#### ABSTRACT

Title of Dissertation:	A BAYESIAN NETWORK PERSPECTIVE ON THE ELEMENTS OF A NUCLEAR POWER PLANT MULTI-UNIT SEISMIC PROBABILISTIC RISK ASSESSMENT
	Jonathan DeJesus Segarra, Doctor of Philosophy, 2021
Dissertation directed by:	Assistant Professor Michelle Bensi Reliability Engineering Program, Department of Civil and Environmental Engineering and Professor Mohammad Modarres Reliability Engineering Program, Department of Mechanical Engineering

Nuclear power plants (NPPs) generated about 10% of the world's electricity in 2020 and about 1/3 of the world's low-carbon electricity production. Probabilistic risk assessments (PRAs) are used to estimate the risk posed by NPPs, generate insights related to strengths and vulnerabilities, and support risk-informed decisionmaking related to safety and reliability. While PRAs are typically carried out on a reactor-by-reactor basis, the Fukushima Dai-ichi accident highlighted the need to also consider multi-unit accidents. To properly characterize the risks of reactor core damage and subsequent radiation release at a multi-unit site, it is necessary to account for dependencies among reactors arising from the possibility that adverse conditions affect multiple units concurrently. For instance, the seismic hazard is one of the most critical threats to NPP

structures, systems, and components (SSCs) because it affects their redundancy. Seismic PRAs are comprised of three elements: seismic hazard analysis, fragility evaluation, and systems analysis.

This dissertation presents a Bayesian network (BN) perspective on the elements of a multi-unit seismic PRA (MUSPRA) by outlining a MUSPRA approach that accounts for the dependencies across NPP reactor units. BNs offer the following advantages: graphical representation that enables transparency and facilitates communicating modeling assumptions; efficiency in modeling complex dependencies; ability to accommodate differing probability distribution assumptions; and facilitating multi-directional inference, which allows for the efficient calculation of joint and conditional probability distributions for all random variables in the BN. The proposed MUSPRA approach considers the spatial variability of the ground motions (hazard analysis), dependent seismic performance of SSCs (fragility evaluation), and efficient BN modeling of systems (systems analysis). Considering the spatial variability of ground motions represents an improvement over the typical assumption that ground motions across a NPP site are perfectly correlated. The method to model dependent seismic performance of SSCs presented is an improvement over the current "perfectly dependent or independent" approach for dependent seismic performance and provides system failure probability results that comply with theoretical bounds. Accounting for these dependencies in a systematic manner makes the MUSPRA more realistic and, therefore, should provide confidence in its results (calculated metrics) and risk insights.

#### A BAYESIAN NETWORK PERSPECTIVE ON THE ELEMENTS OF A NUCLEAR POWER PLANT MULTI-UNIT SEISMIC PROBABILISTIC RISK ASSESSMENT

by

Jonathan DeJesus Segarra

Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, College Park, in partial fulfillment of the requirements for the degree of Doctor of Philosophy 2021

Advisory Committee: Assistant Professor Michelle T. Bensi, Chair Professor Mohammad Modarres, Co-Chair Professor Gregory B. Baecher Associate Professor Katrina M. Groth Dr. Robert J. Budnitz (Special Member) Associate Professor Vedran Lekić (Dean's Representative) © Copyright by Jonathan DeJesus Segarra 2021

## Dedication

To my wife and best friend, Aixa, whose love made me a better person and who encouraged me when the graduate studies journey became frustrating and difficult. This is OUR achievement.

To my children, Jonathan and Isabella, who at the same time I could have used as an excuse for not completing my graduate studies and as the reason to complete my graduate studies. I hope that by me completing this dissertation you take away an illustrative example of perseverance. I'm excited to see you grow and I promise to help you persevere in the challenges that life will throw your way and the ones that you choose to pursue.

To my mother, Maritza, who wholeheartedly supported me when I decided to seek my undergraduate degree in Chemical Engineering, which set the stage of my academic and professional career.

## Acknowledgements

This graduate studies journey was an incredibly humbling experience; there is so much out there that remains to be discovered. This journey made me see that the following quote, attributed to Isaac Newton (although the metaphor can be traced back to Bernard of Chartres, a French philosopher), is remarkably true: "If I have seen further it is by standing on the shoulders of giants."<sup>1</sup>

First of all, I would like to express my sincerest appreciation and admiration for my advisor, Prof. Michelle Bensi. Thank you for accepting me as your student just several weeks after you started as a professor at the University of Maryland (UMD) in 2017. Your attention to detail and work ethic are commendable and have significantly influenced me. I appreciate when you call me out when my math notation gets sloppy, as I always say to you, "thank you for keeping me honest." Your efforts to keep a connection among all of your students and the empathy you showed this past year were outstanding. I will always be honored to be associated with you.

Next is my co-advisor, Prof. Mohammad Modarres. Thank you for accepting me in the Reliability Engineering program and as your student, your patience, and sharing your wisdom with me. One advice that I will not forget, that go beyond academic teaching, is to take criticisms

<sup>&</sup>lt;sup>1</sup> Chen C. (2003) On the Shoulders of Giants. In: Mapping Scientific Frontiers: The Quest for Knowledge Visualization. Springer, London. <u>https://doi.org/10.1007/978-1-4471-0051-5\_5</u>

as if they are being made about an individual's work and not about the individual as a person. I am proud to be your student.

Prof. Katrina Groth, Prof. Gregory Baecher, Dr. Robert Budnitz, and Prof. Vedran Lekić, thank you for enthusiastically accepting being part of the Dissertation Committee, reading this dissertation, and listening to my proposal presentation and dissertation defense.

This accomplishment would not have been possible without the financial and time investment that the U.S. Nuclear Regulatory Commission (NRC) made in me through its graduate fellowship program (GFP). I would like to thank Marissa Bailey (currently Director (Acting), Division of Construction Oversight in NRC's Region II); Michael Franovich (currently Director, Division of Risk Assessment in the NRC's Office of Nuclear Reactor Regulation); (former) Commissioner William Magwood, IV (currently Director-General, Nuclear Energy Agency); and Dr. Dennis Damon (NRC retired) for supporting my application to the GFP position. A special thank you goes to Sean Peters (currently Chief, Human Factors and Reliability Branch in the NRC's Office of Nuclear Regulatory Research) for selecting me for the GFP position and his support throughout my time at UMD; I assure you, Sean, that the skills I gained in my time at UMD significantly contributed to my performance that resulted in the publication of NUREG-2198.

iv

# Table of Contents

Dedication	ii
Acknowled	lgementsiii
Table of C	ontentsv
List of Tab	les xi
List of Fig	uresxiii
List of Abl	previations xvii
Chapter 1	Introduction1
1.1	Motivation1
1.2	Research Objectives
1.3	Methods 6
1.4	Dissertation Outline
Chapter 2	A Brief Introduction to Bayesian Networks
2.1	Bayesian Networks
2.2	Example Bayesian Network Calculations15
2.2.1	Forward Inference
2.2.2	Backward Inference
Chapter 3	Extension of Probabilistic Seismic Hazard Analysis to Account for the Spatial
	Variability of Ground Motions at a Multi-Unit Nuclear Power Plant Hard-Rock Site

3.1	Introduction	. 21
3.2	Background	. 23
3.2.1	Seismic Hazard Curve	. 23
3.2.2	Use of the Seismic Hazard Curve in the Quantification of a Seismic PRA	. 26
3.2.3	Disaggregation of the Seismic Hazard	. 27
3.2.4	Spatial Variability of Ground Motions	. 28
3.3	Framework for Modeling Ground Motion Spatial Variability	. 31
3.3.1	General Bayesian Network to Model Ground Motion Spatial Variability at Hard-	
	Rock Sites	. 32
3.3.2	Streamlined Bayesian Network to Model Ground Motion Spatial Variability at a	
	Hard-Rock Site	. 36
3.3.3	MUSPRA Input for a Hard-Rock Site	. 37
3.4	Example Application of the Proposed Method	. 41
3.4.1	PSHA Results for the Central Illinois Site	. 42
3.4.2	Implementation of the Ground Motion Spatial Variability Model	. 43
3.4	2.1 <i>SA</i> <sup><i>r</i></sup> <sub>1</sub>	. 43
3.4	$M SA_1^r$	. 46
3.4	-2.3 $\Delta SA_{1,2} M \text{ and } P'_{1,2}$	. 47
3.4	-2.4 $SA_2^r   SA_1^r, \Delta SA_{1,2}$ and $SA_2^r   SA_1^r, P_{1,2}'$	. 48
3.4.3	Ground Motion Probability Distributions	. 49

3.5	Summary and Conclusions	54
3.6	Disclaimer and Acknowledgement	56
Chapter 4	A Bayesian Network Approach for Modeling Dependent Seismic Failures in a	
	Nuclear Power Plant Probabilistic Risk Assessment	57
4.1	Introduction	57
4.1.1	Overview	57
4.1.2	Context and Motivation	59
4.1.3	Outline	62
4.2	Background	63
4.2.1	Overview of the Seismic Fragility Modeling Used in the Nuclear Industry	63
4.2.2	Estimation of the Fragility Parameters (Fragility Evaluation)	64
4.2.3	The Reed-McCann Method	67
4.2.4	Lower and Upper Bounds of a System Failure Probability	69
4.3	Bayesian Network Approach for Modeling Dependent Seismic Failures	70
4.3.1	Generation of the CPT for the EES Node	74
4.3.2	Generation of the CPT for the GM EES Node	75
4.3.3	Generation of the CPTs for the $\varepsilon_{AB}^*$ , $\varepsilon_{AC}^*$ , and $\varepsilon_{BC}^*$ Nodes	77
4.3.4	Generation of the CPTs for the $\varepsilon'_A$ , $\varepsilon'_B$ , and $\varepsilon'_C$ Nodes	79
4.3.5	Generation of the CPTs for the $A GM, \varepsilon_{AB}^*, \varepsilon_{AC}^*, \varepsilon_A', B GM, \varepsilon_{AB}^*, \varepsilon_{BC}^*, \varepsilon_B'$ , and	
	$C GM, \varepsilon_{AC}^*, \varepsilon_{BC}^*, \varepsilon_{C}'$ Nodes	81

4.3.6	Generation of CPT for the <i>System</i>   <i>A</i> , <i>B</i> , <i>C</i> Node	84
4.4	Results Comparison	85
4.5	Summary and Conclusions	95
4.6	Disclaimer and Acknowledgements	97
Chapter 5	Multi-Unit Seismic Probabilistic Risk Assessment: A Bayesian Network F	Perspective
5.1	Introduction	
5.1.1	Motivation and Objective	
5.1.2	Context	101
5.1.3	Background	104
5.1.4	Outline	104
5.2	Approach to MUSPRA using Bayesian Networks	104
5.2.1	Single-Unit Seismic PRA Description	105
5.2.2	Mapping of the Single-Unit Seismic PRA into a Bayesian Network	110
5.2.3	Modeling MUSPRA as a Bayesian Network	112
5.2.	.3.1 Seismic Demand Module – Spatial Variability of Ground Motion	112
5.2.	.3.2 System Performance Module – Dependent Seismic Failures	114
5.2.	.3.3 Site-Based Risk Metrics Module	118
5.3	Results and Discussion	120
5.4	Summary and Conclusions	127

5.5	Disclaimer and Acknowledgements	129
Chapter 6	Conclusion, Summary of Contributions, and Future Research	
6.1	Conclusion	
6.2	Summary of Contributions	
6.3	Future Research	
Bibliography		

THIS PAGE IS INTENTIONALLY LEFT BLANK

## List of Tables

Table 1-1	Summary of NPP Sites by Country (as of June 2021) [3], [4]	
Table 3-1	Comparison and Explanation of the Spatial Variability of Ground Motion	Models
	Used in this Study	30
Table 3-2	Definition of the RVs in Figure 3-2(a)	
Table 3-3	Discretization of Seismic Hazard Curve	43
Table 3-4	CPT for Node <i>SA</i> <sup><i>r</i></sup> <sub>1</sub>	45
Table 3-5	CPT for Node <i>M</i>   <i>SA</i> <sup><i>r</i></sup> <sub>1</sub>	46
Table 3-6	CPT for Node $\Delta SA_{1,2} M$	47
Table 3-7	Partial CPT for $SA_2^r   SA_1^r, \Delta SA_{1,2}$	49
Table 4-1	Summary of Dependency Assumptions	69
Table 4-2	Fragility Parameters and Group Dependencies	74
Table 4-3	CPT for the <i>EES</i> Node	75
Table 4-4	Bins and CPT for the <i>GM</i>   <i>EES</i> Node	76
Table 4-5	Bins and CPT for the $\varepsilon_{AB}^*$ Node	
Table 4-6	Bins and CPT for the $\varepsilon'_A$ Node	80
Table 4-7	"Snippet" of the CPT for Component A given GM, $\varepsilon_{AB}^*$ , $\varepsilon_{AC}^*$ , and $\varepsilon_{A}'$	
Table 4-8	System   A, B, C Node CPT for the Parallel System	
Table 4-9	System   A, B, C Node CPT for the Series System	85
Table 4-10	System Fragility of the Parallel Configuration in Section 8.2.4 of	
	NUREG/CR-7237 [15]	89
Table 4-11	System Fragility of the Series Configuration in Section 8.2.5 of	
	NUREG/CR-7237 [15]	90

Table 5-1	Fragility Parameters for basic events	109
Table 5-2	Common Aleatory and Epistemic Uncertainties	116
Table 5-3	Comparison of Results for NPP Site-Based Risk Metrics from 15 Cases	124

# List of Figures

Figure 1-1	Capacity factors of electricity generation by fuel source in the U.S. in 2020 [5],
	[6]. The green and orange bars indicate non-fossil and fossil fuel sources,
	respectively2
Figure 1-2	Illustration of the methods overview in the MUSPRA approach
Figure 2-1	Example of the qualitative part of a BN: the DAG 12
Figure 2-2	BN and CPTs for the EDG failure-to-start problem
Figure 3-1	Example Bayesian network
Figure 3-2	(a) Ground motion amplitude spatial variability BN for a single rock site with two
	units. Illustration of the variables in the BN shown in Figure 3-2(a): (b) Top view
	(c) Elevation view
Figure 3-3	Streamlined BNs to model ground motion spatial variability using the (a) A&S
	[24] and (b) K&M [25] formulations. (c) Illustration of the variables in the
	streamlined BNs ( $M$ and $R$ are not shown in the illustration of the variables) 37
Figure 3-4	(a) BN corresponding to Equation (3.7). (b) BN corresponding to Equation (3.9).
	(c) BN corresponding to Equation (3.8)
Figure 3-5	Mean rock hazard curve at the Central Illinois Site
Figure 3-6	(a) Probability distributions of the ground motion at the reference and non-
	reference units given an earthquake of engineering significance (EES). (b)
	Probability distribution of the ground motion at the reference unit and conditional
	probability distribution at the non-reference unit given that $SA_1^r$ is in the range
	0.4g – 0.56g (BIN-3) and an EES

Figure 3-7	Probability distribution of the ground motion at the reference unit and comparison
	of the conditional probability distributions at the non-reference unit given that
	$SA_1^r$ is in the range 0.4g – 0.56g (BIN-3) and an EES using the A&S [24] and
	K&M [25] formulations estimated using MC simulation and the BN framework 53
Figure 3-8	Conceptual extensions of the BN frameworks based on (a) the A&S [24] and (b)
	K&M [25] formulations to multiple locations on a site
Figure 4-1	Summary of the Reed-McCann method
Figure 4-2	Bayesian network for modeling the fragility of a three-component system74
Figure 4-3	System fragility of the parallel configuration in Section 8.2.4 of
	NUREG/CR-7237 [15] for the case where the GM value is fixed. The shaded area
	represents the lower and upper bounds of the system fragility for the parallel
	configuration
Figure 4-4	System fragility of the series configuration in Section 8.2.5 of NUREG/CR-7237
	[15] for the case where the GM value is fixed. The shaded area represents the
	lower and upper bounds of the system fragility for the series configuration92
Figure 5-1	At-power seismic initiating event tree
Figure 5-2	Fault tree for the seismic event leading to direct core damage (1-01-S-DMG)
	pivotal event
Figure 5-3	Fault tree for the seismic-induced SBO (1-02-S-SBO) pivotal event 108
Figure 5-4	Fault trees for the seismic-induced (a) ATWS (1-06-S-ATWS), (b) large LOCA
	(1-07-S-LLOCA), (c) medium LOCA (1-08-S-MLOCA), (d) small LOCA (1-09-
	S-SLOCA), and (e) LOOP (10-S-LOOP) pivotal events

Figure 5-5	Example of a converging BN structure for a series system with four MCS with its
	СРТ 110
Figure 5-6	Example of a chain-like BN structure for a series system with four MCS with its
	CPT and the CPTs of the chain-like nodes 111
Figure 5-7	Event tree in Figure 5-1 mapped as a chain-like BN structure with the CPTs for
	the failure path events and Unit 1 112
Figure 5-8	BN modules to model a MUSPRA 112
Figure 5-9	BN and CPTs used to model the spatial variability of ground motion 114
Figure 5-10	BN representation of the 1-02-S-SBO fault tree
Figure 5-11	CPTs for various nodes in the BN representing the 1-02-S-SBO fault tree 118
Figure 5-12	BN model of MUSPRA 120
Figure 5-13	CPT for the Site node
Figure 5-14	(a) BN for the no ground motion correlation assumption. (b) BN for the perfect
	ground motion correlation assumption and CPT for node 2-GM 121
Figure 5-15	BN models for the (a) none; (b) partial-intra, none-inter; (c) perfect-intra, partial-
	inter; and (d) perfect SSC fragility correlation assumptions

THIS PAGE IS INTENTIONALLY LEFT BLANK

# List of Abbreviations

annual exceedance frequency
American Nuclear Society
air start motor
American Society of Mechanical Engineers
anticipated transient without scram
Abrahamson and Sykora (i.e., Reference [24])
Bayesian network
common cause failure
core damage
core damage frequency
conservative deterministic failure margin
Central and Eastern United States
conditional probability table
directed acyclic graph
design control document
direct quantification of fault tree using Monte Carlo simulation
emergency diesel generator
earthquake of engineering significance
Electric Power Research Institute
electrical power system
Finite Element Reliability Using Matlab
first-order reliability method
failure path event
fails to start
graduate fellowship program
ground motion
ground motion model
high confidence of low probability of failure
International Atomic Energy Agency
Kawakami and Mogi (i.e., Reference [25])
loss-of-coolant accident
loss of offsite power
Monte Carlo
Monte Carlo simulation (in Chapter 4)

MCS	minimal cut set (in Chapter 5)
MSR	Matrix-based System Reliability analysis method
MUPRA	multi-unit probabilistic risk assessment
MUPSA	multi-unit probabilistic safety assessment
MUSPRA	multi-unit seismic probabilistic risk assessment
ND	not defective
NM	not malfunctioning
NPP	nuclear power plant
NRC	U.S. Nuclear Regulatory Commission
PGA	peak ground acceleration
PMF	probability mass function
PRA	probabilistic risk assessment
PSA	probabilistic safety assessment
PSHA	probabilistic seismic hazard analysis
RCS	reactor coolant system
RV	random variable
SA	spectral acceleration
SBO	station blackout
SG	steam generator
SMA	seismic margin analysis
SPID	screening, prioritization and implementation details
SPRA	seismic probabilistic risk assessment
SSC	structures, systems, and components
SSMRP	seismic safety margins research program
SwC	speed switch circuit
S-LOOP	seismically-induced loss of offsite power
S-SLOCA	seismically-induced small loss-of-coolant accident
TRANS	transient
UMD	University of Maryland

## Chapter 1 Introduction

#### 1.1 <u>Motivation</u>

In 2020, nuclear power plants (NPPs) generated 2,553.2 terawatt-hours of emission-free, lowcarbon baseload electricity, which accounted for about 10% of total global electricity generation and nearly a third of the world's low-carbon electricity production [1]. As shown in Figure 1-1, electricity generation using NPPs in the U.S. has the highest capacity factor<sup>2</sup> of all fuel sources, which makes nuclear power an attractive carbon-free and sustainable baseload source of electricity. It is typical for a NPP site to contain multiple reactor units and other sources of radioactive material (e.g., spent-fuel pool and independent spent fuel storage installation). For instance, as of June 2021, there are 181 NPP sites worldwide with 440 nuclear reactor units in operation [3], [4]. About 89.1 percent of the nuclear reactor units are located on multi-unit NPP sites. Table 1-1 summarizes the NPP site information by country.

<sup>&</sup>lt;sup>2</sup> Capacity factor is the ratio of the electrical energy produced by a generating unit for the period of time considered to the electrical energy that could have been produced at continuous full power operation during the same period [2].



Figure 1-1 Capacity factors of electricity generation by fuel source in the U.S. in 2020 [5], [6]. The green and orange bars indicate non-fossil and fossil fuel sources, respectively.

Probabilistic risk assessments (PRAs) (also referred to as probabilistic safety assessments) are widely used to estimate the risk posed by nuclear power facilities and generate insights related to the strengths and potential vulnerabilities of NPP designs [7]. A PRA answers a set of questions known as the risk triplet: "What can go wrong?" "How likely it is?" and "What are the consequences?" [8]. Despite the presence of multiple reactor units at most NPP sites, the safety and risk analyses are typically performed on a reactor-by-reactor basis. Under such analyses, when one reactor is being analyzed, the other(s) is (are) assumed to be in a safe, stable condition or otherwise unaffected. The Fukushima Dai-ichi accident [9]–[11] highlighted the need to reconsider this assumption and address multi-unit accidents in NPP PRAs. The accident at the

Fukushima Dai-ichi NPP initiated as an earthquake followed by a tsunami that resulted in three of the six units at the site experiencing core damage and subsequent release of radioactive material to the environment. The focus of this dissertation is on the performance of multi-unit PRA (MUPRA) to assess risks from seismic events because "[i]t has long been recognized that external events, particularly seismic and external flooding events, could be substantial contributors to risk because of the potential for multiple common-cause failures" [10]. An earthquake may result in a core damage event, and radiation release. The seismic PRA is intended to estimate the frequency of those adverse events.

Seismic PRAs are comprised of three key elements: (1) seismic hazard analysis, (2) seismic fragility evaluation, and (3) systems (plant response) analysis [12]. The seismic hazard analysis is typically performed using the probabilistic seismic hazard analysis (PSHA) process, which estimates the frequency (per year) of exceeding a certain value of a ground motion parameter (e.g., peak ground acceleration (PGA) or spectral acceleration). The results of a PSHA are typically presented as a seismic hazard curve for a location of interest. The fragility evaluation is used to estimate the seismic capacity and conditional probability of failure of structures, systems, and components (SSCs) as a function of the level (intensity) of a ground motion parameter. The system (plant response) analysis includes the modeling of the NPP's response to the ground motion caused by an earthquake using event trees and fault trees. The plant response also includes the quantification of the appropriate risk metrics (e.g., core damage frequency). To realistically model seismic events in a MUPRA, advancements are needed in each of the key elements of a seismic PRA.

3

Despite the importance of MUPRA, the current state of practice is relatively limited and existing analyses have typically used a number of simplifying assumptions. These simplifying assumptions are the result of limited knowledge, resource constraints, and the limitations of existing PRA tools used in the nuclear industry. For example, current multi-unit seismic PRAs (MUSPRA) (e.g., [13]) assume that all units at the NPP site experience the same ground motion, that is, the ground motions at all units are perfectly correlated. At the NPP site scale, there is expected to be spatial variability in the ground motion at different locations around the site for a given earthquake due to incoherence, wave-passage, attenuation, and site-response effects [14]. Thus, the perfect ground motion correlation assumption is not realistic. Lack of realism in a MUSPRA may lead to distorted risk insights, may mask some risk contributors, and can adversely affect risk-informed decisionmaking. Therefore, an advancement is needed to include the spatial variability of ground motion in a MUSPRA and understand the potential effects of the simplifying assumptions.

Given that ground motions due to an earthquake affect all units at a NPP site (seismic demand) and the redundant SSCs are designed, manufactured, installed, maintained, and analyzed in similar ways (seismic capacity), there is dependency (correlation) between the seismic performance of SSCs. The current approach in seismic PRAs is to assume that the seismic performance of redundant SSCs is either perfectly correlated or independent [15]. These two assumptions are extreme and the truth is likely somewhere in the middle. Therefore, an advancement is needed to consider partial correlation in the seismic performance of redundant SSCs while at the same time incorporating the spatial variability of ground motions in a MUSPRA.

4

The systems (plant response) analysis in seismic PRAs are typically developed using event trees and fault trees, where the underlying assumption is that the pivotal events in event trees and basic events in fault trees are independent [16]. A complementary approach to the event tree/fault tree approach (advancement) is needed to properly account for the spatial variability of ground motion and dependent seismic performance of SSCs in the estimation of NPP site-based risk metrics and development of risk insights in a MUSPRA.

#### 1.2 <u>Research Objectives</u>

This dissertation has three objectives (listed below) which are related to the key elements of a seismic PRA and the needed advancements discussed in Section 1.1.

- Objective 1: Develop a method that models the spatial variability of ground motion in a MUSPRA.
- Objective 2: Develop a method to model the dependent seismic performance of SSCs that integrates with the method that models the spatial variability of ground motion, while leveraging the existing conceptual approaches used in seismic PRAs performed in the nuclear industry.
- Objective 3: Integrate the works performed to accomplish Objectives 1 and 2 into a MUSPRA approach to explore the relative importance of different types of correlation on the NPP site-based risk metrics.

The research performed to address Objectives 1, 2, and 3 is described in Chapter 3, Chapter 4, and Chapter 5, respectively.

