number generators and to meet Capability Category III of the ASME/ANS PRA standard.

2. State-of-knowledge Correlation (SOKC)

2.1 Uncertainties associated with PSA

An important aspect in understanding the PSA results is knowing the sources of uncertainty and assumptions and understanding their potential impact. Uncertainties can be either parameters or model uncertainties, and assumptions can be related either to PSA scope and level of detail or to model uncertainties.

The ASME/ANS standard on PSA[1] requires that both parameter and model uncertainties be addressed. For example, parameter uncertainties are addressed via the quantification process of the core damage and large early release frequencies and model uncertainties also have to be identified and characterized. The ASME/ANS standard notes the following:

> For CCs II and III, the mean and the distribution for the risk metric estimates are usually obtained by propagating the parameter uncertainties of the PRA inputs through the analysis using the Monte Carlo or similar sampling method.
>
> The difference between CC II and CC III is that in CC II the propagation of the uncertainty is only carried out for significant contributors in the significant accident sequences and cutsets, while for CC III the uncertainty distribution for all the input parameters is propagated to obtain the mean of the risk metrics as well as their uncertainty distributions.

Generally speaking, there are two main types of uncertainty; aleatory and epistemic. Aleatory uncertainty is based on the randomness of the nature of the events or phenomena and cannot be reduced by increasing the analyst's knowledge of the systems being modeled. The different types of epistemic uncertainty are completeness, parameter, and model uncertainty.

2.2 Treatment of parameter uncertainty

NUREG-1855[2] is provided on how to address the treatment of parameter uncertainty when using PSA results for risk-informed decision-making. NUREG-1855 addresses the characterization of parameter uncertainty; propagation of uncertainty; assessment of the significance of the state-of-knowledge correlation (SOKC); and comparison of results with acceptance criteria or guidelines. NUREG-1855 notes the following:

> In carrying out the propagation, it is important to consider the state of knowledge correlation (SOKC) between events. The SOKC arises because, for identical or similar components, the state-of-knowledge about their failure parameters is the same. In other words, the data used to obtain mean values and uncertainties of the parameters in the basic event models of these components may come from a common source and, therefore, are not independent, but are correlated.
>
> When the basic event mean values and uncertainty distributions are propagated in the PSA model without accounting for the SOKC, the calculated mean value of the relevant risk metric and the uncertainty about this mean value will be underestimated. The values can be underestimated due to the effect of the SOKC directly, as well as due to incorrect screening out of cutsets in truncation due to neglect of the SOKC in calculating cutset frequencies.

2.3 SOKC

Two of the fundamental premises on which probabilistic analyses are constructed are that: 1) the basic events of the logic model are random, independent variables, and 2) the mean values can be propagated through the logic models. There are at least two challenges to these premises: correlated data and common-cause failures.

The correlated data effect is a statistical effect that occurs when a pool of data is used to characterize the uncertainty distribution for all components of a certain type. Correlated data implies that the same distribution applies to all of these components when they are sampled using a Monte Carlo approach. The effect of correlated variables is a higher mean value than the point estimate value.

EPRI 1009652[3] notes that a significant number of internal events PSAs have propagated the parametric uncertainties through the model including the state-of knowledge correlation. The analyses have resulted in very small differences between the point estimate calculation and the Monte Carlo evaluation, as shown in Table I.

Table I. Comparison of mean and point estimate values

Plant	Data Available	Point Estimate Core Damage Frequency	Monte Carlo Simulation of Core Damage Frequency	Ratio of Difference
LaSalle (2003A) [F-2]	Plant-specific	6.64E-6	6.88E-6	1.04
Oyster Creek (IPE)[1] [F-3]	Generic	3.69E-6	4.67E-6	1.26
NUREG-1150				
Peach Bottom[2] [F-4]	Generic	3.62E-6	4.5E-6	1.24
Zion[2] [F-5]	Generic	2.8E-4	3.4E-4	1.21
Surry[2] [F-6]	Generic	3.3E-5	4.01E-5	1.22
Sequoyah[2] [F-7]	Generic	5.31E-5	5.76E-5	1.08
Grand Gulf[2] [F-8]	Generic	2.07E-6	4.05E-6	1.95
NUREG/CR-4832				
LaSalle[2] [F-9]	Generic	3.14E-5	4.41E-5	1.40

There are two reasons SOKC tends to be of low importance in the total risk metric calculation. First, there tends to be a large number of diverse contributors to core damage frequency (CDF). As shown empirically, the lower the participation fraction of correlated variables in the risk metric, the lower the impact of SOKC. Second, the addition of plant-specific data to the PRA results in reducing the number of correlated data variables in the model. Therefore, extensive use of plant-specific data suppresses SOKC impact by eliminating the number of correlated variables in the model.