#### 1.3 <u>Methods</u>

This dissertation uses Bayesian networks (BNs) as the communication, modeling, and quantification approach to accomplish the research objectives described in Section 1.2. An introduction to BNs is provided in Chapter 2. For a hypothetical dual unit NPP site, Figure 1-2 shows a BN that illustrates an overview of the three methods that are integrated to form the MUSPRA approach proposed by this dissertation.



EES: earthquake of engineering significance GM: ground motion

#### Figure 1-2 Illustration of the methods overview in the MUSPRA approach

First is the method for capturing the effects of spatial variability of ground motion at an NPP site in the MUSPRA, which is discussed in detail in Chapter 3. The proposed method uses the results from an existing PSHA performed for a reference location at the site (e.g., the site seismic hazard curve), which is reflected by the node 1-GM in Figure 1-2 and represents the seismic demand for the SSCs in Unit 1. Then, the method estimates the conditional probability distribution of the ground motion at a non-reference location represented by the node 2-GM in Figure 1-2 and reflects the seismic demand of the SSCs in Unit 2. The method accounts for the spatial variability in ground motion using the node  $P'_{1,2}$  in Figure 1-2. The node EES represents an earthquake of engineering significance (EES) event. An EES event occurs when the ground motion exceeds a minimum threshold. In the context of a seismic PRA, the minimum threshold is the starting value of the first ground motion bin in the discretized seismic hazard curve.

Next is the method for modeling the dependent seismic performance of SSCs, which is discussed in detail in Chapter 4. The proposed method relies on a systematic search of the common sources of uncertainties among the SSCs of interest. In Figure 1-2, the nodes Unit 1 and Unit 2 represent the seismic performance of each unit and at a high level these are the SSCs of interest. The systematic search for the common sources of uncertainty results in the node  $\varepsilon_{U1,U2}^*$ , which is a common parent to the seismic performance of each unit and accounts for the dependency between the units.

Finally, the methods for modeling spatial variability of ground motion and dependent seismic performance are integrated in a MUSPRA approach, whose details are discussed in Chapter 5. The proposed MUSPRA approach creates the MUPRA model based on the single-unit PRA model, maps the event trees and fault trees in the seismic PRA into an equivalent BN, and uses the nodes related to the spatial variability of ground motion and dependent seismic performance to explicitly address the dependencies between the units and estimate the NPP site-based risk metrics (i.e., the node Site in Figure 1-2). The MUSPRA approach also uses efficient BN modeling of systems to complement the typically used event trees and fault trees.

#### 1.4 <u>Dissertation Outline</u>

The remaining of this dissertation is organized as follows. Chapter 2 presents a brief introduction to Bayesian networks and provides an example to illustrate the general concepts behind how the

calculations in subsequent chapters of this dissertation are performed. Chapter 3 presents the method for obtaining the conditional probability distribution of ground motion at a non-reference location(s) on a site given the ground motion at a reference location on the same site. Chapter 4 discusses the method to model dependent seismic performance of SSCs. Chapter 5 presents the MUSPRA approach that integrates the spatial variability of ground motion and dependent seismic performance. Finally, Chapter 6 provides conclusions, summarizes the contributions of this dissertation, and discusses potential future research.

Country	# of Units in Operation	# of Operational Units in Multi-Unit Sites	# of Sites	# of Single- Unit Sites	# of Multi- Unit Sites	# of Sites with 2 Units	# of Sites with 3 Units	# of Sites with 4 Units	# of Sites with 5 Units	# of Sites with 6 Units	# of Sites with 7 Units	# of Sites with 8 Units
Argentina	3	2	2	1	1	1	0	0	0	0	0	0
Armenia	1	0	1	1	0	0	0	0	0	0	0	0
Belarus	1	0	1	1	0	0	0	0	0	0	0	0
Belgium	7	7	2	0	2	0	1	1	0	0	0	0
Brazil	2	2	1	0	1	1	0	0	0	0	0	0
Bulgaria	2	2	1	0	1	1	0	0	0	0	0	0
Canada	19	18	3	1	3	0	0	1	0	1	0	1
China	51	50	14	1	14	7	0	3	1	2	1	0
Czech Republic	6	6	2	0	2	1	0	1	0	0	0	0
Finland	4	4	2	0	2	2	0	0	0	0	0	0
France	56	56	18	0	18	9	0	8	0	1	0	0
Germany	6	0	0	6	0	0	0	0	0	0	0	0
Hungary	4	4	1	0	1	0	0	1	0	0	0	0
India	23	23	7	0	7	3	1	2	0	1	0	0
Iran	1	0	0	1	0	0	0	0	0	0	0	0
Japan	33	27	9	6	9	5	2	1	0	0	1	0
Mexico	2	2	1	0	1	1	0	0	0	0	0	0
Netherlands	1	0	0	1	0	0	0	0	0	0	0	0
Pakistan	6	6	2	0	2	1	0	1	0	0	0	0
Romania	2	2	1	0	1	1	0	0	0	0	0	0
Russia	38	38	11	0	11	2	2	7	0	0	0	0
Slovakia	4	4	2	0	2	2	0	0	0	0	0	0
Slovenia	1	0	0	1	0	0	0	0	0	0	0	0
South Africa	2	2	1	0	1	1	0	0	0	0	0	0
South Korea	24	24	4	0	4	0	0	0	1	2	1	0
Spain	7	4	2	3	2	2	0	0	0	0	0	0
Sweden	6	5	2	1	2	1	1	0	0	0	0	0
Switzerland	4	2	1	2	1	1	0	0	0	0	0	0
Ukraine	15	15	4	0	4	1	1	1	0	1	0	0

Table 1-1Summary of NPP Sites by Country (as of June 2021) [3], [4]

Country	# of Units in Operation	# of Operational Units in Multi-Unit Sites	# of Sites	# of Single- Unit Sites	# of Multi- Unit Sites	# of Sites with 2 Units	# of Sites with 3 Units	# of Sites with 4 Units	# of Sites with 5 Units	# of Sites with 6 Units	# of Sites with 7 Units	# of Sites with 8 Units
United Arab Emirates	1	0	0	1	0	0	0	0	0	0	0	0
United States	93	73	35	20	35	32	3	0	0	0	0	0
United Kingdom	15	14	6	1	6	5	0	1	0	0	0	0
TOTAL	440	392	181	48	133	80	11	28	2	8	3	1

Table 1-1Summary of NPP Sites by Country (as of June 2021) [3], [4]

THIS PAGE IS INTENTIONALLY LEFT BLANK

### Chapter 2 A Brief Introduction to Bayesian Networks

The purpose of this chapter is to introduce the reader to Bayesian networks (BNs) and provide an example, using a relatively simple BN, to illustrate the concepts behind how the calculations in subsequent chapters of this dissertation are performed. The information in this chapter is mainly based on the works by Kjærulff and Madsen [17] and Bensi et al. [18].

Bayesian networks are a type of probabilistic graphical model (also known as probabilistic network). The graphical nature of a BN is a powerful tool for expressing the dependence and independence relationships among the random variables in a problem domain. Typically, BNs are used to represent causal relationships between random variables; however, this is not a requirement. The inference in BNs is based on the well-established theory of probability calculus; therefore, BNs provide a rigorous mathematical framework for "deriving conclusions under uncertainty, where multiple sources of information and complex interaction patterns are involved" [17].

The work in this dissertation focuses on discrete BNs to remain consistent with the discretization typically used in the modeling and quantification of a seismic PRA and to avoid some of the limitations associated with using continuous BNs. In the following section, an introduction to the terminology of BNs (which will be used throughout this dissertation) is provided, along with a brief overview of quantitative inference using BNs.

#### 2.1 <u>Bayesian Networks</u>

A BN is a directed acyclic graph (DAG) consisting of nodes that represent random variables and directed links (also referred to as arcs or edges) that represent the probabilistic dependence

11

among the random variables. The BN defines a factorization of the joint probability distribution over the random variables [17]. A BN has a qualitative part, that is, the DAG and a quantitative part, a set of conditional probability functions. An example of a BN's qualitative part is shown in Figure 2-1.



Figure 2-1 Example of the qualitative part of a BN: the DAG

In Figure 2-1, the circles labeled  $X_i$ , i = 1, ..., 4, are the nodes and the arrows between pairs of nodes are the directed links. In BN terminology, the  $X_1$  and  $X_2$  nodes in Figure 2-1 are *root* nodes and there is a marginal probability distribution associated with each of them. BN terminology uses "family terminology" to describe the relationship between nodes. For example, node  $X_3$  is a *child* of  $X_1$  as indicated by the arrow going from  $X_1$  to  $X_3$ . Conversely,  $X_1$  is referred to as a *parent* of  $X_3$ . Similarly, the parents of node  $X_4$  are nodes  $X_1$  and  $X_2$ . There is a conditional probability distribution associated with each child node that specifies the probability that the child node is in a particular start given the states of the parent node(s).

The quantitative part of the BN is the representation of the joint probability distribution of all nodes whose general continuous representation is [17]:

$$f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f_{X_i | pa(X_i)}(x_i | pa(x_i))$$
(2.1)

where  $f_{X_1,X_2,...,X_n}(x_1, x_2, ..., x_n)$  is the joint probability density function of random variables  $X_1, X_2, ..., X_n; f_{X_i|pa(X_i)}(x_i|pa(x_i))$  is the conditional probability density function of random variable  $X_i$  given its parents, which are denoted by  $pa(X_i)$ ; and n is the number of random variables in the BN. The general discrete representation of the joint probability distribution is:

$$p_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = \prod_{i=1}^n p_{X_i | pa(X_i)}(x_i | pa(x_i))$$
(2.2)

where  $p_{X_1,X_2,...,X_n}(x_1, x_2, ..., x_n)$  is the joint probability mass function of random variables  $X_1, X_2, ..., X_n; p_{X_i|pa(X_i)}(x_i|pa(x_i))$  is the conditional probability mass function of random variable  $X_i$  given its parents, which are denoted by  $pa(X_i)$ ; and n is the number of random variables in the BN.

Applying Equation (2.1) to the BN shown in Figure 2-1 leads to the following joint probability density function:

$$f_{X_1, X_2, X_3, X_4}(x_1, x_2, x_3, x_4) = f_{X_4 | X_1, X_2}(x_4 | x_1, x_2) \cdot f_{X_3 | X_1}(x_3 | x_1) \cdot f_{X_1}(x_1) \cdot f_{X_2}(x_2)$$

$$(2.3)$$

In discrete form, the application of Equation (2.2) to the BN in Figure 2-1 leads to the joint probability mass function and can be written as:

$$p_{X_1,X_2,X_3,X_4}(x_1,x_2,x_3,x_4) = p_{X_4|X_1,X_2}(x_4|x_1,x_2) \cdot p_{X_3|X_1}(x_3|x_1) \cdot p_{X_1}(x_1) \cdot p_{X_2}(x_2)$$
(2.4)

where  $p_{X_1,X_2,X_3,X_4}(x_1, x_2, x_3, x_4) = P(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4)$  refers to the probability that the random variables  $X_1$  through  $X_4$  take on values (states)  $x_1$  through  $x_4$ , respectively.  $p_{X_4|X_1,X_2}(x_4|x_1,x_2) = P(X_4 = x_4|X_1 = x_1,X_2 = x_2)$  is the conditional probability that  $X_4 = x_4$ given that  $X_1 = x_1$  and  $X_2 = x_2$ .  $p_{X_1}(x_1) = P(X_1 = x_1)$  and  $p_{X_2}(x_2) = P(X_2 = x_2)$  are the marginal probabilities that  $X_1 = x_1$  and  $X_2 = x_2$ , respectively.

In principle, any joint (e.g.,  $f_{X_3,X_4}(x_3, x_4)$  or  $p_{X_3,X_4}(x_3, x_4)$ ), marginal (e.g.,  $f_{X_4}(x_4)$  or  $p_{X_4}(x_4)$ ), or conditional probability distribution (e.g.,  $f_{X_4|X_1,X_2}(x_4|x_1, x_2)$  or  $p_{X_4|X_1,X_2}(x_4|x_1, x_2)$ ) can be calculated after establishing the joint probability distribution over all nodes in the BN. Conceptually, inference in BNs consists of the efficient application of Bayes' rule:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$
(2.5)

where P(B|A) is the conditional probability of event *B* given the occurrence of event *A*, which is also known as the *likelihood function*<sup>3</sup> of the observed event *B*. P(A) is the marginal probability of event *A*, which is also known as the *prior probability* of event A and represents the belief about the probability of event *A* prior to getting any information about event *B*. P(B) is a normalization constant obtained using the Theorem of Total Probability. P(A|B) is the posterior probability of event *A* because it represents a belief about the probability of event *A* after observing the occurrence of event *B* [18]. Bayes' rule plays a central role in inference using BNs because "the probability of a cause can be inferred when its effect has been observed" [17].

<sup>&</sup>lt;sup>3</sup> In PRA, the likelihood function may be synonymously referred to as "aleatory model," stochastic model," or "probabilistic model" and "is used to model the process that gives rise to the data [or event *B* in the context of this discussion]" [19].

Specifically, inference in BNs is typically carried out using algorithms that efficiently apply Bayes' rule. The clustering algorithm, which was first proposed by Lauritzen and Spiegelhalter [20] and improved by Jensen et al. [21], is the prevailing inference algorithm in software packages, such as GeNIe [22], and it is based on message passing in a tree structure (junction tree) derived from the structure of the Bayesian network [17]. An explanation about the clustering algorithm and other inference algorithms is beyond the scope of this dissertation, but can be found in the aforementioned references.

#### 2.2 <u>Example Bayesian Network Calculations</u>

The example BN presented in this section was adapted from the work by Kjærulff and Madsen [17] to make it more relevant to NPP applications. Consider an emergency diesel generator (EDG) in a hypothetical NPP. In NPP PRAs, one of the typically modeled failure modes of an EDG is the "failure to start" (i.e., the failure of the EDG to start on demand). In this example, assume that the failure to start of an EDG can be caused by a defective air start motor or a malfunctioning speed switch circuit. The BN for this problem is shown in Figure 2-2, where the nodes "Air Start Motor," "Speed Switch Circuit," and "EDG" represent the possible states of the air start motor, speed switch circuit, and EDG, respectively. For simplicity, each node in Figure 2-2 is assume that one in every 1,000 air start motors is defective and that five out of 100 speed switch circuits malfunction.<sup>4</sup> This information is used to derive the (marginal) probability

<sup>&</sup>lt;sup>4</sup> These values are consistent with the example by Kjærulff and Madsen [17] and, therefore, do not represent actual EDG data from NPPs.
tables for the root nodes "Air Start Motor" (ASM) and "Speed Switch Circuit" (SwC); these are shown in Figure 2-2. In the case of the "EDG" (EDG) node, four conditional probability distributions need to be specified to account for the combination of the states of its parent nodes. Assume that if the air start motor is defective (Def), the EDG will fail to start. This is reflected by the probability of 1 assigned to state "Fails to Start" (FTS) whenever the air start motor is defective. If the air start motor is not defective (ND) and the speed switch circuit is malfunctioning (Mal), the EDG will fail to start in 10 out of 100 EDG demands. If the air start motor is not defective and the speed switch circuit is not malfunctioning (NM), then the EDG will fail to start in one out of 100 EDG demands; this 0.01 EDG failure-to-start probability captures other causes not explicitly considered in this problem. This is information results in the conditional probability table (CPT) for the "EDG" node, which is also shown in Figure 2-2.

CPT for Air Start Motor (ASM)				
Defective (Def)	0.001			
Not Defective (ND)	0.999			

Air Start Motor



CPT for Speed Switch Circuit (SwC)				
Malfunctioning (Mal)	0.05			
Not Malfunctioning (NM)	0.95			

CPT for EDG								
Air Start Motor (ASM)	De	fective (Def)	Not	Defective (ND)				
Speed Switch Circuit (SwC)	Malfunctioning (Mal)	Not Malfunctioning (Not Mal)	Malfunctioning (Mal)	Not Malfunctioning (Not Mal)				
Fails to Start (FTS)	1	1	0.1	0.01				
Does Not Fail to Start	0	0	0.9	0.99				

Figure 2-2 BN and CPTs for the EDG failure-to-start problem

# 2.2.1 Forward Inference

Suppose that we are interested in estimating the (marginal) probability that the EDG will fail to start (i.e., the state "Fails to Start" in the CPT for EDG in Figure 2-2) or P(EDG = FTS). In BN terminology, this is known as forward (predictive) inference because the information is propagating in the direction of the arrows. First, by applying Equation (2.2) to the BN shown in Figure 2-2, the discrete joint probability distribution of the EDG failure-to-start problem is as follows:

$$p(EDG, ASM, SwC) = p(EDG|ASM, SwC) \cdot p(ASM) \cdot p(SwC)$$
(2.6)

In Equation (2.6) and the subsequent discussion, we are omitting the subscripts EDG, ASM, and SwC for brevity. To obtain the marginal probability distribution of EDG, we need to "eliminate" the nodes ASM and SwC from the BN by summing (integrating) over all states of these nodes using the joint probability distribution. Mathematically, this operation is shown as Equation (2.7).

$$p(EDG) = \sum_{ASM} \sum_{SwC} p(EDG|ASM, SwC) \cdot p(ASM) \cdot p(SwC)$$
(2.7)

Applying Equation (2.7) to estimate the (marginal) probability that the EDG will fail to start (one possible state of the random variable EDG) is shown in Equation (2.8).

$$P(EDG = FTS) = P(EDG = FTS|ASM = Def, SwC = Mal) \cdot P(ASM = Def) \cdot P(SwC = Mal) + P(EDG = FTS|ASM = Def, SwC = NM) \cdot P(ASM = Def) \cdot P(SwC = NM) + P(EDG = FTS|ASM = ND, SwC = Mal) \cdot P(ASM = ND) \cdot P(SwC = Mal) + P(EDG = FTS|ASM = ND, SwC = NM) \cdot P(ASM = ND) \cdot P(SwC = NM)$$
(2.8)

All the required quantities to calculate Equation (2.8) can be found in the CPTs shown in Figure 2-2 and its application is as follows:

$$P(EDG = FTS) = 1 \cdot 0.001 \cdot 0.05 + 1 \cdot 0.001 \cdot 0.95 + 0.1 \cdot 0.999 \cdot 0.05 + 0.01 \cdot 0.999 \cdot 0.95$$
  
= 0.00005 + 0.00095 + 0.004995 + 0.0094905 = 0.0154855 \approx 0.0155 (2.9)

The result of Equation (2.9) means that the EDG is estimated to fail to start about 15 times in 1000 demands.

# 2.2.2 Backward Inference

Next, assume that we have observed that the EDG failed to start, which fixes the state of the node EDG to "Fail to Start." We enter this information as "evidence" in the BN; i.e., we specify the state of a node in the BN as a fixed value. We would like to know what might be causing this observation (evidence), that is, we will estimate P(ASM = Def | EDG = FTS) and

P(SwC = Mal|EDG = FTS), which corresponds to backward (diagnostic) inference because the information is propagating in the direction opposite of the direction of the arrows. The estimation of these conditional probabilities starts by applying Bayes' rule:

$$P(ASM = Def|EDG = FTS) = \frac{P(EDG = FTS|ASM = Def) \cdot P(ASM = Def)}{P(EDG = FTS)}$$
(2.10)

$$P(SwC = Mal|EDG = FTS) = \frac{P(EDG = FTS|SwC = Mal) \cdot P(SwC = Mal)}{P(EDG = FTS)}$$
(2.11)

The denominator of Equations (2.10) and (2.11) was estimated using Equations (2.8) and (2.9). The numerator of Equations (2.10) and (2.11) is estimated using the joint probability distribution based on the application of Equation (2.2) to the BN shown in Figure 2-2 and by summing over the states of nodes SwC and ASM, respectively. The application of this process is shown as Equations (2.12) and (2.13).

$$P(EDG = FTS|ASM = Def) \cdot P(ASM = Def)$$
  
=  $\sum_{SwC} P(EDG = FTS|ASM = Def, SwC) \cdot P(ASM = Def) \cdot P(SwC)$  (2.12)

$$P(EDG = FTS|SwC = Mal) \cdot P(SwC = Mal)$$

$$= \sum_{ASM} P(EDG = FTS|ASM, SwC = Mal) \cdot P(ASM) \cdot P(SwC = Mal)$$
(2.13)

Next, the expansion of Equations (2.12) and (2.13) and numerical substitution of the values in the CPTs of Figure 2-2 is as follows:

$$P(EDG = FTS|ASM = Def) \cdot P(ASM = Def)$$

$$= P(EDG = FTS|ASM = Def, SwC = NM) \cdot P(ASM = Def) \cdot P(SwC = NM)$$

$$+ P(EDG = FTS|ASM = Def, SwC = Mal) \cdot P(ASM = Def) \cdot P(SwC = Mal)$$

$$= 1 \cdot 0.001 \cdot 0.95 + 1 \cdot 0.001 \cdot 0.05 = 0.00095 + 0.00005 = 0.001$$

$$P(EDG = FTS|SwC = Mal) \cdot P(ASM = Mal)$$

$$= P(EDG = FTS|ASM = Def, SwC = Mal) \cdot P(ASM = Def) \cdot P(SwC = Mal)$$

$$+ P(EDG = FTS|ASM = ND, SwC = Mal) \cdot P(ASM = ND) \cdot P(SwC = Mal)$$

$$= 1 \cdot 0.001 \cdot 0.05 + 0.1 \cdot 0.999 \cdot 0.05 = 0.00005 + 0.004995 = 0.005045$$

$$(2.14)$$

Finally, we substitute the results from Equations (2.14) and (2.9) into Equation (2.10), and Equations (2.15) and (2.9) into Equation (2.11) to obtain the conditional probabilities of interest.

$$P(ASM = Def|EDG = FTS) = \frac{0.001}{0.0155} = 0.065$$
(2.16)

$$P(SwC = Mal|EDG = FTS) = \frac{0.005045}{0.0155} = 0.325$$
(2.17)

These results mean that a malfunctioning speed switch circuit is more likely to be the cause of the EDG failure to start than a defective air start motor.

THIS PAGE IS INTENTIONALLY LEFT BLANK

# Chapter 3 Extension of Probabilistic Seismic Hazard Analysis to Account for the Spatial Variability of Ground Motions at a Multi-Unit Nuclear Power Plant Hard-Rock Site<sup>5</sup>

# 3.1 Introduction

Nuclear power plant (NPP) sites typically consist of multiple reactors (units). However, the safety and risk analyses (e.g., probabilistic risk assessment (PRA)) at these sites are typically performed on a reactor-by-reactor basis. Under such analyses, when one reactor is being analyzed, the other(s) is (are) assumed to be in a safe, stable condition or otherwise unaffected. In limited cases, when multi-unit effects are considered, the assessment typically makes simplifying assumptions that may affect the resulting safety or risk insights.

The Fukushima Daiichi accident in March 2011 demonstrated the importance of accidents involving multiple units and highlighted the need for considering multi-unit accidents as part of PRAs. To properly characterize the risks at a multi-unit NPP site, it is necessary to account for the dependencies among the reactor units arising from the possibility that adverse conditions may affect multiple units concurrently. Moreover, these dependencies must be addressed in a realistic manner that does not alter important risk insights arising from the assessment.

<sup>&</sup>lt;sup>5</sup> This chapter was published in the journal *Structural Safety*, Vol. 85, p. 101958, Jul. 2020, https://doi.org/10.1016/j.strusafe.2020.101958.

Schroer and Modarres [23] proposed a classification schema for these dependencies to facilitate consideration of multi-unit accidents in a comprehensive manner. One of the dependencies in the classification schema pertains to the initiating events that may affect NPP sites. An earthquake is an example of an initiating event that will affect all units at the site. However, existing seismic PRAs typically neglect multi-unit effects or make the simplifying assumption that all units at the site experience the same ground motion. This assumption of perfect correlation in ground motion is inconsistent with data from dense seismic arrays, which show that there is spatial variability in the ground motion between closely-spaced locations during the same earthquake [24]–[27].

For the purposes of this study, we classify the separation distances between closely-spaced locations using the following scales: structural scale (i.e., up to about 150 meters (m)) [28], local/site scale (i.e., 150 m up to about 1 kilometer (km)) [29], and regional scale (i.e., above 1 km) [30], [31]. The structural, local/site, and regional scales can be viewed as relevant to a single-unit seismic PRA, multi-unit seismic PRA, and multi-site seismic PRA, respectively. In the United States, the separation distance between the units at the same NPP site ranges from 45 m to over 700 m [32]. In this study, we focus on spatial variability of ground motions at the structural and local/site scales, which can be modeled using similar methods and techniques. Conversely, ground motion spatial variability (or correlation) at the regional scale requires alternate conceptual formulations (e.g., [30], [31]).

To pragmatically reflect spatial variability of ground motion at the structure- and site-scales in a multi-unit seismic PRA (MUSPRA), this study proposes a practical method for obtaining the conditional probability distribution of ground motion at a non-reference location(s) on a site given the ground motion at a reference location on the same site. In particular, the method uses

the results of an existing probabilistic seismic hazard analysis (PSHA) of the ground motion at the reference location for a site (i.e., the site seismic hazard curve) in developing this conditional probability distribution. For a hard-rock site (e.g., see Enclosure 1 of [33] and [34]), we first develop a general framework for building the probabilistic model of spatial variability of ground motion, then we provide the implementation details for using the method, including an example case study.

# 3.2 <u>Background</u>

In this section, we introduce key background concepts of relevance to the seismic PRA. In particular, in Section 3.2.1, we first provide a summary of PSHA and how it is used to generate seismic hazard curves. Then, in Section 3.2.2, we provide a brief discussion of how seismic hazard curves are used in the seismic PRA to define the frequency of seismic initiating events. In Section 3.2.3, we shift focus and describe a process related to the PSHA referred to as disaggregation, which will be used in subsequent sections of this paper. Section 3.2.4 provides an introduction to the concept of spatial variability of ground motion.

# 3.2.1 Seismic Hazard Curve

One of the inputs to a seismic PRA is the site seismic hazard curve, which represents the annual exceedance frequency (AEF) of a particular ground motion parameter at a specific location for a specific spectral (or oscillator) frequency and damping (usually assumed as 5% of critical damping). PSHA is the process used to develop the seismic hazard curve. At its most basic level, the PSHA is composed of the five steps discussed by Baker [35].

In this study, we use spectral acceleration (SA) as the ground motion parameter of interest, and thus all subsequent mathematical expressions will use the notation *SA* to represent ground

23

motion intensity. However, unless otherwise noted, all expressions defined herein remain generally valid for other ground motion parameters. In practical PSHA applications, the probability distributions of the earthquake magnitude, source-to-site distance, and ground motion parameter of interest are discretized to facilitate computations; therefore, using the standard PSHA formulation, each point in the seismic hazard curve is obtained as [35]:

$$\lambda(SA > sa) = \sum_{i=1}^{n_{sources}} \lambda_{m,i} \sum_{j=1}^{n_M} \sum_{k=1}^{n_R} P(SA > sa|m_j, r_k) P(M_i = m_j, R_i = r_k)$$
(3.1)

In Equation (3.1),  $\lambda(SA > sa)$  is the rate of exceeding a specific value of spectral acceleration (*sa*);  $n_{sources}$  is the total number of earthquake sources (e.g., faults or areal sources) considered in the PSHA;  $\lambda_{m,i}$  is the rate of exceeding a minimum earthquake magnitude on source *i*, which can be obtained through the application of various earthquake recurrence relationships (e.g., the bounded Gutenberg-Richter recurrence relationship);  $n_M$  and  $n_R$  are the number of magnitude bins into which the magnitude probability distribution is discretized and the number of source-to-site distance bins into which the source-to-site distance probability distribution is discretized, respectively;  $P(SA > sa|m_j, r_k)$  is the conditional probability of exceeding *sa* given specific values of earthquake magnitude (*m*) and source-to-site distance (*r*) and it is obtained using a ground motion model (GMM); and  $P(M_i = m_j, R_i = r_k)$  is the joint probability mass function of the earthquake magnitude (*M*) and source-to-site distance (*R*) for earthquake source *i*, which is obtained through the discretization of the joint probability density function of  $M_i$  and  $R_i$ . Equation (3.1) is calculated multiple times, each time for a different value of spectral acceleration to be exceeded (i.e., *sa* in Equation (3.1)), to obtain the site seismic hazard curve.

Typically, for NPPs, the PSHA results are provided for a "control point" elevation at a particular site reference location (e.g., at the reactor building foundation, at bedrock) [12].

The GMMs used in PSHAs have the following general form [27], [30]:

$$\ln(SA(f)) = g(M, R, \theta, f) + \eta(f) + \epsilon(f)$$
(3.2)

where  $g(M, R, \theta, f)$  is the estimated natural logarithm of the ground motion parameter as a function of M, R, other explanatory variables  $\theta$  (e.g., style-of-faulting), and oscillator frequency f;  $\eta(f)$  is the inter-event variability (or residual); and  $\epsilon(f)$  is the intra-event variability (or residual). The inter- and intra-event variabilities are typically assumed to be independent and normally distributed with zero mean and standard deviations  $\sigma_{\eta}(f)$  and  $\sigma_{\epsilon}(f)$ , respectively. Therefore, the total standard deviation  $\sigma_T(f)$ , which is a measure of aleatory variability [36], [37], is equal to  $\left(\sigma_{\eta}^{2}(f) + \sigma_{\epsilon}^{2}(f)\right)^{0.5}$ . To account for the aleatory variability of the ground motion parameter, the general form of the GMM assumes that the natural logarithm of the ground motion parameter is normally distributed with mean equal to the value predicted by the GMM (i.e.,  $g(M, R, \theta, f)$ ) and standard deviation  $\sigma_T(f)$ . This aleatory variability comes from modeling a complex phenomenon (i.e., ground motion due to an earthquake) using relatively simple (compared to the phenomenon) GMMs [37]. In PSHAs for NPPs, it is typical to include several GMMs (each with its own measure of aleatory variability), among other aspects of ground motion modeling, to account for epistemic uncertainty arising from different technically defensible models. Conventionally, the epistemic uncertainty is systematically considered using logic trees [12] whereby each branch of the logic tree has a weight associated with it and the weights are assigned using the Senior Seismic Hazard Analysis Committee process [38]. Based

on the combination of branches in the PSHA logic tree, the mean seismic hazard curve, as well as the percentiles are estimated. The mean seismic hazard curve is then used in the quantification of the seismic PRA and the percentiles are used to quantify the uncertainty in the seismic PRA results.

3.2.2 Use of the Seismic Hazard Curve in the Quantification of a Seismic PRA The seismic hazard curve provides information regarding the frequency and severity of ground motion and serves as a key input to the seismic PRA. Additional key inputs to the seismic PRA include the seismic equipment list; seismic fragilities of the structures, systems, and components (SSCs); and system models (i.e., event trees and fault trees). A discussion about the technical details of these other seismic PRA elements is beyond the scope of this paper; EPRI report 3002000709 [12] provides detailed information on these topics.

The combination of information about the seismic hazard, seismic performance of SSCs and plant system performance provides the seismic risk estimates for the NPP (e.g., core damage frequency). In a typical seismic PRA using conventional PRA tools, the seismic hazard curve is discretized into several bins (e.g., see Table 3-3) defined by an upper and lower bound ground motion. The representative ground motion parameter value of a bin is generally taken as the geometric mean of the ground motion parameter values at the start and end of the bin; alternatively, it could be taken as the ground motion parameter value at the end of the bin [39], [40]. The occurrence frequency assigned to the bin is taken as the difference between the AEF at the start and end of the bin. The bin frequency is used as the initiating event frequency in the seismic event trees and the representative ground motion parameter value of a bin is used to determine the seismic fragilities (i.e., conditional failure probabilities) of the SSCs in the fault

26

trees. Under the limited current practice in multi-unit seismic PRA, all units at the NPP site experience the same ground motion parameter value; i.e., it is assumed that, in any given earthquake, the ground motion experienced at both units is in the same bin. In contrast to the current practice, our intent with this paper is to provide a method that allow risk analysts to consider different ground motion parameter values on other units at the NPP site; i.e., the ground motion at the different units may fall into different bins in a given earthquake. Consistent with current practice in seismic PRA, the discussion that follows uses discretized model formulations.

## 3.2.3 Disaggregation of the Seismic Hazard

In generating the seismic hazard curve, the PSHA aggregates over all combinations of earthquake characteristics on regional earthquake sources as well as their relative likelihoods of occurrence. However, this means that, when reviewing the results of a PSHA, it can be difficult to understand the earthquake scenarios that are dominant contributors to the seismic hazard at a site [41], [42]. Disaggregation (also referred to as deaggregation) provides a means of retrieving this earthquake scenario information. Because the method proposed in this study will use disaggregation, we briefly introduce the concept here. More comprehensive discussions about the disaggregation of the seismic hazard are provided by McGuire [41], Bazzuro and Cornell [42], and Baker [35].

Disaggregation is used to identify the conditional probability distribution of earthquake scenarios that can cause a ground motion that exceeds a particular level *sa*. These "scenarios" can be defined rather simply (e.g., as the occurrence of an earthquake of a particular magnitude) or in more complicated ways. For the method proposed in this study, we will use disaggregation to identify the conditional probability distribution of earthquake magnitude given the occurrence of

an earthquake with a ground motion that exceeds a particular level *sa*. This is commonly known as the magnitude disaggregation.

Bazzurro and Cornell [42] state that the magnitude disaggregation of the combined seismic hazard from all  $n_{sources}$  is usually obtained by accumulating the contribution of each earthquake magnitude to the total hazard (i.e.,  $\lambda(SA > sa)$ ) during the calculation of Equation (3.1). Then, at the end of the calculations, the accumulated contribution of each earthquake magnitude is divided by the value of the total hazard. This ratio is shown in Equation (3.3) and represents the conditional probability of experiencing an earthquake of magnitude *m* (or more specifically, an earthquake with magnitude in a particular discrete magnitude "bin") given the event that the spectral acceleration exceeds a certain level *sa*.

$$P(M = m|SA > sa) = \frac{\lambda(M = m, SA > sa)}{\lambda(SA > sa)}$$
(3.3)

The denominator of Equation (3.3) is the AEF value calculated using Equation (3.1). The numerator of Equation (3.3) is obtained by omitting the summation over M in Equation (3.1) [35].

# 3.2.4 Spatial Variability of Ground Motions

Fundamental in this work is the concept of spatial variability of ground motions. Zerva [26] defines the phrase "spatial variability of ground motions" as "the differences in the amplitude and phase of seismic ground motions recorded over extended areas." For the present application, we are interested in the amplitude spatial variability in terms of the spectral acceleration. We focus on the spatial variability of spectral acceleration because the GMMs used in nuclear

industry PSHAs provide estimates in terms of *SA* (and other response spectrum ordinates) and PSHA results are typically represented with respect to this quantity. While the spatial variability of ground motions has been explored with respect to measures in the frequency domain (e.g., [26], [43], [44]), such studies are not directly applicable to use in a seismic PRA that leverages the results of an existing PSHA providing seismic hazard curves on amplitude ground motion parameters. The works by Abrahamson and Sykora [24], Kawakami and Mogi [25], and Goda and Atkinson [27] are a few of the limited number of works related to the amplitude spatial variability using *SA* and peak ground acceleration (PGA), which we summarize in the following paragraphs and will leverage in the remainder of this paper.

Abrahamson and Sykora [24] analyzed the variability of *SA* over short distances (less than 100 m) using empirical recordings of seismic ground motions at dense arrays in California, Japan, and Taiwan. They defined the spatial variability of *SA* using an expression for the difference in the logarithm of spectral accelerations at closely-spaced locations during a single earthquake. Kawakami and Mogi [25] examined the spatial variability of PGAs (recorded at about the same epicentral distance) as function of separation distance. They analyzed data from the Chiba array, SMART-1 array, and the SIGNAL database. Kawakami and Mogi [25] state that the dispersion (spatial variability) of PGAs can be represented by either the mean of the PGA ratio ( $\mu_{R'}$ ) or standard deviation of the difference of the logarithm of the two PGAs ( $\sigma_{P'}$ ).

The work by Goda and Atkinson [27] focuses on the correlation of the ground motion residuals. The use of residuals implies that a specific GMM is used to estimate the seismic hazard. However, consistent with the goals of this study, our proposed method starts with the results of an existing PSHA, which aggregates over a number of GMMs. Disaggregating the results to obtain specific GMMs is possible [45], but not practical. Also, Goda and Atkinson [27] noted that they achieved "reliable data coverage only for separation distances greater than 1 km" and advised careful consideration of using their proposed model for distances less than 1 km, which is the separation distance range of interest in our work.

Based on the review of these models of spatial variability of ground motion amplitudes, we will leverage the works by Abrahamson and Sykora [24] and Kawakami and Mogi [25] in the remainder of this paper, which we will abbreviate as A&S and K&M, respectively. We note that these models are conceptually similar. To emphasize this conceptual similarity, Table 3-1 compares and summarizes these model formulations and provides the associated measure of spatial variability of ground motion and its distribution.

	Models Used in this Study	
Authors	Abrahamson and Sykora [24]	Kawakami and Mogi [25]
Spatial variability of ground motion measure	$\Delta SA_{ijk}(f) = \ln[SA_{ij}(f)] - \ln[SA_{ik}(f)]$ where: • $SA_{ij}(f)$ and $SA_{ik}(f)$ are the average horizontal component of the acceleration response spectrum for the <i>j</i> <sup>th</sup> and <i>k</i> <sup>th</sup> stations, respectively, during the <i>i</i> <sup>th</sup> earthquake at the oscillator frequency <i>f</i> • $\Delta SA_{ijk}(f)$ is the difference between the logarithms of the spectral acceleration values of the <i>j</i> <sup>th</sup> and <i>k</i> <sup>th</sup> stations (which are separated by a distance $\xi$ ) from the <i>i</i> <sup>th</sup> earthquake	<ul> <li>P' = ln(PGA<sub>2</sub>) - ln(PGA<sub>1</sub>)</li> <li>where:</li> <li>PGA<sub>1</sub> and PGA<sub>2</sub> are the peak ground accelerations at two closely-spaced locations during the same earthquake, respectively</li> <li>P' is the difference between the logarithms of the peak ground acceleration values at two-closely spaced locations during the same earthquake</li> </ul>
Distribution of the spatial variability of ground motion measure	$\Delta SA_{ijk}(f) \sim N\left(\mu_{\Delta SA_{ijk}(f)}, \sigma_{\Delta SA_{ijk}(f)}\right)$ where: • $\mu_{\Delta SA_{ijk}(f)}$ is the mean of $\Delta SA_{ijk}(f)$ and assumed to be zero • $\sigma_{\Delta SA_{ijk}(f)}$ is the standard deviation of $\Delta SA_{ijk}(f)$ and given by the expression below	<ul> <li>P'~N(μ<sub>P'</sub>, σ<sub>P'</sub>)</li> <li>where:</li> <li>μ<sub>P'</sub> is the mean of P' and, by definition, is zero</li> <li>σ<sub>P'</sub> is the standard deviation of P' and given by the expression below<sup>2</sup></li> </ul>
Standard deviation of the	$\sigma_{\Delta SA_{ijk}}(M,\xi,f) = c_1(f,M)(1 - \exp\{-\xi c_2(f)\})$ where:	$\sigma_{P'}(\xi) = \begin{cases} 2 \times 10^{-4} \cdot \xi + 0.17, & 15 \ m \le \xi \le 230 \ m \\ 5 \times 10^{-5} \cdot \xi + 0.36, & 300 \ m \le \xi \le 2000 \ m \\ \end{cases}$ where:

Table 3-1Comparison and Explanation of the Spatial Variability of Ground Motion<br/>Models Used in this Study

Table 3-1	Comparison and Explanation of the Spatial Variability of Ground Motion
	Models Used in this Study

Authors Abraha	mson and Sykora [24]	Kawakami and Mogi [25]
distribution of the spatial variability of ground measure $\xi$ is the separation stations $j$ ar $f$ is the osc $M$ is the ear $c_1(f, M)$ an estimated b and a maxim	aration distance between and k in meters llator frequency in hertz (Hz) thquake magnitude <sup>1</sup> d $c_2(f)$ are coefficients y A&S [24] using regression num likelihood approach	<ul> <li>ξ is the separation distance between the two closely-spaced locations in meters<sup>3</sup></li> </ul>

<sup>1</sup> The method proposed in this study will address the magnitude dependence through the use of the magnitude disaggregation, as described in Sections 3.3.3 and 3.4.2.

<sup>2</sup> Kawakami and Mogi found an "almost linear relationship" between  $\sigma_{P'}$  and the logarithm of the separation distance; however, they did not provide an expression of such relationship. The expression shown in this table was derived by using Figure 7 in K&M [25] and taking the geometric mean of the  $\sigma_{P'}$  values in the East-West and North-South directions.

<sup>3</sup> The linear relationships in the expression above do not comply with the theoretical insight that at  $\xi = 0$  there is no spatial variability (i.e.,  $\sigma_{P'} = 0$ ) because of the limitations of the approach used to define the relationship. Therefore, we do not use the expression above outside the stated separation distance ranges.

# 3.3 Framework for Modeling Ground Motion Spatial Variability

In this section, the proposed framework for modeling spatial variability of ground motion to support MUSPRA is presented using probabilistic graphical models; specifically, Bayesian networks. As discussed in our previous work [46], Bayesian networks (BNs) are used because of their graphical structure and transparency, as well as the manner in which they facilitate modeling probabilistic dependencies. While we have elected to use BNs as a mechanism for representing the mathematical formulation developed in this paper, the overall formulation is general in its application and does not require the use of BNs for performing probabilistic calculations.

Figure 3-1 shows an example of a BN. In a BN, nodes (ovals) represent random variables (RVs), and arrows indicate a probabilistic dependence between the connected nodes. The direction of the arrow typically represents a causal relationship, though is not a requirement of the BN

modeling framework. In Figure 3-1,  $X_1$  and  $X_2$  are referred to as the *parents* of node *Y*. Conversely, node *Y* is referred to as the *child* of nodes  $X_1$  and  $X_2$ . Each node is associated with a conditional probability table (probability mass function) that provides the probability of the respective RV being in a particular state given the state of its parents. Nodes without parents are described by a marginal probability mass function and are referred to as root nodes. The reader is referred to the works by Kjærulff and Madsen [17] for a comprehensive introduction to BNs and Bensi et al. [18] for a basic introduction to the use of BNs in the context of seismic risk assessment applications.



Figure 3-1 Example Bayesian network

3.3.1 General Bayesian Network to Model Ground Motion Spatial Variability at Hard-Rock Sites

Figure 3-2(a) shows the proposed BN for modeling spatial variability in ground motion. The RVs in Figure 3-2(a) are defined in Table 3-2. In addition, Figure 3-2(b) and (c) provide a companion graphic illustrating the physical meaning of most of the variables (nodes) shown in the BN under a conceptual model of ground motion propagation in which seismic energy radiates from the earthquake source to the hypothetical site (i.e., seismic waves), traveling mostly through rock [47].

The illustration of the variables provided in Figure 3-2(b) is an "overhead (top) view" and shows a hypothetical site with two units along with an earthquake source, which is associated with the

RVs representing the magnitude and location, and the associated source-to-site distance. These RVs are reflected by the nodes *Source*, *M*, *Loc*, and *R*, respectively, in Figure 3-2(a). The node *Source* is a parent of nodes *M* and *Loc*, which represents the source-specific nature of their respective probability distributions. The node *R* is a child of the node *Loc* indicating that the probability distribution of source-to-site distance is dependent on the location of the earthquake. While not shown in this BN, we could also include a link between node *M* and *R* to capture any dependence between those two quantities (e.g., to capture finite fault rupture effects).

Next, as we progress along the BN in Figure 3-2(a), we make use of Figure 3-2(c), which shows the rest of the variables defined in Table 3-2 along with the same hypothetical site shown in Figure 3-2(b). Once the seismic waves arrive to the site, the waves first travel through rock up to the bedrock control point. The resulting ground motion is known as the bedrock ground motion [47] and is denoted with the superscript r (e.g.,  $SA_{CP}^r$ ) in Figure 3-2(a) and Figure 3-2(c). Under the framework proposed in this paper, the probability distribution of the bedrock ground motion at the control point,  $SA_{CP}^r$ , is estimated as a function of earthquake characteristics using a GMM. That is,  $SA_{CP}^r$  is defined as a function of a GMM-predicted value  $\overline{SA_{CP}^r}$  and the associated interand intra-event variabilities ( $\eta$  and  $\epsilon$ ). We note that, when considering regional-scale assessments, there is variability in  $\eta$  and  $\epsilon$  across the spatial domain (e.g., [30], [31]); however, this effect is not significant at the site scale and thus, in this framework,  $\eta$  and  $\epsilon$  are treated as common to both units.



Figure 3-2 (a) Ground motion amplitude spatial variability BN for a single rock site with two units. Illustration of the variables in the BN shown in Figure 3-2(a): (b) Top view (c) Elevation view

RV	Definition
Source	Earthquake source.
М	Earthquake magnitude (taken as the moment magnitude in this study).
Loc	Earthquake location.
R	Distance from the earthquake source to the site (i.e., source-to-site distance).
$\overline{SA_{CP}^r}$	Ground motion at the bedrock control point predicted using a GMM.
$\eta$ and $\epsilon$	Inter-event and intra-event variabilities, respectively.
S Ar	Ground motion at the bedrock control point after incorporating the inter- and
SA <sub>CP</sub>	intra-event variabilities (see Equation (3.2)).
$\xi_{1,CP}$ and $\xi_{2,CP}$	Separation distances (in meters) between Unit 1 and the control point, and Unit 2
	and the control point, respectively.
$\Delta SA_{1,CP}$ and	Ground motion amplitude spatial variability model for Units 1 and 2,
$\Delta SA_{2,CP}$	respectively, each with respect to the site control point using the A&S [24]
_,	formulation. These two nodes would be $P'_{1,CP}$ and $P'_{2,CP}$ using the K&M [25]
	formulation.
$SA_1^r$ and $SA_2^r$	Ground motion at bedrock underneath Units 1 and 2, respectively, after
_	considering the ground motion amplitude spatial variability model.

Table 3-2Definition of the RVs in Figure 3-2(a)

The portion of the BN (Figure 3-2(a)) that includes the node  $SA_{CP}^{r}$  and its "ancestors" (i.e., its parents, their parents, and so on) represents the variables contained in a conventional PSHA. In subsequent portions of the paper, we will describe how we will use the results of an existing PSHA rather than explicitly including in the BN the nodes representing the earthquake characteristics and GMM outputs.

As discussed previously, during any earthquake, the ground motion at the locations of the reactor units will generally differ from the ground motion at the control point. The model proposed herein for hard-rock sites considers a primary source of variability in ground motion between the control point and the individual units: amplitude spatial variability. In this study, the models proposed by A&S [24] and K&M [25] (see Section 3.2.4) will be used to model ground motion amplitude spatial variability; however, the framework proposed herein can be adapted as other models become available.

In particular, the BN captures the ground motion amplitude spatial variability via the nodes  $\Delta SA_{1,CP}$  and  $\Delta SA_{2,CP}$  with the A&S [24] formulation, which are a function of the separation distances  $\xi_{1,CP}$  and  $\xi_{2,CP}$  and earthquake magnitude *M*. If the K&M [25] formulation is used, the amplitude spatial variability is captured by nodes  $P'_{1,CP}$  and  $P'_{2,CP}$ , which are also a function of the separation distances  $\xi_{1,CP}$  and  $\xi_{2,CP}$ , but not a function of the earthquake magnitude *M*; this is the reason why the link (arrow) between nodes *M* and  $\Delta SA_{1,CP}$  (and  $\Delta SA_{2,CP}$ ) in Figure 3-2(a) is a dashed line. The bedrock ground motion underneath each unit (i.e.,  $SA_1^r$  and  $SA_2^r$ ) is then defined as a function of the bedrock ground motion at the control point ( $SA_{CP}^r$ ) and the amplitude spatial variability ( $\Delta SA_{1,CP}$  and  $\Delta SA_{2,CP}$  or  $P'_{1,CP}$  and  $P'_{2,CP}$ ).

Using the probabilistic framework shown in Figure 3-2(a), we can determine the conditional probability distribution of the ground motion hazard at a specific unit (i.e., non-reference ground motion hazard) given the ground motion hazard at the control point (i.e., reference ground motion hazard). As described in our previous work [32], these conditional probability distributions provide input to the MUSPRA and facilitate the consideration of spatial variability of ground motion at a NPP site. Further details are provided below.

# 3.3.2 Streamlined Bayesian Network to Model Ground Motion Spatial Variability at a Hard-Rock Site

The BN in Figure 3-2(a) is a general case for a two-unit hard-rock site assuming that a PSHA was performed for a control point that is located somewhere on the site. In some instances, the control point may coincide with one of the units at the site. If this is the case, the BN contains less variables (nodes) and the structure of the model becomes more streamlined than the general BN presented in Section 3.3.1. This facilitates the explanation on how the BN in Figure 3-2(a) can be used in a MUSPRA. Therefore, without loss of generality, we assume that the control point coincides with Unit 1. That is, in subsequent descriptions, we treat Unit 1 as the location where the ground motion hazard has been characterized using an existing PSHA. We will refer to the unit where the PSHA has been performed as the "reference unit" and the other unit as the "non-reference unit."

Also, for simplicity in the presentation of the derivation that follows, we consider a single earthquake source in the subsequent discussion; however, the framework is equally applicable to multiple earthquake sources. Finally, for clarity in the BN, we note that some parent nodes can be integrated into their children and eliminated from the BN. Specifically in Figure 3-2(a),  $\eta$ ,  $\epsilon$ , and  $\overline{SA_{CP}^r}$  can be integrated into  $SA_{CP}^r$ , and  $\xi_{1,CP}$  and  $\xi_{2,CP}$  can be integrated into  $\Delta SA_{1,CP}$  and  $\Delta SA_{2,CP}$  (or  $P'_{1,CP}$  and  $P'_{2,CP}$ ), respectively, and thus eliminated from the BN. This leads to a streamlined BNs, which are shown in Figure 3-3(a) and (b) for the A&S [24] and K&M [25] formulations, respectively, along with an illustration of the variables in Figure 3-3(c).



Figure 3-3 Streamlined BNs to model ground motion spatial variability using the (a) A&S [24] and (b) K&M [25] formulations. (c) Illustration of the variables in the streamlined BNs (*M* and *R* are not shown in the illustration of the variables).

Similar to the illustration of variables in Figure 3-2(c) and described in Section 3.3.1, the illustration of variables in Figure 3-3 includes a node representing the ground motion hazard at the reference unit  $SA_1^r$ , with a probability distribution defined as a function of M and R based on a GMM. The RV representing the bedrock ground motion at the non-reference unit  $(SA_2^r)$  is defined as a function of the bedrock ground motion hazard underneath the reference unit  $(SA_1^r)$  and the amplitude spatial variability factor ( $\Delta SA_{1,2}$  or  $P'_{1,2}$ ). We will use these "streamlined" BNs in all subsequent descriptions.