2.4 Accounting for SOKC

NUREG-1855 notes the following:

The first step in accounting for the SOKC between basic events is identifying those events that are correlated and grouping them. Each identified group contains basic events that are correlated with each other because the analysts' state-of-knowledge about the parameters for these events is the same.

The groups of basic events correlated via the SOKC should not be confused with groups of common cause failures (CCFs). Although both groups account for statistical correlations between the estimates for component failure of a NPP, they account for different correlations. For this reason, accounting for one type of correlation does not account for the other. A group of correlated basic events can contain several events, including those modeled within a CCF group. For instance, a CCF group may contain one failure mode of all the pumps of a particular system, while a group of correlated basic events may encompass the same failure mode for all the pumps of this type within the NPP. Hence, both types of correlations (i.e., CCF and SOKC) should be included in a PRA model.

EPRI 1009652 notes the following:

The empirical evaluations provide insights into the areas of the PSA that may be influenced by SOKC and in turn may need to be addressed by the decision maker to provide an accurate representation of the results. The following insights are derived from these empirical calculations:

- SOKC can be significant.
- Modeling of correlated events within a CCF group tends to reduce the significance of SOKC.
- The impact of SOKC on risk metrics increases as:
  - EFs increase
  - Fraction of risk metric impacted by the variables that have a SOKC increases
  - Number of coincident correlated variables increases
- High EFs: If the model has high EFs, the correlation effect is generally large. It is also noted that the high EFs also create non-monotonic responses in the Monte Carlo calculation. In general, if the error factors are very large, >30, extreme care in the interpretation of the results is needed.

3. Improved approach accounting for SOKC

3.1 Efficient sampling techniques

The generation of pseudo-random numbers is an important and common task in computer programming of Monte Carlo simulations. The Ziggurat algorithm[4] is a method for efficient random sampling from a probability distribution such as Normal distribution. The following diagram (Fig. 1) demonstrates this algorithm. Note that we operate on one side of the pdf ($x \geq 0$), generating both positive and negative sample values requires that as a final step we randomly flip the sign of the generated non-negative values.

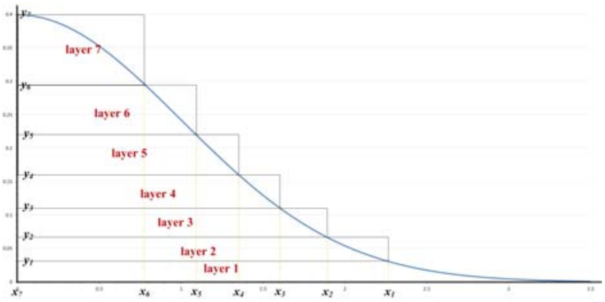

Fig. 1. Example ziggurat with 7 layers of $N(0,1)$

The Ziggurat algorithm gives good performance by using a very simple rejection sampling execution path for the majority of sample points generated, but with more expensive calculations performed to maintain mathematical exactness in some specific corner cases represented by the distribution tail and the far edges of the Ziggurat's rectangles.

Latin hypercube sampling (LHS) is a statistical method for generating a near-random sample of parameter values from a multidimensional distribution. The LHS method is often used to construct computer experiments or for Monte-Carlo integration. In the context of statistical sampling, a square grid containing sample positions is a Latin square if (and only if) there is only one sample in each row and each column. A Latin hypercube is the generalization of this concept to an arbitrary number of dimensions, whereby each sample is the only one in each axis-aligned hyperplane containing it.

In LHS, we must first decide how many sample points to use and for each sample point remember in which row and column the sample point was taken. Note that such configuration is similar to having N rooks on a chess board without threatening each other. Although more efficient than random sampling, this sampling strategy is more difficult to implement since all random samples must be generated simultaneously and to overcome the limitations of memory and speed.

3.2 The proposed approach

The first step of the proposed approach is identifying all basic events that are correlated and grouping them. In this step, we can identify the cutsets that are affected by SOKC effects.

The second step is accounting for the SOKC between CCFs by dividing each CCF into its component failure and its CCF parameter. CCFs occur when multiple (usually identical) components fail due to shared causes. Typical examples of shared causes include impact, vibration, temperature, contaminants, miscalibration and improper maintenance. Although Multiple Greek Letters (MGL), one of parametric models for CCF, is generally used in Korean PSAs, we don't have uncertainty information of MGL parameters. The probability of each CCF can be quantified through multiplying its component failure probability by its CCF parameter. This approach individually provides SOKC grouping of component failure probabilities and CCF parameters.

The third step is propagating the parameter uncertainties of the PSA inputs (including SOKC information) through iterative Monte Carlo simulations with a large number of replicated runs.

3.3 Application to Example PSA model

In order to assess the adequacy of the proposed approach to NPP PSA models, an example PSA model is selected as follows:

- Level 1 internal event PSA model of a plant
- CDF : 1.093E-6/years
- # of MCSs : 24,083
  - SOKC sets : 3,641
  - not SOKC sets : 20,442
- # of basic events : more than 3,800
  - CCFs : more than 1,000
  - SOKC groups : 163

Table II compares the effects of SOKC for Example PSA model. It is shown that its SOKC tends to be of low importance in the total risk metric calculation.

Table II. MC results (from 100 runs of sample size 1E5)

	mean of mean	var. of mean
Random, SOKC	1.101E-6	1.922E-16
Random, w/o SOKC	1.091E-6	2.176E-16
LHS, SOKC	1.096E-6	1.260E-16
LHS, w/o SOKC	1.088E-6	1.024E-16

Table III shows the Monte Carlo simulation results using the proposed CCF separation model. In this simulation, error factors of MGL parameters are assumed 10.

Table III. MC results using new CCF model (from 100 runs of sample size 1E5)

	mean of mean	var. of mean
Random, SOKC	1.085E-6	1.364E-16
Random, w/o SOKC	1.075E-6	1.395E-16
LHS, SOKC	1.083E-6	9.697E-17
LHS, w/o SOKC	1.073E-6	1.274E-16

Table IV compares the effects of two sampling techniques (random sampling and LHS at specific sample sizes) on accuracy of CDF estimates.

Table IV. Effects of sampling methods (100 runs)

Sample size	Sampling	mean of mean	var. of mean
1,000	Random	1.062E-6	1.264E-14
	LHS	1.085E-6	7.527E-15
10,000	Random	1.086E-6	1.386E-15
	LHS	1.080E-6	7.806E-16
100,000	Random	1.085E-6	1.364E-16
	LHS	1.083E-6	9.697E-17

The proposed approach provides a variety of histograms and probability density function plots. Fig. 2 is a histogram based on distribution parameters from a set of MC outputs.

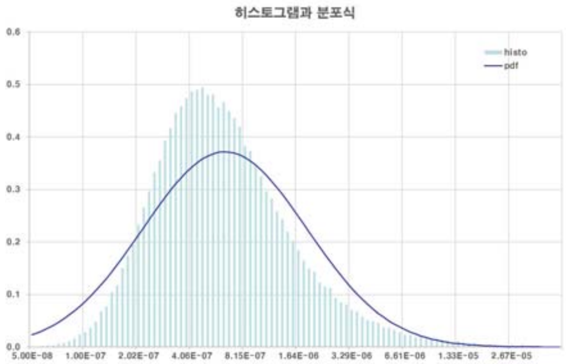

Fig. 2. Generating histogram

4. Conclusions

Audit calculation is needed in PSA regulatory reviews of uncertainty analysis results submitted for licensing. The proposed Monte Carlo UA approach provides a high degree of confidence in PSA reviews. All PSA needs accounting for the SOKC between event probabilities to meet the ASME/ANS PRA standard.

REFERENCES

[1] Standard for Level 1/Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plant Applications, ASME/ANS RA-Sa-2009, ASME/ANS, 2009.
[2] Guidance on the Treatment of Uncertainties Associated with PRAs in Risk-Informed Decision making, NUREG-1855 Rev.1, USNRC, 2013.
[3] Guideline for the Treatment of Uncertainty in Risk-Informed Applications", EPRI 1009652, EPRI, 2004.
[4] J. Choi, Generating Log-normally Distributed Random Numbers by Using the Ziggurat Algorithm, KNS Spring Meeting, Jeju, May 12-13, 2016.