# 3.3.3 MUSPRA Input for a Hard-Rock Site

In this section, we describe the development of the MUSPRA input that reflects the spatial variability of ground motions across an NPP hard-rock site. In general, we are interested in estimating the conditional probability distribution of a non-reference ground motion hazard (i.e., Unit 2) given the reference ground motion hazard (i.e., Unit 1) as:

$$p_{SA_2^r|SA_1^r}(sa_2^r|sa_1^r) = \frac{p_{SA_1^r,SA_2^r}(sa_1^r,sa_2^r)}{p_{SA_1^r}(sa_1^r)}$$
(3.4)

In the above expression,  $p_{SA_1^r,SA_2^r}(sa_1^r, sa_2^r)$  represents the joint probability mass functions of the reference and non-reference ground motion hazards for rock sites. The quantity  $p_{SA_1^r}(sa_1^r)$  is the marginal probability mass functions of the reference ground motion hazard for rock sites. In particular, we describe how the conditional probability distribution of the ground motion at Unit 2 given the ground motion at Unit 1 is developed for a rock site using two different models for ground motion spatial variability [24], [25]. The process for developing the conditional probability distribution for a rock site is reflected by the BNs and accompanying graphic in Figure 3-3. In the subsequent figures in this section, we omit the separation distance between the units,  $\xi_{1,2}$ , because for a site-specific analysis, this variable would be known (i.e., it is not a RV). For the ease of notation in the subsequent discussion, we will write  $p_{SA_1^r}(sa_1^r)$  as  $P(SA_2^r|SA_1^r)$  for a conditional probability mass function,  $p_{SA_1^r,SA_2^r}(sa_1^r, sa_2^r)$  as  $P(SA_1^r, SA_2^r)$  for a joint probability mass function, and  $p_{SA_1^r}(sa_1^r)$  as  $P(SA_1^r)$  for a marginal probability mass function.

The joint probability mass functions of the variables in the BNs shown in Figure 3-3 (a) and Figure 3-3(b) are

$$P(M, R, \Delta SA_{1,2}, SA_1^r, SA_2^r) = P(SA_2^r | \Delta SA_{1,2}, SA_1^r) P(\Delta SA_{1,2} | M) P(SA_1^r | M, R) P(M) P(R)$$
(3.5)

$$P(M, R, P'_{1,2}, SA^r_1, SA^r_2) = P(SA^r_2 | P'_{1,2}, SA^r_1) P(P'_{1,2}) P(SA^r_1 | M, R) P(M) P(R)$$
(3.6)

Next, we sum over the variable *R* from Equation (3.5) and over the variables *M* and *R* from Equation (3.6), which corresponds to the elimination of the node(s) from the respective BNs. This leads to the following expressions and the BNs shown in Figure 3-4(a) and (c):

$$P(M, \Delta SA_{1,2}, SA_1^r, SA_2^r) = P(SA_2^r | \Delta SA_{1,2}, SA_1^r) P(\Delta SA_{1,2} | M) P(M) \sum_R P(SA_1^r | M, R) P(R)$$

$$= P(SA_2^r | \Delta SA_{1,2}, SA_1^r) P(\Delta SA_{1,2} | M) P(M) P(SA_1^r | M)$$

$$P(P_{1,2}', SA_1^r, SA_2^r) = P(SA_2^r | P_{1,2}', SA_1^r) P(P_{1,2}') \sum_M \sum_R P(SA_1^r | M, R) P(M) P(R)$$

$$= P(SA_2^r | P_{1,2}', SA_1^r) P(P_{1,2}') P(SA_1^r)$$
(3.7)
(3.7)
(3.7)
(3.7)



Figure 3-4 (a) BN corresponding to Equation (3.7). (b) BN corresponding to Equation (3.9). (c) BN corresponding to Equation (3.8).

In the case of the BN using the K&M [25] formulation, Figure 3-4(c) and Equation (3.8), we have the variables we need to obtain the conditional probability distribution of the ground motion hazard at the non-reference unit. In the case of the BN using the A&S [24] formulation, we use Bayes' rule on the term  $P(SA_1^r|M)$  of Equation (3.7), and obtain Equation (3.9) and the BN shown in Figure 3-4(b). The effect of using Bayes' rule is reversing the link between *M* and  $SA_1^r$  in Figure 3-4(a), which then becomes Figure 3-4(b). The joint probability distribution over the resulting BN is then given as:

$$P(M, \Delta SA_{1,2}, SA_1^r, SA_2^r) = P(SA_2^r | \Delta SA_{1,2}, SA_1^r) P(\Delta SA_{1,2} | M) P(M | SA_1^r) P(SA_1^r)$$
(3.9)

Equations (3.8) and (3.9) provide all the necessary information to compute the probability distributions in which we are interested. Moreover, as outlined below, all required quantities can be developed using existing PSHA results. In particular:

- $P(SA_1^r)$  is obtained through discretization of the reference rock seismic hazard curve.
- $P(M|SA_1^r)$  is obtained from the magnitude disaggregation at the reference unit.
- $P(\Delta SA_{1,2}|M)$  is obtained from the A&S [24] ground motion amplitude spatial variability model for a rock site.
- $P(SA_2^r | \Delta SA_{1,2}, SA_1^r)$  is an indicator function based on the definition of  $\Delta SA_{1,2}$  (see Table 3-1, i.e.,  $SA_2^r = SA_1^r \exp(\Delta SA_{1,2})$ ).  $P(SA_2^r | \Delta SA_{1,2}, SA_1^r)$  is equal to one when the value of  $SA_2^r$ , given the values of  $\Delta SA_{1,2}$  and  $SA_1^r$ , falls within the interval (bin) of  $SA_2^r$  being considered and zero otherwise, that is,

$$P(SA_{2}^{r}|\Delta SA_{1,2}, SA_{1}^{r}) = \begin{cases} 1, & (SA_{2}^{r})_{low} \le SA_{1}^{r} \exp(\Delta SA_{1,2}) < (SA_{2}^{r})_{high} \\ 0, & otherwise \end{cases}$$
(3.10)

- *P*(*P*'<sub>1,2</sub>) is obtained from the K&M [25] ground motion amplitude spatial variability model.
- $P(SA_2^r|P_{1,2}, SA_1^r)$  is similar to  $P(SA_2^r|\Delta SA_{1,2}, SA_1^r)$ , where  $P_{1,2}'$  substitutes  $\Delta SA_{1,2}$  in Equation (3.10).

To estimate the conditional probability distribution of a non-reference ground motion hazard (i.e., Unit 2) given the reference ground motion hazard (i.e., Unit 1) using Equation (3.4), we first compute the joint probability mass function of the reference and non-reference ground motion hazards from the joint probability mass function of all the variables contained in the BNs shown in Figure 3-4(b) and (c), respectively, that is,

$$P(SA_{1}^{r}, SA_{2}^{r}) = \sum_{M} \sum_{\Delta SA_{1,2}} P(SA_{2}^{r} | \Delta SA_{1,2}, SA_{1}^{r}) P(\Delta SA_{1,2} | M) P(M | SA_{1}^{r}) P(SA_{1}^{r})$$
(3.11)

$$P(SA_1^r, SA_2^r) = \sum_{P_{1,2}'} P(SA_2^r | P_{1,2}', SA_1^r) P(P_{1,2}') P(SA_1^r)$$
(3.12)

Next, the results from Equations (3.11) and (3.12) are divided by the marginal probability mass function of the reference ground motion hazard. This leads to Equations (3.13) and (3.14), which use the A&S [24] and K&M [25] formulations, respectively, and establish our proposed models for specifying the conditional probability distribution of the ground motion hazard at a non-reference location on an NPP hard-rock site given the ground motion hazard at the reference location on the same NPP hard-rock site.

$$P(SA_{2}^{r}|SA_{1}^{r}) = \frac{\sum_{M} \sum_{\Delta SA_{1,2}} P(SA_{2}^{r}|\Delta SA_{1,2}, SA_{1}^{r}) P(\Delta SA_{1,2}|M) P(M|SA_{1}^{r}) P(SA_{1}^{r})}{P(SA_{1}^{r})}$$
(3.13)  
$$P(SA_{2}^{r}|SA_{1}^{r}) = \frac{\sum_{P_{1,2}^{r}} P(SA_{2}^{r}|P_{1,2}^{r}, SA_{1}^{r}) P(P_{1,2}^{r}) P(SA_{1}^{r})}{P(SA_{1}^{r})}$$
(3.14)

$$P(SA_1^r) = P(SA_1^r)$$
(5.14)

The implementation of Equations (3.13) and (3.14) in the BNs is illustrated in Section 3.4.2 using a case study.

## 3.4 <u>Example Application of the Proposed Method</u>

In this section, we apply the method described in Section 3.3.3 to a hypothetical rock site and assume a separation distance between units of 100 m. The seismic hazard curve and disaggregation results are calculated using software that is available to the U.S. Nuclear Regulatory Commission (NRC) staff. The site used in this example corresponds to the Central Illinois site used in the demonstration hazard calculations in NUREG-2115 [48] and the Electric Power Research Institute (EPRI) GMM Review Project [49]. These two projects selected the Central Illinois site (latitude 40.0, longitude -90.0) because of the "[h]azard from [the] New Madrid seismic zone and paleo-earthquake zones in central Illinois" [48], [49]. The seismic

hazard shown here is not reflective of the hazard at any nuclear facility, and results are presented here solely to illustrate the implementation of the proposed framework.

# 3.4.1 PSHA Results for the Central Illinois Site

The mean seismic hazard curve at bedrock for the reference location at the Central Illinois site is shown in Figure 3-5. The ground motion parameter used is PGA, which we assume corresponds to *SA* at 100 Hz. We elect to use this ground motion parameter because PGA is the typical ground motion parameter used in seismic PRAs to describe the seismic demand to the NPP SSCs. In the subsequent discussion, we will use *PGA* and *SA* interchangeably.



Figure 3-5 Mean rock hazard curve at the Central Illinois Site

Results for the magnitude disaggregation were obtained for the following return periods in years: 2,500; 10,000; 50,000; 100,000; 250,000;  $1 \times 10^6$ ; and  $1 \times 10^7$ . These return periods correspond to the probability of exceeding the following PGA values: 0.08g, 0.18g, 0.4g, 0.56g, 0.86g, 1.5g, and 3.32g, respectively. The return periods and corresponding PGA values were selected to coincide with the start and end of the PGA bins used in the seismic PRA. Specifically, the seismic hazard curve shown in Figure 3-5 was discretized as shown in columns 1–5 in Table 3-3.

To ensure consistency with general practice, the PGA discretization bins used in this case study were adapted from a representative seismic PRA summary report for a NPP [50]. The bin frequency values were calculated using the seismic hazard curve in Figure 3-5.

Table 5-5 Discretization of Seisine Hazard Curve										
Seismic Bin Bin		smic Bin Bin Bin		Bin	Magnitude					
Initiator	Start	End	Representative	Frequency	Disaggregation					
Bin	PGA (g)	PGA (g)	PGA (g)	(yr <sup>-1</sup> )	Used in Bin					
BIN-1	0.08	0.18	0.12	3.67E-4	$P(M = m   0.08g < PGA \le 0.18g)$					
BIN-2	0.18	0.4	0.27	7.6E-5	$P(M = m   0.18g < PGA \le 0.4g)$					
BIN-3	0.4	0.56	0.47	9.58E-6	$P(M = m   0.4g < PGA \le 0.56g)$					
BIN-4	0.56	0.86	0.69	5.92E-6	$P(M = m   0.56g < PGA \le 0.86g)$					
BIN-5	0.86	1.5	1.14	2.89E-6	$P(M = m   0.86g < PGA \le 1.5g)$					
BIN-6	1.5	3.32	2.23	9.19E-7	$P(M = m   1.5g < PGA \le 3.32g)$					
BIN-7	3.32		3.32	1.01E-7	P(M = m   PGA > 3.32g)					

Table 3-3Discretization of Seismic Hazard Curve

3.4.2 Implementation of the Ground Motion Spatial Variability Model

BNs are used to implement the ground motion spatial variability model. In this section, we discuss how to obtain the conditional probability tables (CPTs) for each of the nodes in the BNs shown in Figure 3-4(b) and (c). As noted previously, while we use BNs as a preferred communication and calculation mechanism in this paper, the approach can be implemented more generally without using BNs as the calculation framework (i.e., by performing calculations directly using the equations presented in previous sections of this report and any general-purpose computing software).

# $3.4.2.1 \quad SA_1^r$

To generate the conditional probability table (conditional probability mass function (PMF)) assigned to node  $SA_1^r$ , the domain of  $SA_1^r$  is discretized into a set of contiguous bins consistent

with the discretization of the seismic hazard curve used in the seismic PRA, as described in Section  $3.2.2.^6$  This is reflected in columns 1–5 in Table 3-3.

The binning in Table 3-3 reflects typical seismic PRA practice in which the seismic PRA does not consider all possible ground motion values. Particularly, ground motion values at or below a certain threshold of engineering significance are excluded because they are not expected to cause damage to the NPP SSCs of relevance to the seismic PRA. In this example, that threshold is 0.08g and is the ground motion value that defines the start of the first PGA bin. However, for probabilistic consistency, the columns of the CPTs in a BN are required to sum up to 1.0. Consequently, for implementing our proposed method, we condition the node  $SA_1^r$  and all the calculations on the occurrence of an earthquake of engineering significance (EES). The probability of the EES is numerically equivalent to the AEF of the PGA at the start of the first bin; i.e.,  $P(EES) = P(PGA > 0.08g) \approx \lambda(PGA > 0.08g)$ . The conditional probability that the PGA value is in BIN-1 given an EES is then:<sup>7</sup>

<sup>&</sup>lt;sup>6</sup> We note that, in the context of seismic PRA, the Poisson distribution is typically assumed to describe the number of occurrences of earthquakes of interest in a given time interval; therefore, the probability of at least one event (earthquake) in time *t* is  $P(at \ least \ 1 \ event) = 1 - \exp(-\lambda t)$ . Using t = 1yr, and the typical values of  $\lambda(SA > sa)$  in a seismic PRA, which range from  $10^{-7}$  per year to  $10^{-3}$  per year, the resulting P(SA > sa) is numerically equivalent to  $\lambda(SA > sa)$ .

<sup>&</sup>lt;sup>7</sup> The change from probability to frequency is done because in seismic PRA these quantities are numerically equivalent.

$$P(0.08g < PGA \le 0.18g | PGA > 0.08g) = \frac{P(0.08g < PGA \le 0.18g \cap PGA > 0.08g)}{P(PGA > 0.08g)}$$
$$= \frac{P(0.08g < PGA \le 0.18g)}{P(PGA > 0.08g)}$$
$$= \frac{\lambda(PGA > 0.08g) - \lambda(PGA > 0.18g)}{\lambda(PGA > 0.08g)}$$
(3.15)

Repeating this calculation for the remaining PGA bins and using  $\lambda(PGA > 0.08g) =$ 

 $4.62 \times 10^{-4}$  /yr results in the CPT for  $SA_1^r$ , which is shown in Table 3-4.

Table 3-4	CPT fo	or Node <b>S</b> A	$\mathbf{I}_1^r$
Seismic Initiator Bin	pga <sub>start</sub> (g)	pga <sub>end</sub> (g)	$P(pga_{start} < PGA \le pga_{end}   PGA > 0.08g)$
BIN-0	0	0.08	0
BIN-1	0.08	0.18	$7.94 \times 10^{-1}$
BIN-2	0.18	0.4	$1.64 \times 10^{-1}$
BIN-3	0.4	0.56	$2.07 \times 10^{-2}$
BIN-4	0.56	0.86	$1.28 \times 10^{-2}$
BIN-5	0.86	1.5	$6.26 \times 10^{-3}$
BIN-6	1.5	3.32	$1.99 \times 10^{-3}$
BIN-7	3.32		$2.19 \times 10^{-4}$

We note that any numerical results obtained from calculations under the condition that an earthquake of engineering significance has occurred must be multiplied by  $\lambda(PGA > 0.08g) = 4.62 \times 10^{-4}$  /yr (i.e., the frequency of occurrence of the EES) to obtain the "unconditional rates" that would be used as input to the seismic PRA. Further, we note that we have opted to include a "BIN-0" in Table 3-4 even though this corresponds to an earthquake of non-engineering significance and is not included in the seismic PRA. The rationale for introducing the first row of Table 3-4 (BIN-0) is provided in Section 3.4.2.4.

# $3.4.2.2 \qquad M|SA_1^r|$

The CPT (conditional PMF) for node *M* is generated through the use of the magnitude disaggregation results, which provide the conditional distribution of *M* for each value of  $SA_1^r$  and are a standard output of the PSHA performed for the control point (reference location). In Table 3-3, the quantities in the column labeled "Magnitude Disaggregation Used in Bin" are not directly obtained from the PSHA results—except the last quantity, P(M = m|PGA > 3.32g). However, these quantities can be calculated using the PSHA results as follows:

$$P(M = m_i | sa_{start} < SA \le sa_{end})$$
  
= 
$$\frac{P(M = m_i | SA > sa_{start})\lambda(SA > sa_{start}) - P(M = m_i | SA > sa_{end})\lambda(SA > sa_{end})}{\lambda(SA > sa_{start}) - \lambda(SA > sa_{end})}$$
(3.16)

where  $P(M = m_i | SA > sa_{start})$  and  $P(M = m_i | SA > sa_{end})$  are the magnitude disaggregation results that correspond to exceeding the values of *SA* at the start and end of the bin, respectively; and  $\lambda(SA > sa_{start})$  and  $\lambda(SA > sa_{end})$  are the AEF of the values of *SA* at the start and end of the bin, respectively. For this case study, implementing Equation (3.16) results in the CPT (conditional PMF) for  $M|SA_1^r$ , which is shown in Table 3-5. The conditional probability increases for  $m_i = 7.75$  in BIN-1, BIN-2, and BIN-3 due to the New Madrid seismic source and it is characteristic for the Central Illinois site.

1 4010	5-5								
ЛЛ		$PGA$ (or $SA_1^r$ ) bins (g)							
<i>M</i> hing	m <sub>i</sub>	BIN-1	BIN-2	BIN-3	BIN-4	BIN-5	BIN-6	BIN-7	
DIIIS		0.08-0.18	0.18-0.4	0.4–0.56	0.56-0.86	0.86-1.5	1.5-3.32	> 3.32	
5-5.5	5.25	0.1809	0.2773	0.3543	0.3631	0.3463	0.3238	0.3672	
5.5-6	5.75	0.1532	0.2194	0.2738	0.2891	0.2910	0.2699	0.2187	
6-6.5	6.25	0.1038	0.1431	0.1722	0.1841	0.1974	0.2074	0.1871	
6.5–7	6.75	0.0604	0.0747	0.0878	0.0944	0.1052	0.1231	0.1297	
7–7.5	7.25	0.0859	0.0526	0.0389	0.0391	0.0446	0.0569	0.0685	
7.5-8	7.75	0.3180	0.1697	0.0510	0.0225	0.0139	0.0184	0.0279	

Table 3-5 CPT for Node  $M|SA_1^r$ 

Table .	5-5	CPT IOF N	ode $M$   $SA_1$	l				
M bins				PGA (	or $SA_1^r$ ) bins	(g)		
	$m_i$	BIN-1	BIN-2	BIN-3	BIN-4	BIN-5	BIN-6	BIN-7
		0.08-0.18	0.18-0.4	0.4-0.56	0.56-0.86	0.86-1.5	1.5-3.32	> 3.32
8-8.5	8.25	0.0978	0.0632	0.0221	0.0077	0.0017	0.0005	0.0009

Table 3-5 CPT for Node  $M|SA_1^r$ 

# $3.4.2.3 \qquad \Delta SA_{1,2}|M \text{ and } P'_{1,2}$

The CPT (conditional PMF) of  $\Delta SA_{1,2}$  is generated by first defining a discrete number of bins representing contiguous intervals within the domain of  $\Delta SA_{1,2}$ . The selected discretization may vary by application, desired tolerance for discretization errors, and available computing capabilities. Then, a probability mass is assigned to each bin by discretizing a normal distribution with zero mean and standard deviation that depends on the state (i.e., value) of parent node *M*, in accordance with the equation for  $\sigma_{\Delta SA_{ijk}}(M,\xi,f)$  shown in Table 3-1. Implementing this process results in the CPT for  $\Delta SA_{1,2}|M$ , which is shown in Table 3-6.

1 4010 5 0							
$\Delta SA_{1,2}$	M bins						
bins	5-5.5	5.5–6	6-6.5	6.5–7	7–7.5	7.5–8	8-8.5
-0.460.35	0.0013	0	0	0	0	0	0
-0.350.23	0.0212	0	0	0	0	0	0
-0.230.12	0.1357	0.0101	0.0101	0.0101	0.0101	0.0101	0.0101
-0.12 - 0	0.3418	0.4899	0.4899	0.4899	0.4899	0.4899	0.4899
0 - 0.12	0.3418	0.4899	0.4899	0.4899	0.4899	0.4899	0.4899
0.12 - 0.23	0.1357	0.0101	0.0101	0.0101	0.0101	0.0101	0.0101
0.23 - 0.35	0.0212	0	0	0	0	0	0
0.35 - 0.46	0.0013	0	0	0	0	0	0

Table 3-6 CPT for Node  $\Delta SA_{1,2}|M$ 

We note that the conditional distribution of  $\Delta SA_{1,2}$  for the *M* bins from 5.5–6 to 8–8.5 are the same because, for rock sites, there is no data for these ranges of the earthquake magnitude and Abrahamson and Sykora [24] provided the values for  $c_1(f, M)$  only up to 25 Hz. Due to this, we applied a single coefficient  $c_1(f, M)$  to all magnitude ranges in excess of 5.5. That is, based on a review of the functional shape of the models developed by Abrahamson and Sykora [24], it was

judged to be reasonable to estimate the value of  $c_1(f, M)$  applied above magnitude 5.5 by linear extrapolation in two dimensions (magnitude and frequency) and application of a "lower cap" such that the coefficient was not less than 0.05.

In the case of  $P'_{1,2}$ , we followed the same process as we have just described, except that  $P'_{1,2}$  does not have a dependency on *M* and we use the equation for  $\sigma_{P'}(\xi)$  shown in Table 3-1 to define the standard deviation. We decided to not show the CPT for  $P'_{1,2}$  because it has a similar structure to that of  $SA_1^r$ .

# 3.4.2.4 $SA_2^r | SA_1^r, \Delta SA_{1,2} \text{ and } SA_2^r | SA_1^r, P'_{1,2}$

The CPT (conditional PMF) of node  $SA_2^r$  is generated by first defining a set of contiguous bins that discretize the domain of  $SA_2^r$  and then determining the probability associated with each discrete bin of  $SA_2^r$  given the states of its parents. We defined the  $SA_2^r$  bins to be the same as those of  $SA_1^r$ . In the simplest form, this conditional probability can be achieved by considering the representative values for each state of  $\Delta SA_{1,2}$  (or  $P'_{1,2}$ ) and  $SA_1^r$  and then using calculations that identify into which bin of  $SA_2^r$  the resulting value falls. However, this introduces (potentially significant) discretization error. Instead, we propose to use Monte Carlo simulation to determine the conditional distribution of  $SA_2^r$  for each combination of the states of its parent nodes. Specifically, we uniformly sampled from the range associated with each parent nodes' state to account for the uncertainty within the bin and discretized the resulting synthetic distribution. The "snippet" of the resulting CPT is presented in Table 3-7 for the BN using the A&S [24] formulation. Given that  $SA_1^r$ ,  $\Delta SA_{1,2}$ , and  $SA_2^r$  has eight bins each, the CPT for  $SA_2^r |SA_1^r, \Delta SA_{1,2}$  given the purposes of this example). The CPT for  $SA_2^r | SA_1^r, P_{1,2}'$  has a similar structure; therefore, it is not presented.

We now draw attention to the reason we opted to include BIN-0 in Table 3-4. Under a combination of the low PGA values for the reference unit  $(SA_1^r)$  and small values of  $\Delta SA_{1,2}$  (or  $P'_{1,2}$ ), it is possible that the ground motion at the non-reference unit  $(SA_2^r)$  falls in the range of an earthquake of non-engineering significance. Therefore, we include a PGA bin in the range of  $0g < PGA \le 0.08g$  (i.e., BIN-0 in Table 3-4) to allow for the assignment of probability to these cases. At the other extreme, we note that high  $SA_1^r$  values combined with high  $\Delta SA_{1,2}$  (or  $P'_{1,2}$ ) values can result in extremely high  $SA_2^r$  values. These are assigned to the last PGA bin (BIN-7 in our example). In a seismic PRA, a bin such as BIN-7 is usually assumed to result in core damage directly.

$SA_2^r$	$\Delta SA_{1,2}$ bins							
bins	-0.460.35	-0.350.23	-0.230.12	-0.12 - 0	0-0.12	0.12 - 0.23	0.23 - 0.35	0.35 - 0.46
(g)	$SA_1^r = 0.4g - 0.56g$							
0-0.08	0	0	0	0	0	0	0	0
0.08-0.18	0	0	0	0	0	0	0	0
0.18-0.4	1	0.8436	0.4774	0.1432	0	0	0	0
0.4–0.56	0	0.1564	0.5226	0.8568	0.8041	0.4311	0.1194	0
0.56-0.86	0	0	0	0	0.1959	0.5689	0.8806	0.9808
0.86-1.5	0	0	0	0	0	0	0	0.0192
1.5-3.32	0	0	0	0	0	0	0	0
> 3.32	0	0	0	0	0	0	0	0

Table 3-7 Partial CPT for  $SA_2^r | SA_1^r, \Delta SA_{1,2}$ 

#### 3.4.3 Ground Motion Probability Distributions

We perform probabilistic calculations using the BNs shown in Figure 3-4(b) and (c) and the GeNIe software [22] with CPTs populated as described in Section 3.4.2. As mentioned in

Section 3.4.2, we conditioned the BNs on the occurrence of an EES. The ground motion probability distributions in this section are shown as discrete because the seismic hazard curve is discretized when quantifying a seismic PRA.

Figure 3-6(a) shows the probability distributions of the ground motion at the reference and nonreference units given the occurrence of an EES. It can be observed from the A&S [24] and K&M [25] formulations in Figure 3-6(a) that the probability mass of the ground motion at the nonreference unit is more disperse than the probability mass at the reference unit, reflecting the additional variability. This includes the possibility of the non-reference unit experiencing an earthquake of non-engineering significance when the reference unit is experiencing such an event. Nonetheless, most of the probability mass of the ground motion at the non-reference unit is at BIN-1, as expected, because the distributions of  $\Delta SA_{1,2}$  and  $P'_{1,2}$  have a mean of zero. As the ground motion increases, we observe that the probability mass of the ground motion at the nonreference unit is slightly higher than the probability mass of the ground motion at the reference unit.

To determine the conditional probability distribution of ground motion at the non-reference unit given the ground motion at the reference unit (as would be required to support the performance of the MUSPRA), we use the updating aspects of the BNs by sequentially setting the *SA* for the reference unit to each of the possible ground motion bins as an "evidence case" in the BNs and perform the inference for each case. Figure 3-6(b) shows an example of the conditional probability distributions of *SA* at the non-reference unit using both formulations [24], [25] for the case when the *SA* at the reference unit is in BIN-3. Similar to the distributions in Figure 3-6(a), most of the probability mass of the ground motion at the non-reference unit is in the same bin as

50

the ground motion at the reference unit. However, because of the spatial variability of ground motion models we have included, the probability distribution of ground motion at the nonreference unit is dispersed around the ground motion at the reference unit.



Figure 3-6 (a) Probability distributions of the ground motion at the reference and nonreference units given an earthquake of engineering significance (EES). (b) Probability distribution of the ground motion at the reference unit and conditional probability distribution at the non-reference unit given that  $SA_1^r$  is in the range 0.4g - 0.56g (BIN-3) and an EES.

The calculations provided above were done using the BN as the computational platform. However, the calculations of the example distributions shown in this case study can also be performed using other computational methods. To demonstrate this, we also performed the calculations shown in Figure 3-6(b) using Monte Carlo (MC) simulation and, as expected, obtained similar results as those from using the BN framework. For example, Figure 3-7 compares the calculated conditional distributions using the BN and MC (with one million simulations) for the distribution obtained from the A&S [24] and K&M [25] formulations. Both the MC and BN estimates begin with a hazard discretization that is defined by the PRA-defined ground motion bins. The MC simulation then takes random draws from the specified
distributions, while the BN framework uses the discretized form of the distributions. However, to minimize the effects of discretization in the BN, the CPTs in the BN framework are developed by sampling uniformly within the bin (rather than using a representative value) or, when possible, account for the shape of the underlying distribution as described in Bensi et al. [18]. Under the MC simulation approach, simulations must be performed separately to generate results for each evidence case, while the simulations to generate the BN CPTs need only be performed initially and information updating is performed using the BN inference algorithms. While MC simulation and the BN framework yield similar results, the BN framework offers the following advantages: (1) BNs provide a graphical representation that facilitates communication of modeling assumptions, (2) they are effective at modeling dependencies, and (3) they easily facilitate information updating, which allows for the efficient calculation of conditional distributions for all variables shown in the BN. While MC simulation is not subject to the discretization limitations associated with the BN, this advantage is reduced, in part, by the discretization of the ground motion used as input to the MUSPRA as well as the discrete nature of other PSHA output (e.g., magnitude disaggregation). Moreover, the effects of discretization in the BN can be minimized through integration of MC techniques in the CPT development as in Bensi et al. [18].

The BN framework can be readily extended to model the spatial variability of ground motion at more than two locations by including additional nodes in the BN and making inferences for the additional locations based on the reference ground motion hazard estimated at the NPP site. For example, the BN model formulation linking the ground motion at multiple locations to a single reference location ( $SA_{ref}$ ) and assuming conditional independence of the variability terms ( $\Delta SA_n$  and  $P'_n$ ) from location to location would take the form shown in Figure 3-8(a) and (b) (such a formulation can be likewise implemented via MC simulation). More sophisticated formulations that account for the dependency in the variability terms from location to location would require the development of alternate formulations that require an extension of the existing spatial variability of ground motion models, which compare two locations at a time. BNs have been developed for modeling spatial ground motion correlation (variability) at multiple locations across regional scales as described in Bensi et al. [51].



Figure 3-7 Probability distribution of the ground motion at the reference unit and comparison of the conditional probability distributions at the non-reference unit given that  $SA_1^r$  is in the range 0.4g – 0.56g (BIN-3) and an EES using the A&S [24] and K&M [25] formulations estimated using MC simulation and the BN framework



Figure 3-8 Conceptual extensions of the BN frameworks based on (a) the A&S [24] and (b) K&M [25] formulations to multiple locations on a site

# 3.5 <u>Summary and Conclusions</u>

We proposed a method for modeling ground motion spatial variability across an NPP hard-rock site to support MUSPRA. We present this method as a more realistic alternative to the conventional assumption that ground motions across an NPP site are perfectly correlated. In particular, we propose a pragmatically-motivated approach to developing the conditional probability distribution(s) of ground motion at a non-reference unit(s) at an NPP hard-rock site given the ground motion at a reference unit for which an existing PSHA has been performed. The proposed method is presented using a BN framework because of their graphical structure and capacity to facilitate probabilistic modeling of dependencies; however, the proposed approach is not limited to implementation via the BN framework.

An example application using a PSHA for a hypothetical rock site illustrated the development of the requisite quantities and the means by which they can be obtained using existing resources and a site-specific PSHA. The two model formulations of spatial variability of ground motion used in the case study [24], [25] should be viewed as a representation of the type of models that

are available in the literature (i.e., model functional forms and input variables), rather than an assessment of the validity of these specific models for the target locations. Nonetheless, if a site-specific analysis is carried out using different models of spatial variability of ground motion and these models remain a function of the separation distance alone or a function of the separation distance and the earthquake magnitude, then the formulations presented in this paper remain applicable under differing numerical assumptions (e.g., different distribution parameter values).

Based on the models of spatial variability of ground motion we used in the case study [24], [25], we observed that the probability distribution of the ground motion at the non-reference unit is more dispersed and slightly shifted to the right relative to the distribution of the ground motion at the reference unit, allowing for the possibility of higher levels of ground motion at the nonreference location. Consequently, this can change the seismic failure probability of the associated SSCs due to the possibility of different ground motions. In particular, depending on the shape of the fragility curve, the probability assigned to higher ground motion values can result in an increase in the estimated failure probability of SSCs. Further, inclusion of variability can have impacts on estimated system risk metrics, as described in Ferson et al. [52] and our previous work [46]. Because of this, we note that the perfect correlation assumption with respect to the ground motion at multiple locations/units at an NPP site may not necessarily be conservative. This observation should remain valid regardless of the model of spatial variability of ground motion that is used because any model should estimate that there is a non-negligible probability that the ground motion hazard at the non-reference unit may be different than that of the reference unit.

55

Future studies will address the development of the implementation details and case studies for the MUSPRA input for a soil site. Similar to the implementation for a rock site, the focus for a soil site will be on ensuring that the implementation is based on practical considerations and that the model formulations address the range of existing information that is available about the seismic hazard for the site, soil properties, and site response.

# 3.6 Disclaimer and Acknowledgement

This paper was prepared, in part, by employees of the NRC on their own time apart from their regular duties. The NRC has neither approved nor disapproved its technical content. The views expressed in this paper are not necessarily those of the NRC.

We would like to acknowledge and express appreciation for the comments and suggestions of the reviewers, which helped refocus and improve this paper.

# Chapter 4 A Bayesian Network Approach for Modeling Dependent Seismic Failures in a Nuclear Power Plant Probabilistic Risk Assessment<sup>8</sup>

# 4.1 Introduction

## 4.1.1 Overview

The safety and reliability assurances of nuclear power plant (NPP) structures and components are critical during the engineering design and operation. Safety analyses of NPPs are typically complemented using probabilistic risk assessments (PRA), and the nuclear industry heavily relies on PRAs to address such assurances. Seismic PRA is an integral part of PRA because the seismic hazard is one of the most critical threats to the safety and reliability of NPP structures and components. Seismic PRAs are performed to support risk-informed decisionmaking related to earthquake safety and reliability at NPPs. Seismic PRAs are comprised of three key elements: (1) seismic hazard analysis, (2) seismic fragility evaluation, and (3) plant response model. In this paper, we focus on the seismic fragility evaluation, which estimates the conditional probability of failure of structures, systems, and components (SSCs) whose failure may lead to unacceptable effects on the NPP (e.g., core damage) [12]. Since the ground motion produced by earthquakes

<sup>&</sup>lt;sup>8</sup> This chapter was published in the journal *Reliability Engineering & System Safety*, Vol. 213, p. 107678, Sep. 2021, <u>https://doi.org/10.1016/j.ress.2021.107678</u>.

affects all SSCs at the NPP, it raises the dependency between the seismic failures of SSCs in earthquake scenarios.

Dependency between seismic failures can arise due to correlation in the seismic response and correlation in the seismic capacity<sup>9</sup> of the SSCs [53]. Seismic response refers to the load (demand) experienced by an SSC at its location due to the ground motion that results from an earthquake (e.g., the floor response spectra for a component, which considers the free-field ground motion (i.e., the seismic hazard curve) and the soil-structure interaction). The seismic response correlation between SSCs is due to a single earthquake ground motion affecting all the buildings in the NPP as well as the foundation medium (i.e., rock or soil) that may be similar for some or all the buildings across the site and the layout and design of different buildings that may be similar [15]. Dependency between seismic capacities generally arises because the SSCs have common materials and fabrication and have their fragilities calculated for the same failure modes and using the same models [15].

In this work, we propose a method to model "fragility correlation" using Bayesian networks (BNs), while leveraging the existing conceptual approaches used in seismic PRAs performed in the nuclear industry. The proposed approach intends to leverage modern tools for modeling dependencies and quantifying system failure probabilities (i.e., BNs), which are not subject to the quantitative limitations of existing quantification tools. We seek to leverage these advances while maintaining the *conceptual* knowledge associated with existing practices in the modeling

<sup>&</sup>lt;sup>9</sup> Bohn and Lambright [53] used the term "fragility" to refer to the capacity of an SSC (see Section 2.1.3.4 of Budnitz et al. [15]).

of seismic fragility for NPP SSCs. In this paper, we outline the proposed approach to constructing the BN model and provide practical implementation guidance for developing the requisite components. We use an established case study to compare the results of the proposed BN-based approach against existing quantification methods used in practice as well as theoretical bounds. We show that the proposed BN approach offers several modeling advantages, is more generalizable, and performs more consistently relative to theoretical bounds than the existing standard of practice.

#### 4.1.2 Context and Motivation

The issue of dependent seismic failures and its quantification has been discussed since the 1980s, starting with the Seismic Safety Margins Research Program (SSMRP) [54], [55]. Based on the SSMRP, Bohn and Lambright [53] developed a set of thumb rules for the response correlation and capacity correlation (see Section 3.7.2.2 of NUREG/CR-4840 [53]). As noted by Budnitz et al. [15], in practice, seismic PRAs have applied a correlation coefficient of 1 "to the situation of similar SSCs exposed to the same earthquake load (typically, SSCs located near each other), and zero everywhere else." Both of these cases are extremes and the truth is likely to be somewhere in the middle; however, the approach of using 0 or 1 as the correlation coefficient may be due to the lack of a practical method for analyzing dependent seismic failures as well as the limitations of modeling tools. In Section 1.4 of NUREG/CR-7237 [15], Budnitz et al. state that the fundamental question addressed by the method for modeling dependent seismic failures is:

"what is the joint probability of two or more seismic caused failures, conditional on the occurrence of an earthquake of a given 'size,' and how and why the joint probability may be different from the situation in which those failures are essentially independent." If the failures are essentially independent, the joint probability of failure is calculated as the product of the failure probabilities of the individual SSCs. After considering several methods for analyzing dependent seismic failures of SSCs, Budnitz et al. [15] found the method proposed by Reed et al. [56] (also known as the "Reed-McCann method") to be the most promising method to answer their above stated fundamental question. The Reed-McCann method is further described in Section 4.2.3.

It should be noted that Zhou et al. [57] found issues with the Reed-McCann method's quantification of the seismic failure probability of a series system (i.e., OR failure logic). Due to this issue, Zhou et al. [40] proposed another method to account for the dependency of seismic failures using copulas. They applied it in the context of a multi-unit seismic PRA. Specifically, Zhou et al. [40] used the Gaussian copula to remain consistent with the conventions of the current state of practice of seismic fragility modeling; that is, the assumption that seismic capacities are lognormally distributed. The Gaussian copula requires the estimation of the correlation coefficient and, in their application to a multi-unit seismic PRA, Zhou et al. [40] used the "equi-correlated" model [58], [59] as a reasonable characterization of the correlation. The "equi-correlated" model means that only one correlation coefficient needs to be specified between similar or identical SSCs [40]. They assumed that the "correlation would be provided by the seismic fragility analysts through separating the common sources of uncertainties among the interested SSCs" (see the notes in Figures 3 and 6 of Zhou et al. [40]) and performed sensitivity analyses for the correlation coefficient. Another method that relies on the estimation of correlation coefficients is the work by Watanabe et al. [60], where the effect of correlation/dependency between components (i.e., correlation in seismic responses and correlation in seismic capacities) in the system failure probability is taken into account using

60

direct quantification of fault tree using Monte Carlo simulation (DQFM) method. Kwag et al. [61] proposed to enhance the efficiency of the DQFM method by strategically reusing samples and changing the sampling technique from Monte Carlo simulation to Latin Hypercube sampling. The proposed enhancements by Kwag et al. [61] feature a way of transforming the output from Electric Power Research Institute (EPRI) seismic fragility evaluation [62], which we briefly explain in Sections 4.2.1 and 4.2.2, to the output from the Japan Atomic Energy Research Institute seismic fragility evaluation [63] for use as input to the systems analysis.

In addition to the issues with the Reed-McCann method's quantification identified by Zhou et al. [57], we also found issues with the Reed-McCann method's quantification of the seismic failure probability of a parallel system (i.e., AND failure logic) (see Section 4.4 for more details). Specifically, we found that the Reed-McCann method's quantification can provide system failure probability results outside theoretical bounds. Based on the quantification issues identified for the Reed-McCann method [56] and the need to estimate a correlation coefficient in order to use the methods proposed by Zhou et al. [40], Watanabe et al. [60], and Kwag et al. [61], we propose an alternate approach for analyzing dependent seismic failures based on BNs, which provides a general structure (independent of probability distribution form) for constructing the joint probability distribution of involved random variables using conditional relationships.

As noted previously, a seismic PRA is comprised of three key elements: (1) seismic hazard analysis, (2) seismic fragility evaluation, and (3) plant response model. In our previous work [64], we focused on item (1), proposing a method that uses BNs to account for the spatial variability of ground motions in an NPP multi-unit seismic PRA at a hard-rock site. Our proposed method is a more realistic alternative to the conventional assumption that ground motions across an NPP site are perfectly correlated. In this work, we focus on the modeling of dependencies under the second element of the seismic PRA. The proposal herein is to use BNs to calculate the fragility of a system. The use of BNs is offered as an alternative to the Reed-McCann method, which has been proposed for use in multi-unit seismic PRA as evidenced by its discussion in Annex II of the International Atomic Energy Agency (IAEA) Safety Report Series No. 96 [65] and in EPRI 3002018229 [13].

We propose the BN approach as an alternative to calculate the system fragility of multiple components because of the ease of the proposed integration with other BN-based PRA realism enhancements (e.g., modeling of the spatial variability of ground motion), the overall quantitative robustness relative to theoretical bounds (as demonstrated in Section 4.4), and the inherent characteristics of BNs. Specifically, (1) BNs provide a graphical representation that enables transparency and facilitates communication of modeling assumptions, (2) they are effective at modeling complex dependencies, and (3) they facilitate information updating, which allows for the efficient calculation of joint and conditional probability distributions for all random variables in the BN.

#### 4.1.3 Outline

This paper is organized as follows. Section 4.2 provides background related to the fragility model used in the nuclear industry (Section 4.2.1); evaluation of fragility parameters (Section 4.2.2); the Reed-McCann method, which is expected to be used in the nuclear industry (Section 4.2.3); and estimation of the lower and upper bounds of a system failure probability (Section 4.2.4). Section 4.3 discusses the development of the Bayesian network for modeling dependent seismic failures. Section 4.4 compares the fragility results of a three-component system assuming

62

parallel and series configurations using the Reed-McCann method, the proposed BN approach, the First-Order Reliability Method, and Monte Carlo simulation. The selected three-component system is the illustrative example presented in Sections 8.2.4 and 8.2.5 of NUREG/CR-7237 [15], which analyzed the correlation of seismic performance in similar SSCs. The example system is described as "different components on different floors" [15]. Finally, Section 4.5 summarizes our work and provides conclusions.

# 4.2 <u>Background</u>

4.2.1 Overview of the Seismic Fragility Modeling Used in the Nuclear Industry Seismic fragility (hereafter referred to as "fragility") is the conditional failure probability of an SSC given a value of a ground motion parameter [12]. It is the probability that the seismic capacity ( $\Theta$ ) of the SSC is less than the ground motion parameter (or demand).<sup>10</sup> The ground motion parameter in the fragility must be consistent with the ground motion parameter used in the probabilistic seismic hazard analysis (PSHA) performed at the site of interest [12]. The seismic capacity for the failure mode of interest of an SSC is typically modeled as the product of a "best estimate" seismic capacity (usually the median,  $\Theta_m$ ) and two random variables  $E_R$  and  $E_U$ [66]:

$$\Theta = \Theta_{\rm m} \cdot {\rm E}_R \cdot {\rm E}_U. \tag{4.1}$$

<sup>&</sup>lt;sup>10</sup> In the literature related to NPP seismic fragilities (e.g., [66]), the letter "A" is typically used to represent the seismic capacity of an SSC. However, we are using the Greek letter "Θ" for the same purpose in an attempt to avoid confusion with component A, which we use in the description of the proposed BN approach in Section 3.

The random variable  $E_R$  represents the inherent randomness (aleatory variability) about the median and  $E_U$  represents the epistemic uncertainty in the median value. These two random variables are assumed to be lognormally distributed with medians equal to 1 and logarithmic standard deviations  $\beta_R$  and  $\beta_U$ , respectively. Consequently, the seismic capacity  $\Theta$  is a lognormally distributed random variable with median (scale parameter)  $\Theta_m$  and logarithmic standard deviation (shape parameter) ( $\beta_R^2 + \beta_U^2$ )<sup>0.5</sup>. If a point estimate of the fragility is desired for a given set of parameters  $\Theta_m$ ,  $\beta_R$ , and  $\beta_U$ , it is obtained by using the composite variability,  $\beta_C = (\beta_R^2 + \beta_U^2)^{0.5}$ , as shown in Equation (4.2). By varying the value of  $\theta$  in Equation (4.2), the composite (or mean [67]) fragility curve is obtained.

$$P(\Theta < \theta) = P(SSC \ failure | \theta) = \Phi\left[\frac{\ln\left(\frac{\theta}{\Theta_m}\right)}{\beta_c}\right]$$
(4.2)

In Equation (4.2),  $\Phi[\cdot]$  is the standard normal cumulative distribution function of the term in the brackets;  $\theta$  is the ground motion parameter from the PSHA, which is typically related to the acceleration experienced by the SSC due to the earthquake;  $\Theta_m$  is the median seismic capacity of the SSC; and  $\beta_c$  is the composite variability.

# 4.2.2 Estimation of the Fragility Parameters (Fragility Evaluation)

The objective of the fragility evaluation is to estimate realistic parameters related to the probability distribution of seismic capacity for the failure mode of interest of an SSC (i.e., the parameters  $\Theta_m$ ,  $\beta_U$ , and  $\beta_R$  to use in Equation (4.2)). There are two methods for developing fragilities used in the nuclear industry [68]: the Conservative Deterministic Failure Margin (CDFM) approach [69] and the separation of variables approach (also known as the fragility

analysis method) [62]. The CDFM approach first calculates the high-confidence-of-lowprobability-of-failure (HCLPF) seismic capacity, that is, the acceleration value at which the SSC has a probability of failure of less than 1 percent [70]. Then, it calculates the median seismic capacity using the HCLPF seismic capacity and an assumed composite variability ( $\beta_c$ ) [68]. The equation to calculate the median seismic capacity is shown in Equation (4.3) [70].

$$\Theta_m = \frac{HCLPF}{\exp(-2.33\beta_C)} \tag{4.3}$$

The separation of variables approach calculates the seismic capacity ( $\Theta$ ) as the product of a reference seismic capacity ( $\Theta_{REF}$ ) (i.e., a reference value of a ground motion parameter) and a factor of safety (*F*) [62]:

$$\Theta = F \cdot \Theta_{REF} \tag{4.4}$$

The value of  $\Theta_{REF}$  is often taken as the peak ground acceleration (PGA) from either the sitespecific uniform hazard response spectrum<sup>11</sup> calculated at  $10^{-4}/yr$  or the ground motion response spectrum.<sup>12</sup> The factor of safety is assumed to be lognormally distributed with median  $F_m$  and logarithmic standard deviations  $\beta_R$  and  $\beta_U$ , which are identical to the aleatory and epistemic uncertainties of the seismic capacity [66]. The factor of safety is estimated as the product of a series of factors. The factors used differ based on whether the fragility evaluation is

<sup>&</sup>lt;sup>11</sup> The work by Baker [71] explains how to calculate the uniform hazard response spectrum.

<sup>&</sup>lt;sup>12</sup> The U.S. Nuclear Regulatory Commission's Regulatory Guide 1.208 [72] explains how the ground motion response spectrum is calculated.

being performed on a structure or equipment. For a structure, the factor of safety is estimated as the product of a capacity factor ( $F_{Cap}$ ) and structure response factor ( $F_{RS}$ ):  $F = F_{Cap} \cdot F_{RS}$ . The capacity factor ( $F_{Cap}$ ) is calculated as a product of strength and ductility factors. The structure response factor ( $F_{RS}$ ) is calculated as the product of series of factors accounting for structural response variability.

For equipment qualified by analysis, the factor of safety *F* includes an additional factor ( $F_{RE}$ ) to account for differences in equipment response relative to the structural response; that is,  $F = F_{Cap} \cdot F_{RS} \cdot F_{RE}$ . The equipment response factor ( $F_{RE}$ ) is, in turn, estimated as the product of factor associated with dynamic response [62]. Futher details about the factors discussed above and the method to obtain the seismic capacity of equipment qualified by testing are discussed in EPRI 3002012994 [62].

Based on the assumptions that the factor of safety (*F*) is lognormally distributed and that all the constituent factors are independent, the median factor of safety  $F_m$  and the logarithmic standard deviation  $\beta_F$  can, in general, be estimated as

$$F_m = F_{Cap_m} \cdot F_{RS_m} \cdot F_{RE_m} \tag{4.5}$$

and

$$\beta_F = \left(\beta_{F_{Cap}}^2 + \beta_{F_{RS}}^2 + \beta_{F_{RE}}^2\right)^{0.5},\tag{4.6}$$

respectively. In Equation (4.5),  $F_{Cap_m}$ ,  $F_{RS_m}$ , and  $F_{RE_m}$  are the median values of the capacity, structure response, and equipment response factors, respectively, which are obtained by the product of the medians of their constituent factors. In Equation (4.6),  $\beta_{F_{Cap}}$ ,  $\beta_{F_{RS}}$ , and  $\beta_{F_{RE}}$  are

the variabilities of the capacity, structure response, and equipment response factors, respectively, which are in turn obtained from the variabilities of the constituent factors. The individual  $\beta$ -values in Equation (4.6) are further divided into randomness ( $\beta_R$ ) and uncertainty ( $\beta_U$ ) [62] from which the overall  $\beta_R$  and  $\beta_U$  values are derived and then the composite variability is calculated as  $\beta_C = (\beta_R^2 + \beta_U^2)^{0.5}$ . The median seismic capacity used in the fragility calculation (see Equation (4.2)) is calculated as shown in Equation (4.7) [62].

$$\Theta_m = F_m \cdot \Theta_{REF} \tag{4.7}$$

#### 4.2.3 The Reed-McCann Method

As noted in Section 4.1, the current state of practice of modeling dependent seismic failures is to use a correlation coefficient equal to one for similar SSCs located next to each other and zero otherwise. The method by Reed et al. [56] (i.e., the Reed-McCann method), which was adopted by Budnitz et al. [15] as the most promising method to model dependent seismic failures, starts by examining the process in which the factors of safety in the fragility evaluation (see Section 4.2.2) are developed. In turn, the examination allows the quantification of the dependency in the structural parameters. Dependency relationships are established by systematically identifying the common  $\beta_U$  and  $\beta_R$  values among the SSCs of interest [56].

The Reed-McCann method quantifies the system failure probability considering dependence in the capacities of the constituent components and consists of two stages, which are summarized in Figure 4-1. Figure 4-1 includes detailed quantification steps based on the Mathcad files used by Budnitz et al. [15] to implement the Reed-McCann method [56] using Latin Hypercube sampling. Both stages of the Reed-McCann method begin with input from the fragility analyst, who is expected to provide the median seismic capacities and associated logarithmic standard deviations for all system components,  $i = 1, ..., N_c$ , where  $N_c$  is the number of components, which will be typically obtained using the procedures summarized in Section 4.2.2. The Reed-McCann method is based on the premise that, for any pair of system components, both the aleatory variability and epistemic uncertainty can be partitioned into a "common" element and a component-specific ("reduced") element. In Figure 4-1,  $\beta_{U_{ij}}^*$  and  $\beta_{R_{ij}}^*$  are parameters that represent the common epistemic and aleatory uncertainties, respectively, that exist between the component under consideration (i.e., component *i*) and other components ( $i \neq j$ ). The  $\beta'_{U_i}$  and  $\beta'_{R_i}$  parameters represent the reduced epistemic and aleatory uncertainties, respectively, of the seismic capacity of component *i* and are calculated using the equations shown in Steps 1 and 4 of Figure 4-1, respectively. In general, several  $\beta^*_{U_{ij}}$  and  $\beta^*_{R_{ij}}$  values are required to represent the different dependency groups among components within a system.



Figure 4-1 Summary of the Reed-McCann method.

# 4.2.4 Lower and Upper Bounds of a System Failure Probability

When modeling dependent seismic failures in seismic PRA, the main interest is when there is positive dependence because it affects the redundancy incorporated in NPP systems. The lower and upper bounds of the failure probability for parallel (i.e., AND failure logic) and series (i.e., OR failure logic) systems with positive dependence can be expressed as follows (generalized from the work by Ferson et al. [52]):

Parallel System

$$P(A) \cdot P(B) \cdot P(C) \le P(A \cap B \cap C) \le \min(P(A), P(B), P(C))$$

$$(4.8)$$

• Series System  $\max(P(A), P(B), P(C)) \le P(A \cup B \cup C) \le 1 - (1 - P(A))(1 - P(B))(1 - P(C))$ (4.9)

where  $P(A \cap B \cap C)$  is the parallel system failure probability,  $P(A \cup B \cup C)$  is the series system failure probability, and P(A), P(B), and P(C) refer to the failure probability of components A, B, and C, respectively. In the context of seismic fragilities, the failure probabilities of the components (see Equation (4.2)) and systems are a function of the ground motion. The lower and upper bounds of the system failure probability whose components are positively dependent correspond to the dependency assumptions summarized in Table 4-1.

Table 4-1         Summary of Dependency Assumptions									
System Configuration	Lower Bound	Upper Bound							
Parallel	Independent	Perfect dependence							
Series	Perfect dependence	Independent							

As shown in Table 4-1, a dependency assumption can give either the lower or upper bound of the system failure probability depending on the system configuration. For example, the

independence assumption gives the lower bound of the failure probability for a parallel system, but it gives the upper bound of the failure probability for a series system.

#### 4.3 Bayesian Network Approach for Modeling Dependent Seismic Failures

The proposed BN approach for modeling dependent seismic failures relies on separating common sources of uncertainties among the SSCs of interest, consistent with the conceptual approaches used in existing practice. We propose the use of BNs to model dependent seismic failures because of the transparency offered by their graphical nature, as well as the ease and completeness with which they facilitate modeling probabilistic dependencies and quantification of system failure probabilities. Also, BNs facilitate information updating, which allows for the efficient calculation of joint and conditional probability distributions for all random variables in the BN. Section 3 of our previous work [64] provides a brief introduction to BNs.

Figure 4-2 shows the BN used to analyze the system-level seismic fragility of three components (A, B, and C). The node *EES* represents an earthquake of engineering significance (EES) event (i.e., an event in which the ground motion exceeds a specific minimum value of relevance to the nuclear facility) and all inferences using the BN are conditioned on such an event. The node *GM* represents the event that a ground motion (GM) is in a range of values (referred to in practice as a "ground motion bin"), where each range of GM values is a state. The ground motion parameter of interest in our work is PGA because it is typically used in the quantification of seismic PRAs; however, the proposed approach is equally applicable to other response spectrum ground motion parameters. For the remainder of this paper, the terms "PGA value(s)" and "GM value(s)," and "*PGA*" and "*GM*" are synonymous.

The nodes *A*, *B*, and *C* represent the state (i.e., failure or survival) of components A, B, and C, respectively. The state of components A, B, and C depends on the ground motion experienced at the component's location and the capacity of the component. In this work, we assume all components experience the same ground motion, as reflected by the common parent *GM* node to nodes *A*, *B*, and *C*. The BN proposed in our previous work relaxes this assumption, that is, it accounts for the spatial variability of ground motion experienced by multiple components at multiple locations [64].

The capacity of a component is defined as in Equation (4.1). Here, we use the log-transformation and then separate both the aleatory variability ( $E_R$ ) and epistemic uncertainty ( $E_U$ ) into the common parts (i.e.,  $E_{R_{ij}}^*$  and  $E_{U_{ij}}^*$ ) and component-specific parts (i.e.,  $E_{R_i}'$  and  $E_{U_i}'$ ). We define the capacity of component *i*, *i* = *A*, *B*, *C*, as:

$$\ln(\Theta_{i}) = \ln(\Theta_{m_{i}}) + \ln(E_{R_{i}}) + \ln(E_{U_{i}})$$
  
=  $\ln(\Theta_{m_{i}}) + \ln(E_{R_{i}}') + \ln(E_{U_{i}}') + \sum_{i \neq j} \left[ \ln(E_{R_{ij}}^{*}) + \ln(E_{U_{ij}}^{*}) \right]$  (4.10)

where all the terms are defined in Section 4.2.1 and above. Then, we set

$$\ln(E'_{i}) = \ln(E'_{R_{i}}) + \ln(E'_{U_{i}})$$
(4.11)

$$\ln(\mathbf{E}_{ij}^*) = \ln\left(\mathbf{E}_{R_{ij}}^*\right) + \ln\left(\mathbf{E}_{U_{ij}}^*\right)$$
(4.12)

where  $E'_i$  is the component-specific uncertainty for component *i* and  $E^*_{ij}$  is the common uncertainty between components *i* and *j* ( $i \neq j$ ). Then, we define the following:

$$\varepsilon_i' = \ln(\mathsf{E}_i') \tag{4.13}$$

$$\varepsilon_{ij}^* = \ln(\mathbf{E}_{ij}^*) \tag{4.14}$$

Next, we substitute Equations (4.11) and (4.12) into Equation (4.10) and use the definitions in Equations (4.13) and (4.14) to obtain:

$$\ln(\Theta_i) = \ln(\Theta_{m_i}) + \ln(E'_i) + \sum_{i \neq j} [\ln(E^*_{ij})]$$
  
= 
$$\ln(\Theta_{m_i}) + \varepsilon'_i + \sum_{i \neq j} [\varepsilon^*_{ij}]$$
(4.15)

In Figure 4-2, the component-specific capacity uncertainty for components A, B, and C is reflected in nodes  $\varepsilon'_A$ ,  $\varepsilon'_B$ , and  $\varepsilon'_C$ , respectively. The common sources of uncertainty among the components are captured by nodes  $\varepsilon^*_{AB}$ ,  $\varepsilon^*_{AC}$ , and  $\varepsilon^*_{BC}$ , which introduces the dependency between the components. The median seismic capacity of each component ( $\Theta_{m_i}$ ) is not explicitly shown as a node in the BN due to its deterministic nature. Finally, the node *System* represents the system state (i.e., "Failure" or "Success") given the states of components A, B, and C. The BN in Figure 4-2 does not explicitly distinguish between the aleatory and epistemic parts of uncertainty. Instead, the  $\varepsilon$ -nodes include both the aleatory and epistemic parts of uncertainty because the ultimate quantification is theoretically invariant to whether they are combined or treated separately; however, inclusion of additional nodes increases the computational demand and may increase the effect of discretization error.

The BN structure in Figure 4-2 is based on the relatively low number of components and our desire to explicitly model all random variables that affect the system fragility. However, a BN

for the same system may have a different structure. In general, the nodes in a BN that represent the dependency between components should be explicitly modeled (e.g., nodes  $\varepsilon_{AB}^*$ ,  $\varepsilon_{AC}^*$ , and  $\varepsilon_{BC}^*$ in Figure 4-2) while nodes that only affect one component may be integrated into the node that represents the component (e.g., node  $\varepsilon_A'$  may be integrated into node *A* in Figure 4-2) [18], [73]. In addition, if a system has a relatively large number of components, Bensi et al. [73] proposed to model the *System* node using a chain-like BN structure instead of a converging BN structure (e.g., nodes *A*, *B*, and *C* converge into the node *System* in Figure 4-2) to make the inferences in the BN more efficient.

We illustrate the application of the BN in Figure 4-2 using an established case study example system in separate parallel and series configurations. The example system is described as "different components on different floors" and has been previously used to demonstrate the modeling of dependent fragilities [15]. In the next subsections, we will explain how the conditional probability table (CPT) for each node in Figure 4-2 is generated. The parameters used to characterize the capacity of the components are provided in Table 4-2. They correspond to the parameters used in the examples in Sections 8.2.4 and 8.2.5 of NUREG/CR-7237 [15] and uses the notation introduced in Section 4.2.<sup>13</sup>

<sup>&</sup>lt;sup>13</sup> These parameters are assumed to be provided by the fragility analyst after the fragility evaluation (Section 4.2.2) and the systematic identification of the common sources of uncertainty among the SSCs of interest. According to Budnitz et al. [15], "the fragility analyst is well equipped to make this judgement [i.e., the identification of common sources of uncertainty among the SSCs of interest] because he/she has in intimate knowledge (acquired through the review of design documents and from the plant walkdown) of how the components are designed, qualified and installed."



Figure 4-2 Bayesian network for modeling the fragility of a three-component system.

Table 4-2 Fragility Parameters and Group Dependencies									
Component Properties									
Component ( <i>i</i> )	$\Theta_{m_i}$	$\beta_{U_i}$	$\beta_{R_i}$						
A	0.7g	0.35	0.20						
В	1.3g	0.30	0.25						
С	0.9g	0.40	0.20						
	<b>Group Dependencies</b>								
		Common Variability							
Group	Components ( <i>ij</i> )	$eta_{U_{ij}}^*$	$eta_{R_{ij}}^*$						
1	AB	0.10	0.10						
2	AC	0.15	0.05						
3	BC	0.20	0.10						

#### 4.3.1 Generation of the CPT for the EES Node

The CPT for the *EES* node is based on the ground motion value that has the potential to meaningfully affect the site's SSCs. In general, this ground motion value is the starting ground motion value in the first bin of the discretized seismic hazard curve used in the quantification of the seismic PRA. In our example, an EES occurs when the PGA exceeds 0.1g. Based on the PSHA results from the Central Illinois site we used in our previous work [64], we obtained a function for the annual exceedance frequency (AEF) of the seismic hazard [74]; it is shown in Equation (4.16) and is valid from 0.02g through 1.7g.

$$AEF = \lambda(PGA > pga) = 2.67 \times 10^{-6} \cdot PGA^{-2.05}$$
(4.16)

Using Equation (4.16) with PGA = 0.1g and based on the Poisson assumption,<sup>14</sup> the (annual) probability of an EES is  $3 \times 10^{-4}$ . The CPT for the *EES* node is shown as Table 4-3, in which "EES no" and "EES yes" represent the states of the *EES* node. The (annual) probability of the "EES no" state is  $1 - P(EES \ yes) = 1 - 3 \times 10^{-4} = 0.9997$ .

Table 4-3CPT for the EES NodeStateP(State)ESS no0.9997ESS yes0.0003

#### 4.3.2 Generation of the CPT for the *GM*|*EES* Node

The method for generating the CPT for the *GM* node is explained in our previous work [64], which we summarize as follows. The domain of *GM* is divided into several bins consistent with the discretization of the seismic hazard curve used in the quantification of the seismic PRA. Typically, six to eight ground motion bins are used in the seismic PRA (e.g., [50]). In this example, the start and end PGA values for each of the *GM* bins were selected so that their geometric mean, which is the typical representative PGA value used in the quantification of a seismic PRA, is similar to the PGA values used by Budnitz et al. in their implementation of the Reed-McCann method [15] (see the  $GM_{rep}$  values in the first column of Table 4-10 and Table 4-11). To ensure probabilistic consistency (i.e., ensure that the CPT columns for the *GM* node sum up to 1.0), we conditioned the *GM* node on the occurrence of an EES. The expressions of

<sup>&</sup>lt;sup>14</sup> Under the Poisson assumption, the AEF of a PGA value is numerically equivalent to the probability of exceeding the same PGA value (i.e.,  $\lambda(PGA > pga) \approx P(PGA > pga)$ ).

the conditional probability for bins  $GM_1$  through  $GM_7$  and  $GM_8$  given an EES are shown as Equations (4.17) and (4.18), respectively. The  $\lambda(PGA > pga_{start})$  and  $\lambda(PGA > pga_{end})$ values in Equations (4.17) and (4.18) are calculated using Equation (4.16). The denominator in Equations (4.17) and (4.18) is based on the ground motion value to be exceeded that would result in an EES, which in our example is 0.1g.

$$P(pga_{start} < PGA \le pga_{end} | PGA > 0.1g) = \frac{\lambda(PGA > pga_{start}) - \lambda(PGA > pga_{end})}{\lambda(PGA > 0.1g)}$$
(4.17)

$$P(PGA > pga_{start} | PGA > 0.1g) = \frac{\lambda(PGA > pga_{start})}{\lambda(PGA > 0.1g)}$$
(4.18)

Applying Equations (4.17) and (4.18) yields the conditional probabilities for bins (states)  $GM_1$  through  $GM_8$ , which are shown in the last column of Table 4-4 (i.e., "P(State|EES yes)"). The probability of an EES in state  $GM_0$  is zero because the PGA values in state  $GM_0$  are less than or equal to the PGA value we selected as the EES. The probabilities in the "P(State|EES no)" column of Table 4-4 are determined based on the fact that when the event is not an EES, the PGA values must be less than or equal to 0.1g. Therefore, the probability of  $GM_0$  given that the node EES is in its "EES no" state is 1, which leads to the conclusion that the probabilities for the remaining GM bins (i.e., states  $GM_1$  through  $GM_8$ ) are zero.

		the up		
GM Bins	GM Bins	EES	EES no	EES yes
$GM_{start} - GM_{end} \left(g\right)$	$GM_{rep}\left(g ight)$	State	P(State EES no)	P(State EES yes)
0 - 0.10	-	$GM_0$	1	0
0.10 - 0.15	0.122	$GM_1$	0	5.65 x 10 <sup>-1</sup>
0.15 - 0.20	0.173	$GM_2$	0	1.94 x 10 <sup>-1</sup>
0.20 - 0.30	0.245	$GM_3$	0	1.36 x 10 <sup>-1</sup>
0.30 - 0.48	0.379	$GM_4$	0	6.50 x 10 <sup>-2</sup>
0.48 - 0.71	0.584	$GM_5$	0	2.21 x 10 <sup>-2</sup>
0.71 - 0.95	0.821	$GM_6$	0	8.08 x 10 <sup>-3</sup>

Table 4-4Bins and CPT for the *GM* | *EES* Node

			the up			
GM Bins		GM Bins	EES	EES no	EES yes	
	$GM_{start} - GM_{end} \left(g\right)$	GM <sub>rep</sub> (g)	State	P(State EES no)	P(State EES yes)	
	0.95 - 1.50	1.194	$GM_7$	0	6.02 x 10 <sup>-3</sup>	
	1.50 - 1.94	1.706	$GM_8$	0	3.88 x 10 <sup>-3</sup>	

Table 4-4Bins and CPT for the *GM* | *EES* Node

4.3.3 Generation of the CPTs for the  $\varepsilon_{AB}^*$ ,  $\varepsilon_{AC}^*$ , and  $\varepsilon_{BC}^*$  Nodes

As noted in Section 4.2.1, the typical seismic PRA practice assumes that  $E_R$  and  $E_U$  are lognormally distributed with medians equal to 1 and logarithmic standard deviations  $\beta_R$  and  $\beta_U$ , respectively. Using the partitioning introduced in Equations (4.11) and (4.12), the common uncertainty between component *i* and other components ( $i \neq j$ ) is comprised of both aleatory and epistemic components:  $\varepsilon_{ij}^* = \ln (E_{R_{ij}}^*) + \ln (E_{U_{ij}}^*)$ . It follows that  $\varepsilon_{ij}^*$  is normally distributed<sup>15</sup> with zero mean and standard deviation  $\beta_{C_{ij}}^*$ , which represents the common composite variability of components *i* and *j* and is calculated using Equation (4.19).

$$\beta_{C_{ij}}^{*} = \left[ \left( \beta_{U_{ij}}^{*} \right)^{2} + \left( \beta_{R_{ij}}^{*} \right)^{2} \right]^{0.5}$$
(4.19)

<sup>&</sup>lt;sup>15</sup> Since  $E_{R_{ij}}^*$  and  $E_{U_{ij}}^*$  are lognormally distributed with medians equal to 1 and logarithmic standard deviations,  $\beta_{R_{ij}}^*$  and  $\beta_{U_{ij}}^*$ , respectively, their respective logarithms are normally distributed with means equal to zero (=  $\ln(median) = \ln(1)$ ) and standard deviations  $\beta_{R_{ij}}^*$  and  $\beta_{U_{ij}}^*$ , respectively. Then, the sum of two independent, normally distributed random variables is also normally distributed with mean equal to the sum of the means of the two random variables and standard deviation equal to the square root of the sum of the variances of the two random variables (see Equation (4.19)).

In Equation (4.19),  $\beta_{U_{ij}}^*$  and  $\beta_{R_{ij}}^*$  are parameters that represent the common epistemic and aleatory uncertainties, respectively, that exist between the component under consideration (i.e., component *i*) and other components ( $i \neq j$ ). The parameters  $\beta_{U_{ij}}^*$  and  $\beta_{R_{ij}}^*$  would be provided by the fragility analysts as the result of the separation of common sources of uncertainties among the SSCs of interest; the values used in this paper are shown in Table 4-2.

The CPTs of the "common uncertainty" nodes  $\varepsilon_{AB}^*$ ,  $\varepsilon_{AC}^*$ , and  $\varepsilon_{BC}^*$  are generated by first discretizing their respective continuous domains. The selected discretization may vary by application, desired tolerance for discretization errors, and available computing capabilities [64]. The probability mass of each bin is developed by discretizing a normal distribution with zero mean and standard deviation  $\beta_{C_{ij}}^*$ .

To generate the CPT for node  $\varepsilon_{AB}^*$ , we assign probability to each state of  $\varepsilon_{AB}^*$  by "discretizing" the distribution of  $\varepsilon_{AB}^*$ . That is, the domain of  $\varepsilon_{AB}^*$  is first linearly discretized into *n* contiguous bins centered between approximately  $\pm 3$  standard deviations. Each discrete bin corresponds to one state in the CPT for node  $\varepsilon_{AB}^*$ . It is generally preferred for *n* to be equal to an odd number to ensure one of the bins is centered at zero, which is the mean of the normal distribution being discretized in this case. The upper edge of bin *n* and lower edge of bin 1 are replaced by  $\infty$  and  $-\infty$ , respectively, to ensure all probability is included in the CPT for  $\varepsilon_{AB}^*$ . The probability assigned to bin (state) *i*, *i* = 1, ..., *n*, with lower and upper bin edges equal  $\varepsilon_{AB_{start(i)}}^*$  and  $\varepsilon_{AB_{end(i)}}^*$ , is the probability that  $\varepsilon_{AB}^*$  falls within the bin edges. This is computed as:

$$P\left(\varepsilon_{AB_{start(i)}}^{*} \le \varepsilon_{AB}^{*} \le \varepsilon_{AB_{end(i)}}^{*}\right) = \Phi\left(\frac{\varepsilon_{AB_{end(i)}}^{*}}{\beta_{C_{AB}}^{*}}\right) - \Phi\left(\frac{\varepsilon_{AB_{start(i)}}^{*}}{\beta_{C_{AB}}^{*}}\right)$$
(4.20)

where  $\Phi(\cdot)$  is the standard normal cumulative distribution function of the term inside the parenthesis. While the above approach was used in this case study, alternate strategies for discretization may readily be employed for the development of CPTs.

Using the values in Table 4-2 and Equation (4.19), the values of the parameters  $\beta_{C_{AB}}^*$ ,  $\beta_{C_{AC}}^*$ , and  $\beta^*_{C_{BC}}$  in our example are 0.14, 0.16, and 0.22, respectively. The CPT for  $\varepsilon^*_{AB}$  is shown as Table 4-5. We note that the CPTs for  $\varepsilon_{AC}^*$  and  $\varepsilon_{BC}^*$  have a similar structure as the CPT of  $\varepsilon_{AB}^*$  and are thus not repeated here.

Table 4-5 Bins and	d CPT for t	he $\varepsilon_{AB}^*$ Node
$arepsilon_{AB}^*  ext{ Bins } \ arepsilon_{AB_{start}}^* - arepsilon_{AB_{end}}^*$	State	P(State)
-0.42 to -0.34	$\varepsilon_{AB_1}^*$	7.05E-3
-0.34 to -0.27	$\varepsilon^*_{AB_2}$	2.11E-2
-0.27 to -0.19	$\mathcal{E}_{AB_{3}}^{*}$	5.82E-2
-0.19 to -0.11	$\varepsilon^*_{AB_4}$	1.20E-1
-0.11 to -0.04	$\mathcal{E}_{AB_{5}}^{*}$	1.86E-1
-0.04 to 0.04	$\mathcal{E}_{AB_{6}}^{*}$	2.15E-1
0.04 to 0.11	$\mathcal{E}_{AB_{7}}^{*}$	1.86E-1
0.11 to 0.19	$\varepsilon^*_{AB_8}$	1.20E-1
0.19 to 0.27	$\mathcal{E}_{AB9}^{*}$	5.82E-2
0.27 to 0.34	$\varepsilon_{AB}^{*}{}_{10}$	2.11E-2
0.34 to 0.42	$\varepsilon_{AB_{11}}^*$	7.05E-3

Generation of the CPTs for the  $\varepsilon'_A$ ,  $\varepsilon'_B$ , and  $\varepsilon'_C$  Nodes 4.3.4

The process for generating the CPTs of the "component-specific uncertainty" nodes  $\varepsilon'_A$ ,  $\varepsilon'_B$ , and  $\varepsilon_{\mathcal{C}}'$  is similar to the process described in Section 4.3.3. The only difference is that the standard deviation of the zero mean normal distribution used to generate the probability mass of each bin is  $\beta'_{C_i}$ . The parameter  $\beta'_{C_i}$  is the component-specific (reduced) composite variability of component *i*, which is calculated using Equation (4.21).

$$\beta_{C_i}' = \left[ \left( \beta_{U_i}' \right)^2 + \left( \beta_{R_i}' \right)^2 \right]^{0.5}$$
(4.21)

The parameters  $\beta'_{U_i}$  and  $\beta'_{R_i}$  are the component-specific (reduced) epistemic uncertainty and aleatory variability of the seismic capacity of component *i*, and are calculated using the equations shown in Steps 1 and 4 of Figure 4-1, respectively. Using the values in Table 4-2, the equations in Steps 1 and 4 of Figure 4-1, and Equation (4.21), the values of the parameters  $\beta'_{C_A}$ ,  $\beta'_{C_B}$ , and  $\beta'_{C_C}$  in our example are 0.34, 0.29, and 0.35, respectively. The CPT for  $\varepsilon'_A$  is shown as Table 4-6. We note that the CPTs for  $\varepsilon'_B$  and  $\varepsilon'_C$  have a similar structure as the CPT of  $\varepsilon'_A$  and are thus not shown.

Table 4-0	DIIIS a		$\Gamma$ the $\mathcal{E}_A$ Node
$arepsilon_A'  ext{ Bins} \ arepsilon_{Astart}' - arepsilon_A'$	Aend	State	P(State)
-1.02 to -0	0.83	$\varepsilon'_{A_1}$	7.05E-3
-0.83 to -0	.65	$\varepsilon'_{A_2}$	2.11E-2
-0.65 to -0	.46	$\varepsilon'_{A_3}$	5.82E-2
-0.46 to -0	0.28	$\varepsilon'_{A_4}$	1.20E-1
-0.28 to -0	0.09	$\varepsilon'_{A_5}$	1.86E-1
-0.09 to 0	.09	$\varepsilon'_{A_6}$	2.15E-1
0.09 to 0.	28	$\varepsilon'_{A_7}$	1.86E-1
0.28 to 0.	46	$\varepsilon'_{A_8}$	1.20E-1
0.46 to 0.	65	$\varepsilon'_{A_9}$	5.82E-2
0.65 to 0.	83	$\varepsilon'_{A_{10}}$	2.11E-2
0.83 to 1.	02	$\mathcal{E}'_{A_{11}}$	7.05E-3

Table 4-6 Bins and CPT for the  $\varepsilon'_4$  Node

An astute reader may notice that the probability mass of each bin is the same for  $\varepsilon_{AB}^*$  and  $\varepsilon_A'$ . This is because the respective normal distributions are discretized using the same number of bins; however, the values for each of the variables in the bins are different.

4.3.5 Generation of the CPTs for the 
$$(A|GM, \varepsilon_{AB}^*, \varepsilon_{AC}^*, \varepsilon_A')$$
,  $B|GM, \varepsilon_{AB}^*, \varepsilon_{BC}^*, \varepsilon_B'$ , and  $C|GM, \varepsilon_{AC}^*, \varepsilon_{BC}^*, \varepsilon_C'$  Nodes

The limit state function  $g_i = \ln(\Theta_i) - \ln(GM)$ , i = A, B, C, is used to compute the failure probability associated with each component. For example, using Equation (4.15) for component A the limit state function takes the following form:

$$g_A = \ln(\Theta_{m_A}) + \varepsilon'_A + \varepsilon^*_{AB} + \varepsilon^*_{AC} - \ln(GM)$$
(4.22)

The limit-state function is in its failure state when  $g_A \leq 0$ . Therefore, the failure probability of component A is:

$$P(g_A \le 0) = P\left(\ln(\Theta_{m_A}) + \varepsilon'_A + \varepsilon^*_{AB} + \varepsilon^*_{AC} - \ln(GM) \le 0\right)$$
  
$$= P\left(\ln(\Theta_{m_A}) + \varepsilon'_A + \varepsilon^*_{AB} + \varepsilon^*_{AC} \le \ln(GM)\right)$$
  
$$= P\left(\exp\left[\ln(\Theta_{m_A}) + \varepsilon'_A + \varepsilon^*_{AB} + \varepsilon^*_{AC}\right] \le GM\right)$$
(4.23)

In Equation (4.23),  $\Theta_{m_A}$  is the median seismic capacity of component A and *GM* is the seismic demand (ground motion) due to an earthquake, and all other terms are as defined above. The values of the median seismic capacity for components A, B, and C are given in Table 4-2. To develop the CPT for the node *A* using the above expressions, we apply Monte Carlo simulation. Monte Carlo simulation is used to reduce the effect of the discretization of the parent nodes on the computed failure probabilities assigned to node *A*.

To illustrate this process, let  $n_{\varepsilon'_A}$ ,  $n_{\varepsilon^*_{AB}}$ ,  $n_{\varepsilon^*_{AC}}$ , and  $n_{GM}$  denote the number of bins (states) associated with nodes  $\varepsilon'_A \varepsilon^*_{AB}$ ,  $\varepsilon^*_{AC}$ , and GM, respectively. Recall that each state of  $\varepsilon'_A \varepsilon^*_{AB}$ ,  $\varepsilon^*_{AC}$ , and GM is associated with a discretized range of the associated random variable. For each combination of the states of  $\varepsilon'_A \varepsilon^*_{AB}$ ,  $\varepsilon^*_{AC}$ , and GM, we draw  $N_{Sim}$  (independent) simulations from the distributions of  $\varepsilon'_A \varepsilon^*_{AB}$ ,  $\varepsilon^*_{AC}$ , and GM truncated and renormalized to the discretized range associated with the relevant states. For each simulation, we then compute the value of the limit state function  $g_A = \ln(\Theta_A) - \ln(GM)$ . We estimate the failure probability for each combination of parent node states as the fraction of  $N_{sim}$  simulations for which  $g_A \leq 0$ .

For example, consider one combination of the states of the parents to node (component) A:

$$\left\{i_{\varepsilon_{A}^{\prime}},i_{\varepsilon_{AB}^{*}},i_{\varepsilon_{AC}^{*}},i_{GM}\right\},i_{\varepsilon_{A}^{\prime}}=1,\ldots,n_{\varepsilon_{A}^{\prime}};\ i_{\varepsilon_{AB}^{*}}=1,\ldots,n_{\varepsilon_{AB}^{*}};\ i_{\varepsilon_{AC}^{*}}=1,\ldots,n_{\varepsilon_{AC}^{*}};\ i_{GM}=1,\ldots,n_{GM}$$

We generate  $N_{sim}$  (independent) simulations of each quantity using the following distributions:

• the truncated normal distribution: 
$$f_{\varepsilon_A'}\left(\varepsilon_A' \middle| \varepsilon_{A'} \atop start(i_{\varepsilon_A'}) \le \varepsilon_A' < \varepsilon_{A'} \atop end(i_{\varepsilon_A'})\right)$$

• the truncated normal distribution: 
$$f_{\varepsilon_{AB}^*}\left(\varepsilon_{AB}^*\right|\varepsilon_{AB}^*\left[\varepsilon_{AB}^*\right] \le \varepsilon_{AB}^* < \varepsilon_{AB}^* < \varepsilon_{AB}^*\right)$$

• the truncated normal distribution: 
$$f_{\varepsilon_{AC}^*}\left(\varepsilon_{AC}^*\left|\varepsilon_{AC}^*\right|\varepsilon_{AC}^*\right) \le \varepsilon_{AC}^* < \varepsilon_{AC}^*\right) \le \varepsilon_{AC}^* < \varepsilon_{AC}^*$$

• the truncated uniform distribution<sup>16</sup>:  $f_{GM}(gm|GM_{start(i_{GM})} \leq GM < GM_{end(i_{GM})})$ where the notation  $f_X(x|x_1 \leq X < x_2)$  denotes the probability density function of the random variable *X*, truncated and renormalized to the interval  $[x_1, x_2)$ . The mean of the normal distributions being truncated is zero and the respective standard deviations are  $\beta'_{C_A}$ ,  $\beta^*_{C_{AB}}$ , and  $\beta^*_{C_{AC}}$ . For each combination of simulations for the parent nodes, we estimate the probability of failure as the fraction of  $N_{sim}$  simulations for which  $\exp[\ln(\Theta_{m_A}) + \varepsilon'_A + \varepsilon^*_{AB} + \varepsilon^*_{AC}] \leq GM$ . Then, the success probability for component A is just the complement of its probability of failure. The process is repeated for each combination of the states of the parents of node *A* to generate the CPT for node (component) A. Similar calculations are performed for nodes (components) B and C.

Based on the number of states of the parent nodes (i.e., 9 *GM* states and 11 states for  $\varepsilon_{AB}^*$ ,  $\varepsilon_{AC}^*$ , and  $\varepsilon_A'$  each) and node *A* (i.e., 2 states), the CPT for the seismic performance of component A has 23,958 entries (= 9 · 11 · 11 · 11 · 2), which makes the presentation of the CPT of node *A* impractical. However, a "snippet" of the CPT of node *A* is shown as Table 4-7. In Table 4-7, the "F" and "S" represent the failure state and success state of component A, respectively, and, for example, " $\varepsilon_{A6}' - \varepsilon_{A11}'$ " means that the state probabilities of component A are the same for states 6

<sup>&</sup>lt;sup>16</sup> As will be described in Section 4.4, we also implemented an alternative approach in which the GM is represented by a single representative value for each bin. This is necessary to enable an "apples-to-apples" comparison between the Reed-McCann method and the BN approach.

through 11 of the node  $\varepsilon'_A$  given  $\varepsilon^*_{AC_1}$ ,  $\varepsilon^*_{AB_1}$ , and  $GM_1$ . This is not general and it will vary depending on the states of the random variables.

_10	Table 4-7 Simplet of the CT 1 for Component A given $GM$ , $\mathcal{E}_{AB}$ , $\mathcal{E}_{AC}$ , and $\mathcal{E}_{A}$														
	GM <sub>1</sub>														
	$arepsilon^*_{AB_1}$														
	$\varepsilon_{AC_1}^* \qquad \varepsilon_{AC_2}^* \qquad \varepsilon_{AC_3}^*$														
	$\varepsilon'_{A_1}$	$\varepsilon'_{A_2}$	$\varepsilon_{A_3}'$	$\varepsilon'_{A4}$	$\varepsilon'_{A_5}$	$\varepsilon_{A_6}' - \varepsilon_{A_{11}}'$	$\varepsilon_{A_1}'$	$\varepsilon'_{A_2}$	$\varepsilon'_{A_3}$	$\varepsilon'_{A_4}$	$arepsilon_{A_5}^{\prime} - arepsilon_{A_{11}}^{\prime}$	$\varepsilon'_{A_1}$	$\varepsilon'_{A_2}$	$\varepsilon'_{A_3}$	$arepsilon_{A_4}^{\prime} - arepsilon_{A_{11}}^{\prime}$
F	0.45	0.25	0.09	0.02	0.0007	0	0.22	0.08	0.013	0.0003	0	0.06	0.009	0.0002	0
S	0.55	0.75	0.91	0.98	0.9993	1	0.78	0.92	0.987	0.9997	1	0.94	0.991	0.9998	1

Table 4-7 "Snippet" of the CPT for Component A given GM,  $\varepsilon_{AB}^*$ ,  $\varepsilon_{AC}^*$ , and  $\varepsilon_{A}'$ 

Lastly, the CPT of component A has a dependency on state  $GM_0$ ; however, the  $GM_0$  state has no effect on the system fragility results because all inferences using the BN are conditioned on the occurrence of an EES, which set the conditional probability of  $GM_0$  given the state "EES yes" to zero (as shown in Table 4-4).

# 4.3.6 Generation of CPT for the *System* | *A*, *B*, *C* Node

The CPT for the *System* node is based on Boolean failure logic. For the parallel (AND failure logic) and series (OR failure logic) systems, the CPTs are shown as Table 4-8 and Table 4-9, respectively. An example of how these CPTs should be interpreted is as follows. In the parallel system, the *System* node is in its "Failure" state with probability of 1 only if components A, B, and C are each in the "Failure" state; otherwise, the *System* node is in its "Success" state with probability of 1.

Α		Fai	lure		Success						
В	Fai	lure	Suc	cess	Fai	Success					
С	Failure	Success	Failure	Success	Failure	Success	Failure	Success			
Failure	1	0	0	0	0	0	0	0			
Success	0	1	1	1	1	1	1	1			

 Table 4-8
 System | A, B, C Node CPT for the Parallel System

Α		Fai	lure		Success					
В	Fai	lure	Success		Failure		Success			
С	Failure	Success	Failure	Success	Failure	Success	Failure	Success		
Failure	1	1	1	1	1	1	1	0		
Success	0	0	0	0	0	0	0	1		

Table 4-9System | A, B, C Node CPT for the Series System

#### 4.4 <u>Results Comparison</u>

In this section, we compare the results of the system fragility analysis for two system configurations, parallel and series, using the Reed-McCann method, the proposed BN approach, the First-Order Reliability Method (FORM), and Monte Carlo simulation (MCS). The reader is referred to the work by Der Kiureghian [75] for an explanation of FORM.

We implemented the Reed-McCann method [56] using the Mathcad [76] files developed by Budnitz et al. in NUREG/CR-7237 [15]. We performed inferences using the BN in Figure 4-2 and the GeNIe software [22]. We implemented FORM using the Matrix-based System Reliability analysis method (MSR) within the Finite Element Reliability Using Matlab (FERUM) software [77]. MCS was implemented using Matlab [78] with  $1 \times 10^8$  samples. Since the logarithms of the seismic capacities are joint-normally distributed, direct MCS under varying assumptions of correlation (i.e., independence, perfect dependence, and using a specified correlation matrix) is relatively straightforward. Moreover, the limited size of the system makes direct MCS computationally tractable. Zhou et al. [40] provided strategies for simulation under more general joint probability distributions using copulas. They further proposed the use of importance sampling to increase the efficiency of simulations, which may be needed for more complex system configurations. The specification of limit-state functions and correlation coefficients between the involved random variables are required to implement FORM and MCS. The limit-state function we used for each component i is shown in Equation (4.24). The correlation coefficient of the logarithm of the seismic capacities of each pair of components is based on Equation (5) in the work by Reed et al. [56] and is shown in Equation (4.25).

$$g_{i} = \exp(\Theta_{i}) - GM$$

$$(4.24)$$

$$\rho_{\ln(\Theta_{i}),\ln(\Theta_{j})} = \frac{\left(\beta_{c_{ij}}^{*}\right)^{2}}{\beta_{c_{i}}\beta_{c_{j}}}$$

$$(4.25)$$

In Equation (4.24),  $\Theta_i$  is the seismic capacity of component *i* and *GM* is the ground motion value. The seismic capacity of component *i*,  $\Theta_i$ , is assumed to be normally distributed with mean  $\ln(\Theta_{m_i})$  and standard deviation  $\beta_{C_i}$ ,<sup>17</sup> where  $\Theta_{m_i}$  is the median seismic capacity of component *i* and  $\beta_{C_i}$  is a parameter that represents the composite variability of component *i* and is equal to  $(\beta_{U_i}^2 + \beta_{R_i}^2)^{0.5}$ . The terms of Equation (4.25) have been previously defined. The values for  $\Theta_{m_i}$ ,  $\beta_{U_i}$ , and  $\beta_{R_i}$  for components A, B, and C and their respective group dependency parameters,  $\beta_{U_{ij}}^*$ and  $\beta_{R_{ij}}^*$ , are provided in Table 4-2.

For the ground motion value, GM, we developed two sets of calculations. The first set of calculations assumes that GM is fixed at the geometric mean of  $GM_{start}$  and  $GM_{end}$  for each GM

<sup>&</sup>lt;sup>17</sup> This is equivalent to the formulation that the seismic capacity of component *i* is lognormally distributed with median  $\Theta_{m_i}$  and logarithmic standard deviation  $\beta_{C_i}$ .

bin. We refer to this as the representative GM value  $(GM_{rep})$  of the bin. The second set of calculations assumes that *GM* is uniformly distributed between  $GM_{start}$  and  $GM_{end}$  for each bin and is used to account for the uncertainty within the bin. The calculations with the fixed *GM* value were performed to compare the system fragility results obtained using the BN, FORM, and MCS to those from the Reed-McCann method [56], which uses a fixed *GM* value, as implemented by Budnitz et al. [15] (i.e., the typical practice used in the nuclear industry). The calculations with the uniformly distributed *GM* values were performed to compare the system fragility results obtained using the system.

In the BN shown in Figure 4-2, all inferences are conditioned on the occurrence of an EES, that is, the "EES yes" state of the *EES* node is set as evidence. Next, we used the updating aspects of the BN by setting each of the states of the *GM* node (i.e.,  $GM_1$  through  $GM_8$ ) as evidence and performed the (forward) inference for each case.

In terms of the computational times, once the problem is set up (i.e., all computer modeling components "coded"), the Reed-McCann method takes several minutes to calculate the system fragility for all GM bins in each system. FORM and MCS take several minutes to calculate the system fragility for each ground motion bin. The generation of the CPTs for the BN (using MCS) also takes several minutes. But once the CPTs are created (which only needs to be performed once), the inference for each ground motion bin to calculate the system fragility takes less than one second. These computational times were obtained using a personal computer (Intel® Core<sup>TM</sup> i7-6500U processor with 16 GB of RAM).

The results from the four approaches/methods along with the theoretical lower and upper bounds of the system fragility for parallel and series configurations are shown in Table 4-10 and Table
4-11, respectively. In Table 4-10 and Table 4-11, the results from the calculations where the GM value was fixed are shown in parenthesis and the results where the GM value is uniformly distributed within the bin are not in parenthesis. The shaded values indicate cases where the results are outside the theoretical lower and upper bounds of the system fragility. The lower and upper bounds of the system fragility for the calculations where the GM value was fixed were obtained by calculating the fragility of each component using Equation (4.2) and then using Equation (4.8) for the parallel system and Equation (4.9) for the series system. For the calculations where the GM is uniformly distributed within the bin, the lower and upper bounds were obtained using MCS to account for the uncertainty within the ground motion bin and calculate the fragility of each component. Then, the system fragilities were calculated under the assumptions of independent seismic capacities and perfect dependence in seismic capacities. These results are also plotted in Figure 4-3 (parallel system) and Figure 4-4 (series system) for the case where the GM value is fixed. We did not plot the case where the GM value is uniformly distributed because, currently, the Reed-McCann method cannot be implemented using a randomly distributed GM value.

		1231 [13]				
GM Bin GM <sub>start</sub> - GM <sub>end</sub> (GM <sub>rep</sub> )	Lower Bound <sup>1</sup> [independence]	Upper Bound <sup>1</sup> [perfect dependence]	Reed-McCann Quantification <sup>2,5</sup>	Bayesian Network <sup>3,4,5</sup>	FORM <sup>3</sup>	MCS <sup>1,3,5</sup>
BIN-1 0.10g – 0.15g (0.122g)	~0 (2.0E-20)	~0 (6.9E-10)	N/A (0)	~0 (~0)	Note 6 (1.0E-15)	~0 (~0)
BIN-2 0.15g – 0.20g (0.172g)	~0 (3.0E-15)	2.7E-7 (1.1E-7)	N/A (0)	2.5E-13 (~0)	6.9E-11 (8.6E-12)	~0 (~0)
BIN-3 0.20g – 0.30g (0.242g)	~0 (5.8E-11)	2.3E-5 (8.4E-6)	N/A (5.1E-14)	1.5E-7 (3.0E-8)	1.9E-7 (1.5E-8)	7.0E-8 (3.0E-8)
BIN-4 0.30g - 0.48g (0.377g)	1.2E-5 (1.2E-6)	1.6E-3 (7.6E-4)	N/A (1.5E-7)	1.6E-4 (7.7E-5)	1.6E-4 (2.1E-5)	9.5E-5 (2.4E-5)
BIN-5 0.48g - 0.71g (0.582g)	2.5E-3 (1.1E-3)	2.6E-2 (2.0E-2)	N/A (3.9E-4)	8.5E-3 (6.8E-3)	8.4E-3 (3.8E-3)	6.5E-3 (3.8E-3)
BIN-6 0.71g - 0.95g (0.819g)	4.0E-2 (3.2E-2)	1.3E-1 (1.2E-1)	N/A (2.5E-2)	7.4E-2 (7.0E-2)	6.9E-2 (5.5E-2)	6.3E-2 (5.5E-2)
BIN-7 0.95g – 1.50g (1.19g)	3.1E-1 (2.7E-1)	4.4E-1 (4.1E-1)	N/A (2.7E-1)	3.5E-1 (3.4E-1)	3.6E-1 (3.2E-1)	3.5E-1 (3.2E-1)
BIN-8 1.50g – 1.94g (1.7g)	6.9E-1 (6.9E-1)	7.6E-1 (7.5E-1)	N/A (7.1E-1)	7.1E-1 (7.1E-1)	7.1E-1 (7.1E-1)	7.1E-1 (7.1E-1)

Table 4-10System Fragility of the Parallel Configuration in Section 8.2.4 of<br/>NUREG/CR-7237 [15]

Notes

1. Since we used  $10^8$  samples with MCS, we use "approximately zero" (~0) for the case when the simulation contains no failures.

2. The Reed-McCann method uses the fixed GM value shown as "GM<sub>rep</sub>" in the first column of this table. The Reed-McCann method cannot be implemented using a randomly distributed GM value. Therefore, we use "N/A" for the results of the calculation where the GM values are uniformly distributed within the bin.

3. The fixed GM value used with the BN, FORM, and MCS is the geometric mean of  $GM_{start}$  and  $GM_{end}$ , which varies from  $GM_{rep}$  at the 3<sup>rd</sup> decimal place (compare with the  $GM_{rep}$  values in the second column of Table 4-4).

4. The BN results shown as "approximately zero" (~0) are because the inference in GeNIe showed the guaranteed success of component B and this guarantees the success of the parallel system. This is due to the amount of simulations used in the generation of the CPT for component B.

5. The shaded cells indicate that the calculated system fragility is outside the lower and upper bounds.

6. We were unable to obtain a result because FORM did not converge when estimating the fragility of component B.

		n 1201 [10]				
GM Bin GM <sub>start</sub> - GM <sub>end</sub> (GM <sub>rep</sub> )	Lower Bound [perfect dependence]	Upper Bound [independence]	Reed-McCann Quantification <sup>1,3</sup>	Bayesian Network <sup>2,3</sup>	FORM <sup>2,3</sup>	MCS <sup>2,3</sup>
BIN-1 0.10g – 0.15g (0.122g)	1.7E-5 (7.3E-6)	2.7E-5 (1.1E-5)	N/A (9.7E-8)	2.7E-6 (2.7E-7)	Note 4 (1.2E-5)	2.5E-5 (1.2E-5)
BIN-2 0.15g – 0.20g (0.172g)	3.6E-4 (2.5E-4)	5.1E-4 (3.6E-4)	N/A (1.2E-4)	2.4E-4 (1.4E-4)	6.0E-4 (3.8E-4)	5.2E-4 (3.8E-4)
BIN-3 0.20g – 0.30g (0.242g)	6.7E-3 (4.2E-3)	9.3E-3 (5.9E-3)	N/A (9.1E-3)	7.9E-3 (5.1E-3)	1.1E-2 (6.4E-3)	9.3E-3 (6.4E-3)
BIN-4 0.30g – 0.48g (0.377g)	8.1E-2 (6.2E-2)	1.1E-1 (8.7E-2)	N/A (1.0E-1)	1.1E-1 (8.4E-2)	1.2E-1 (8.9E-2)	1.1E-1 (8.9E-2)
BIN-5 0.48g - 0.71g (0.582g)	3.4E-1 (3.2E-1)	4.7E-1 (4.5E-1)	N/A (5.2E-1)	4.3E-1 (4.1E-1)	4.6E-1 (4.3E-1)	4.5E-1 (4.3E-1)
BIN-6 0.71g - 0.95g (0.819g)	6.6E-1 (6.5E-1)	8.2E-1 (8.2E-1)	N/A (8.7E-1)	7.7E-1 (7.6E-1)	7.9E-1 (7.9E-1)	7.9E-1 (7.9E-1)
BIN-7 0.95g – 1.50g (1.19g)	9.0E-1 (9.1E-1)	9.8E-1 (9.9E-1)	N/A (9.5E-1)	9.6E-1 (9.6E-1)	9.6E-1 (9.8E-1)	9.7E-1 (9.8E-1)
BIN-8 1.50g – 1.94g (1.7g)	9.9E-1 (9.9E-1)	1 (1)	N/A (1)	1 (1)	1 (1)	1 (1)

Table 4-11System Fragility of the Series Configuration in Section 8.2.5 of<br/>NUREG/CR-7237 [15]

Notes

1. The Reed-McCann method uses the fixed GM value shown as "GM<sub>rep</sub>" in the first column of this table. The Reed-McCann method cannot be implemented using a randomly distributed GM value. Therefore, we use "N/A" for the results of the calculation where the GM values are uniformly distributed within the bin.

2. The fixed GM value used with the BN, FORM, and MCS is the geometric mean of  $GM_{start}$  and  $GM_{end}$ , which varies from  $GM_{rep}$  at the 3<sup>rd</sup> decimal place (compare with the  $GM_{rep}$  values in the second column of Table 4-4).

3. The shaded cells indicate that the calculated system fragility is outside the lower and upper bounds.

4. We were unable to obtain a result because FORM did not converge when estimating the fragility of component B.



Figure 4-3 System fragility of the parallel configuration in Section 8.2.4 of NUREG/CR-7237 [15] for the case where the GM value is fixed. The shaded area represents the lower and upper bounds of the system fragility for the parallel configuration.



Figure 4-4 System fragility of the series configuration in Section 8.2.5 of NUREG/CR-7237 [15] for the case where the GM value is fixed. The shaded area represents the lower and upper bounds of the system fragility for the series configuration.

For the parallel system (Table 4-10 and Figure 4-3), we observed that the results from the BN approach remained inside the lower and upper bounds of the system fragility in six out of eight GM values. The BN approach results outside the bounds are associated with the "low end" of the fragility curve (i.e., BIN-1 and BIN-2) in which the MCS used to generate the CPTs and the discretization in the BN impacted the results, and the estimated system fragility bounds are very small. Refinement of the discretization and modification of the MCS approach used in generating CPTs can help to improve the BN performance in estimating these small probabilities. However, from the perspective of the seismic PRA quantification, these small system fragility results at the "low end" of the fragility curve may not be included in the seismic PRA results because of the

truncation limits that are used in the seismic PRA quantification and the truncation limits vary from application to application.<sup>18</sup> In other words, if the value of the cut set, which includes the contribution from the AEF of the seismic hazard, is less than the truncation limit, that cut set is not considered in the seismic PRA quantification. The Reed-McCann method likewise generated system fragility results outside the bounds in the "low end" of the fragility curve. However, in the "middle" of the fragility curve (i.e., BIN-3 through BIN-6), the Reed-McCann method also failed to give results that stay within the lower and upper bounds of the system fragility. In contrast, the other three approaches yielded results within the lower and upper bounds of the system fragility in the "middle" of the fragility curve. The BN, FORM, and MCS approaches provided similar results (i.e., in the same order of magnitude) in the "middle" of the fragility curve. All four approaches/methods provided similar results that stay within the lower and upper bounds of the system fragility at the "high end" of the fragility curve (i.e., BIN-7 and BIN-8). For the case where the GM value is uniformly distributed within each bin, the BN approach results remained within the lower and upper bounds of the system fragility for all GM bins. Also, the BN approach provided similar results as those from FORM and MCS. This provides confidence that the BN approach appropriately captures the dependencies between the seismic capacities of multiple components in a parallel system.

<sup>&</sup>lt;sup>18</sup> Truncation limits may be in the range of  $1 \times 10^{-13}$  to  $1 \times 10^{-10}$ ; however, the truncation limit for a specific application is selected in accordance with the supporting requirements QU-B2 and QU-B3 of the American Society of Mechanical Engineers (ASME)/American Nuclear Society (ANS) PRA standard [79].

For the series system (Table 4-11 and Figure 4-4), we also observed that the results from the BN approach remained inside the lower and upper bounds of the system fragility in six out of eight GM values, with similar trends observed in the "low end" of the fragility curve (i.e., BIN-1 and BIN-2) as in the parallel system case. In the "middle" of the fragility curve (i.e., BIN-3 through BIN-6), the BN approach results fell between the lower and upper bounds of the system fragility. In contrast, the Reed-McCann method results fell outside the lower and upper bounds of the system fragility with all the results being greater than the upper bound of the system fragility, that is, the Reed-McCann method results are overly conservative. The BN, FORM, and MCS provided similar results, but in the case of BIN-3 and BIN-4, FORM and MCS gave a system fragility slightly greater than the upper bound of the system fragility. Similar to the parallel system, all four approaches/methods provided similar results that remained within the lower and upper bounds of the system fragility at the "high end" of the fragility curve (i.e., BIN-7 and BIN-8). For the case where the GM value is uniformly distributed within each bin, MCS has better performance than the BN and FORM at the "low end" of the fragility curve as judged by the performance within the theoretical bounds. In the "middle" and "high end" of the fragility curve, the BN and MCS had similar performance and provided results that fell between the lower and upper bounds of the system fragility. FORM produced results that were slightly greater than the upper bounds of the system fragility for BIN-3 and BIN-4 and for the remaining bins (i.e., BIN-5 through BIN-8), FORM has similar performance to the BN and MCS. For the series system, the BN appropriately captures the dependency between multiple components in BIN-3 through BIN-8. For BIN-1 and BIN-2 (i.e., low ground motion values), it may be appropriate to use the independence assumption for the series system.

The system fragility results for the parallel and series systems tend to be higher for the case where the GM value within each bin is uniformly distributed than those where the representative GM value within each bin is fixed. This is expected because the case where the GM value within each bin is uniformly distributed considers the possibility of GM values that are greater than GM<sub>rep</sub> when estimating the fragilities of each component and that of the system within that bin. This tendency for higher system fragility results is balanced by the fact that the case where the GM value within each bin is uniformly distributed also considers GM values that are less than GM<sub>rep</sub>. Thus, the system fragility estimated when considering the ground motion uncertainty within the GM bin will differ from the value estimated when considering a fixed GM value at the representative GM value of the bin. The magnitude of the difference depends on the shape of the fragility curve over the range of GM values covered by the bin. This observation may support arguments made by Zhou et al. [40] and the IAEA [39] against the use of the geometric mean as the representative GM value of a GM bin. Specifically, Zhou et al. [40] and the IAEA [39] support using GM<sub>end</sub> as the representative GM value of a GM bin. We did not explore using GM<sub>end</sub> as the representative GM value of a GM bin because of the relative ease with which the BN facilitates consideration of "within bin uncertainty." Therefore, we do not offer commentary on the use of alternate representative GM values.

#### 4.5 <u>Summary and Conclusions</u>

We proposed an alternative approach to the Reed-McCann method based on BNs for modeling dependent seismic failures. Based on the results (Section 4.4), the BN approach we proposed in this paper is better suited than the Reed-McCann method (as implemented in NUREG/CR-7237 [15]) for modeling dependent seismic failures as judged by the comparison against the theoretical lower and upper bounds of the system fragility. Also, as demonstrated by the

95

illustrative example, the BN approach does not need an explicit estimate of the correlation coefficient as is the case for the methods proposed by Zhou et al. [40], Watanabe et al. [60], and Kwag et al. [61], but instead only requires information regarding the relative "split" between common and component-specific uncertainties. We also compared the performance of the BN to results based on FORM and MCS. The BN, FORM, and MCS provided similar results. In addition, the BN offers the following advantages over the Reed-McCann method, FORM, and MCS: (1) BNs provide a graphical representation that enables transparency and facilitates communication of modeling assumptions; (2) they are effective at modeling complex dependencies and can accommodate differing probability distribution assumptions; and (3) they facilitate information updating, which allows for the efficient calculation of joint and conditional probability distributions for all random variables in the BN. In this paper, we demonstrated advantage (1) by using Figure 4-2 to provide a graphical representation of the system, advantage (2) by computing the system fragility which includes complex dependencies represented by the  $\varepsilon_{ii}^*$  nodes, and advantage (3) by calculating the parallel and series system fragilities for each ground motion bin using the updating aspects of the BN. As an additional example of advantage (3), the parallel system fragility in ground motion BIN-5 changes from  $6.8 \times 10^{-3}$  to  $2.1 \times 10^{-2}$ when the failure of component A is entered as evidence; obtaining this result did not require recalculating the CPTs. In combination with the extension to PSHA to account for the spatial variability of ground motion we previously proposed [64], the BN approach presented in this paper can be used to simultaneously and realistically account for dependent seismic failures and spatial variability of ground motion in both single-unit and multi-unit seismic PRAs.

#### 4.6 <u>Disclaimer and Acknowledgements</u>

This paper was prepared, in part, by an employee of the U.S. Nuclear Regulatory Commission (NRC) on his own time apart from his regular duties. The NRC has neither approved nor disapproved its technical content. The views expressed in this paper are not necessarily those of the NRC.

We would like to thank the following individuals from the NRC staff: Robert Roche Rivera and Richard Rivera Lugo for facilitating training material related to the development of seismic fragilities in the nuclear industry, and José Pires for sharing the Mathcad files that showed the implementation of the Reed-McCann method. We greatly appreciate the assistance from Constantinos Frantzis, graduate student in the Department of Civil and Environmental Engineering at the University of Maryland, College Park, in setting up CalREL [80], which was an instrumental precursor to our implementation of MSR within FERUM. Finally, we are grateful to the reviewers for their time in reviewing the original manuscript and for their comments and suggestions, which improved this paper. THIS PAGE IS INTENTIONALLY LEFT BLANK

# Chapter 5 Multi-Unit Seismic Probabilistic Risk Assessment: A Bayesian Network Perspective<sup>19</sup>

#### 5.1 Introduction

#### 5.1.1 Motivation and Objective

The Fukushima Dai-ichi accident [9]–[11] highlighted the need to consider multi-unit accidents in probabilistic risk assessments (PRA) of nuclear power plants (NPP). The accident at the Fukushima Dai-ichi NPP initiated as an earthquake followed by a tsunami that resulted in three of the six units at the site experiencing core damage and subsequent release of radioactive material to the environment. In response to the accident, the nuclear industry has developed additional strategies to address risks from natural hazards (e.g., use of diverse and flexible coping strategies equipment [81]). In parallel, there has been an increased interest in multi-unit probabilistic risk assessment (MUPRA) (e.g., [13], [65]). Modarres et al. [82] note that "MUPRA refers broadly to an extension of the traditional PRA techniques to assess the risks of multi-unit sites. This includes single-unit PRAs that consider the accident sequences that may propagate from one unit to another, fully integrated PRA models that address accident sequences that may involve any combination of reactor units and radiological sources, and hybrids of these models."

<sup>&</sup>lt;sup>19</sup> This chapter was submitted to the journal *Reliability Engineering & System Safety* on May 18, 2021 (Manuscript number: JRESS-D-21-00695) and as of May 22, 2021, it is "Under Review."

In the context of this paper, we focus on the application of MUPRA to seismic events because "[i]t has long been recognized that external events, particularly seismic and external flooding events, could be substantial contributors to risk because of the potential for multiple commoncause failures" [10]. This motivates the use of seismic PRAs to characterize the risk contribution of earthquakes to the NPP site, rather than the risk contribution on a unit-by-unit basis.

The elements of a seismic PRA (SPRA) are probabilistic seismic hazard analysis (PSHA), fragility evaluation, and systems (plant response) analysis [12]. The PSHA is used to estimate the frequency (per year) of exceeding a certain value of a ground motion parameter (e.g., peak ground acceleration (PGA) or spectral acceleration), which is represented as a seismic hazard curve. The fragility evaluation is used to estimate the seismic capacity and conditional probability of failure of structures, systems, and components (SSCs). The system (plant response) analysis includes the modeling of the NPP's response to the ground motion caused by an earthquake using event trees and fault trees, which may result in a core damage event and radiation release, and quantification of the appropriate risk metrics (e.g., core damage frequency).

The objective of this paper is to integrate our previous works of extending a PSHA to account for the spatial variability of ground motion [64] and estimating the correlated (dependent) seismic failures (i.e., fragilities) of SSCs [83] into a multi-unit seismic probabilistic risk assessment (MUSPRA) approach to explore the relative importance of different types of correlation. The approach is demonstrated by modeling a pair of nuclear reactors at a hypothetical NPP site.

100

#### 5.1.2 Context

Schroer and Modarres [23] proposed an event classification schema to account for dependencies among multiple units at an NPP site. The event classification schema is composed of seven categories, where six of them have the potential to affect multiple units and one does not affect multiple units (i.e., independent events). The six categories that have the potential to affect multiple units are (1) initiating events, (2) shared connections, (3) identical components, (4) proximity dependencies, (5) human dependencies, and (6) organizational dependencies. Subsequently, there was an international workshop on multi-unit probabilistic safety assessment (MUPSA) [84], where one of the technical challenges identified was site-based accident sequence quantification and risk metrics, including the "need to address variations in site response to the same earthquake and correlation among component fragilities in the MUPSA context."

Several works have been published about MUSPRA, with some pre-dating the Fukushima Daichi accident. The works by Hataka [85] and Ebisawa et al. [86] use the Seismic Safety Margins Research Program (SSMRP) fragility method [53], [54], [87]. Therefore, these works consider the correlation in response and correlation in capacity (fragility), which are later combined into a failure correlation using the following equation [53]:

$$\rho_{1,2} = \frac{\beta_{R1} \cdot \beta_{R2}}{\sqrt{\beta_{R1}^2 + \beta_{F1}^2} \cdot \sqrt{\beta_{R2}^2 + \beta_{F2}^2}} \cdot \rho_{R1,R2} + \frac{\beta_{F1} \cdot \beta_{F2}}{\sqrt{\beta_{R1}^2 + \beta_{F1}^2} \cdot \sqrt{\beta_{R2}^2 + \beta_{F2}^2}} \cdot \rho_{F1,F2}$$
(5.1)

In Equation (5.1),  $\rho_{1,2}$  is the correlation coefficient between the failure of components 1 and 2;  $\beta_{R1}$  and  $\beta_{R2}$  are the standard deviations of the logarithms of the responses of components 1 and 2, respectively;  $\beta_{F1}$  and  $\beta_{F2}$  are the standard deviations of the logarithms of the capacities (fragilities) of components 1 and 2, respectively,  $\rho_{R1,R2}$  is the correlation coefficient between the responses of components 1 and 2; and  $\rho_{F1,F2}$  is the correlation coefficient between the capacities (fragilities) of components 1 and 2. Typically, under the SSMRP fragility method, it is assumed that there is not enough information to determine the capacity (fragility) correlation coefficient between components 1 and 2 ( $\rho_{F1,F2}$ ) or that capacity (fragility) components 1 and 2 are independent; therefore,  $\rho_{F1,F2} = 0$  (e.g., [86]). Alternatively, the effect of  $\rho_{F1,F2}$  in the results of the seismic PRA may be determined through sensitivity analysis. Hakata [85] proposed rules for assigning the response correlation coefficient across units and these rules are assigned by detailed response analysis or expert judgment. Ebisawa et al. [86] implemented the failure correlation coefficient across units using a large fault tree composed of the core damage sequences from each unit.

Other works in MUSPRA, such as those by Zhou et al. [40] and Jung et al. [88], are carried out in the context of the Electric Power Research Institute (EPRI) fragility methodology, which is documented in EPRI 3002012994 [62]. The EPRI fragility methodology expresses the fragility of an SSC as a function of a global ground motion parameter (e.g., PGA), and assumes that the seismic capacity is lognormally distributed with median seismic capacity  $\Theta_m$  and logarithmic standard deviations reflecting the epistemic uncertainty ( $\beta_U$ ) and randomness ( $\beta_R$ ) in the seismic capacity [62], [66]. Zhou et al. [40] used the Gaussian copula with a judged correlation coefficient between the seismic capacities of SSCs to model the correlated (dependent) seismic failures. Then, they calculated the equivalent  $\beta$ -factor common cause failure (CCF) parametric model used in the traditional single-unit PRAs as the ratio of the number of correlated failures to the number of all failures. Subsequently, the  $\beta$ -factor is used to define a seismic CCF model, which was also implemented as part of a large fault tree representing the entire MUSPRA. Jung et al. [88] converted correlated seismic failures into seismic CCF basic events by solving a system of simultaneous equations, where the number of simultaneous equations and corresponding basic events depends on the number of SSCs in the correlated group (i.e., the number of basic events is equal to  $2^{(\# of SSCs in correlated group)} - 1$ ).

This paper proposes a different approach to perform MUSPRA. It relies on Bayesian networks (BNs) to represent the entire MUSPRA. This is because, based on its graphical nature and rigorous probabilistic calculations, BNs offer a systematic, easy-to-follow approach to model MUSPRA where modeling dependencies are critical. Bayesian networks have been used in PRAs to model correlated (basic) events under multiple hazards [89], validate PRA models [90], and estimate multi-unit seismic core damage probability [91]. Previously, we proposed using BNs to model the spatial variability of ground motion in a MUSPRA [64] and dependent seismic failures [83]. The proposed approach for modeling dependent seismic failures using BNs [83] is performed in the context of the EPRI fragility methodology [62]. In contrast with the other works that use the EPRI fragility methodology, the proposed approach to model MUSPRA using BNs explicitly considers the spatial variability of ground motion for the global ground motion parameter when estimating NPP site-based risk metrics. In this paper, we estimate the NPP site-based risk metrics listed below [40].

• Site core damage frequency (CDF): the frequency of one or more core damage events; this definition corresponds to the union (logic OR) of core damage events in multiple units.

• Concurrent CDF: the frequency of multiple core damage events nearly simultaneously; this definition corresponds to the intersection (logic AND) of core damage events in multiple units.

#### 5.1.3 Background

There are many published works that describe the elements of a seismic PRA and provide an information about BNs. In our previous works [64], [83], we provided a background on these topics in the context of a MUSPRA and will not repeat them here. For a more comprehensive background on these topics, an interested reader may consult EPRI 3002000709 [12] for a description of the elements of a seismic PRA and the works by Kjærulff and Madsen [17] and Koller and Friedman [92] for information about BNs.

#### 5.1.4 Outline

This paper is organized as follows. Section 5.2 explains the approach to model a MUSPRA using BNs by describing the single-unit seismic PRA used as the basis for the MUSPRA (Section 5.2.1), the mapping of the event tree and fault tree logic into a BN (Section 5.2.2), and the modeling of the MUSPRA as a BN (Section 5.2.3). Section 5.3 presents the NPP site-based risk metric results and discusses the rationale (insights) for the observed result trends. Finally, Section 5.4 summarizes our work and provides conclusions.

#### 5.2 Approach to MUSPRA using Bayesian Networks

The approach used in this paper to model a MUSPRA using BNs is summarized as follows:

- Adopt the Modarres et al. [82] methodology, that is, create the MUPRA model based on the single-unit PRA model (i.e., the event trees and fault trees). The single-unit seismic PRA is described in 5.2.1.
- 2. Map the event trees and fault trees into an equivalent Bayesian network; the event tree pivotal events (fault trees) are modeled based on the minimal cut sets. The mapping is described in Section 5.2.2.
- 3. Create nodes in the BN that explicitly address the dependencies between the SSCs across the units and estimate the NPP site-based risk metrics. In terms of a seismic PRA, the explicit nodes that address dependencies are the ones used to model the spatial variability of ground motion and dependent seismic failures. This process is described in Section 5.2.3.

#### 5.2.1 Single-Unit Seismic PRA Description

The single-unit seismic PRA used in this paper is a simplified version of the PRA-based seismic margin analysis (SMA) documented in Section 19.1.5.1 of the Design Control Document (DCD) for the design certification of the Advanced Power Reactor 1400 (APR1400) NPP [93], [94]. The at-power seismic initiating event tree is shown in Figure 5-1, and it models the pivotal events leading directly to core damage:

- seismic-induced failures leading to direct core damage in Unit 1 (1-01-S-DMG),
- seismic-induced station blackout (SBO) in Unit 1 (1-02-S-SBO),
- seismic-induced anticipated transient without scram (ATWS) in Unit 1 (1-06-S-ATWS),
- seismic-induced large loss-of-coolant accident (LOCA) in Unit 1 (1-07-S-LLOCA), and

• seismic-induced medium LOCA (1-08-S-MLOCA) in Unit 1.<sup>20</sup>

Seismic-induced small LOCA in Unit 1 (1-09-S-SLOCA) and loss of offsite power (LOOP) in all units (10-S-LOOP) are also included in the seismic initiating event tree, but the modeling is terminated at the small LOCA, LOOP, and transient (TRANS) end states to keep this problem as a relatively simple seismic PRA. The LOOP pivotal event was assumed to affect all units based on the Schroer and Modarres event classification schema [23], which categorizes LOOP events as "definite" initiating events.



#### Figure 5-1 At-power seismic initiating event tree

The pivotal events shown in Figure 5-1 are each associated with a top event of the following fault trees:

<sup>&</sup>lt;sup>20</sup> Some simplifications to the PRA-based SMA involve not modeling the loss of all instrumentation and control (1-03-S-IC), seismic-induced main steam line break (1-04-S-MSLB), and seismic-induced total loss of component cooling water (1-05-S-TLOCCW), which result in core damage, and reducing the number of trains in the 1-02-S-SBO fault tree from four to two.

- The "Seismic event leading to direct core damage in Unit 1" pivotal event (top event in Figure 5-2) models the seismic-induced failure of major NPP structures and reactor coolant system (RCS) components. The RCS of the APR1400 is arranged as two closed loops connected in parallel to the reactor vessel. Each loop consists of one outlet hot leg, one steam generator (SG), two cold legs, and two reactor coolant pumps (RCPs) [95].
- The "Seismic-induced SBO in Unit 1" pivotal event (top event in Figure 5-3) models the failure of the major components of the onsite electrical power system (EPS), including two redundant onsite EPS trains and the emergency diesel generator (EDG) building.
- The "Seismic-induced ATWS in Unit 1" pivotal event (top event in Figure 5-4(a)) models the failure due to the binding of the control element drive mechanism extension shaft.
- The different size seismic-induced LOCA pivotal events are due to the seismic-induced failure of the pressurizer surge line (top event in Figure 5-4(b)), pressurizer spray nozzle (top event in Figure 5-4(c)), and RCS small piping (top event in Figure 5-4(d)).
- The "Seismic-induced LOOP in all units" pivotal event (top event in Figure 5-4(e)) is due to the failure of the ceramic insulators.

The fragility parameters for the basic events shown in Figure 5-2 through Figure 5-4 are provided in Table 5-1.



Figure 5-2 Fault tree for the seismic event leading to direct core damage (1-01-S-DMG) pivotal event



Figure 5-3 Fault tree for the seismic-induced SBO (1-02-S-SBO) pivotal event



# Figure 5-4 Fault trees for the seismic-induced (a) ATWS (1-06-S-ATWS), (b) large LOCA (1-07-S-LLOCA), (c) medium LOCA (1-08-S-MLOCA), (d) small LOCA (1-09-S-SLOCA), and (e) LOOP (10-S-LOOP) pivotal events

Table 5-1	Fragility Parameters for b	asic ever	nts		
Basic Event(s)	Failure Mode	$\Theta_m(\mathbf{g})$	$\beta_R$	$\beta_{U}$	Reference
1-01-AUXB	Shear wall	1.48	0.31	0.26	Record # 1212 of [96]
1-01-NI	Sliding	1.3	0.23	0.42	Record # 1243 of [96]
1-01-TB	NS braced frame	1.48	0.33	0.3	Record # 1297 of [96]
1-01-CTMT	Shear	2.46	0.26	0.29	Record # 173 of [96]
1-01-RPV	Structural failure	1.8	0.3	0.4	Table 6-1 of [70]
1-01-PZR	Structural failure of support	2.5	0.3	0.4	Table 6-1 of [70]
1-01-SG1 1-01-SG2	Nozzle	2.46	0.4	0.5	Record # 302 of [96] (See Note 1)
1-01-RCP1A 1-01-RCP1B 1-01-RCP2A 1-01-RCP2B	Structural failure of support	2.5	0.3	0.4	Table 6-1 of [70]
1-02-EDGB	Shear wall	2.09	0.31	0.26	Record # 1214 of [96]
1-02-EDG1A 1-02-EDG1B	Structural	0.92	0.12	0.23	Record # 277 of [96]
1-02-SWGR1A 1-02-SWGR1B	Structural	3.26	0.26	0.25	Record # 1183 of [96]
1-02-DCCC1A 1-02-DCCC1B	Structural	1.48	0.23	0.28	Record # 1181 of [96]
1-02-BATT1A 1-02-BATT1B	Functional	1.94	0.24	0.48	Record # 10 of [96]
1-06-CEDM	Excess bending	2	0.3	0.4	Record # 214 of [96]
1-07-PZR-SLN	Structural failure of pressurizer surge line nozzle	1.73	0.3	0.4	(See Note 2)
1-08-PZR-SPN	Structural failure of pressurizer spray nozzle	1.61	0.3	0.4	(See Note 2)
1-09-RCS-SPI	Generic	0.95	0.3	0.4	(See Note 2)
10-CERINS	Failure of ceramic insulators	0.3	0.3	0.45	Table 6-1 of [70]
Notes					

1. The  $\beta_R$  and  $\beta_U$  values were selected such that the resulting high confidence of low probability of failure (HCLPF) capacity is approximately the HCLPF capacity provided in the APR1400 DCD [93] using the following equation:  $HCLPF = \Theta_m \cdot \exp[-1.65 * (\beta_R + \beta_U)]$ .

2. The  $\beta_R$  and  $\beta_U$  values are generic and the  $\Theta_m$  values were calculated from the HCLPF capacity provided in the APR1400 DCD [93] using the following equation:  $\Theta_m = HCLPF / \exp[-1.65 * (\beta_R + \beta_U)]$ .

#### 5.2.2 Mapping of the Single-Unit Seismic PRA into a Bayesian Network

Because seismic PRAs are typically carried out using event trees and fault trees, in the BN-based MUSPRA, we need to convert the event tree and fault tree logic into a BN. Bobbio et al. [97] proposed an approach to map fault trees into BNs, which we initially adopt in this study. The resulting BN from a fault tree with a large number of minimal cut sets (MCS) may be inefficient (and potentially intractable) to quantify because of a converging structure, as shown in Figure 5-5 along with the corresponding conditional probability table (CPT) for the *System* node.



Figure 5-5 Example of a converging BN structure for a series system with four MCS with its CPT

To overcome this potential inefficiency, we use the method proposed by Bensi et al. [73], where the converging BN structure shown in Figure 5-5 can be modeled as a chain-like BN structure as shown in Figure 5-6. In Figure 5-6, the nodes  $FPE_i$ , i = 1, 2, 3, 4, represent a failure path sequence, which Bensi et al. [73] define as "a chain of events, corresponding to a MCS, in which the terminal event in the sequence indicates whether or not all components in the MCS are in the fail state" and each node in the failure path sequence (i.e.,  $FPE_1$ ,  $FPE_2$ ,  $FPE_3$ , and  $FPE_4$ ) is a failure path event (FPE). In the case of Figure 5-6, the chain of events corresponds to the system with the terminal event (i.e.,  $FPE_4$ ) indicating whether any MCS is in the fail state.

				$\sim$		_								
CPT fo	or $FPE_1$		(					(MC			CPT f	or $FPE_4$		
$MCS_1$	Fail	Safe		$MCS_1$	$MCS_2$	) ( <sup>m</sup>	$\iota S_3$		.34	FPE <sub>3</sub>	F	ail	Si	afe
$\mathit{FPE}_1$ Fail	1	0		×	<b>T</b>		Ŧ		$\leq$	MCS <sub>4</sub>	Fail	Safe	Fail	Safe
$FPE_1$ Safe	0	1	(					Гер		FPE <sub>4</sub> Fail	1	1	1	0
				<i>ГРЕ</i> 1		<u>г</u> ,				FPE <sub>4</sub> Safe	0	0	0	1
	CPT f	or FPE			$\bigcirc$				$\leq$	CPT fo	r Syste	<i>m</i>	7	
	<b>.</b>	0.1122	-					(					4	
$FPE_1$	F	ail	S	afe				Syst	tem )	FPE <sub>4</sub>	Fail	Safe		
$MCS_2$	Fail	Safe	Fail	Safe				$\sim$	$\leq$	System Fail	1	0		
$FPE_2$ Fail	1	1	1	0		CPT f	or FPE <sub>3</sub>			System Safe	0	1	1	
FPE <sub>2</sub> Safe	0	0	0	1	FPE <sub>2</sub>	F	ail	S	afe		-		-	
					MCS <sub>3</sub>	Fail	Safe	Fail	Safe					
					FPE <sub>3</sub> Fail	1	1	1	0					
					FPE <sub>3</sub> Safe	0	0	0	1					

Figure 5-6 Example of a chain-like BN structure for a series system with four MCS with its CPT and the CPTs of the chain-like nodes

The efficiency gains are realized by the size of the CPT (or the number of entries) for the *System* (child) node. In general, the size of the CPT for the child node in a converging structure where the parent nodes have binary states is  $2^{(\#of \ parent \ nodes)} \cdot (\#of \ states \ of \ child \ node)$ . In Figure 5-5, the size of the CPT for the *System* node is 32, whereas the size of the CPT for the *System* node in Figure 5-6 is 4.

The event tree in Figure 5-1 also needs to be mapped into a BN. We also use a chain-like BN structure to limit the size of the CPT of the node representing the state of the unit, which has four states (i.e., core damage, S-SLOCA, S-SLOOP, and TRANS). Since the event tree in Figure 5-1 has seven binary-state pivotal events, the node representing the state of the unit would have 512  $(= 2^7 \cdot 4)$  entries as a converging BN structure. The chain-like BN structure is shown in Figure 5-7, where the size of the CPT for the node representing the state of the unit is reduced to 16.

CPT for 1	-FPE <sub>DMG</sub>	i		C	CPT for 1	-FPE <sub>SBO</sub>					с	PT 1	for 1-F	PEATWS					CPT for :	L-FPE <sub>LLOC</sub>		
1-01-S-DMG	Fail	Safe	1-FI	PE <sub>DMG</sub>	Unit	1 CD	Unit 1	Not C	D	1-FP	E <sub>SBO</sub>		Unit 1	. CD	Unit 1	Not CD	1-FP	E <sub>ATWS</sub>	Uni	t 1 CD	Unit 1	Not CD
Unit 1 CD	1	0	1-02	-S-SBO	Fail	Safe	Fail	Safe	e	1-06-S-	06-S-ATWS Fail Safe I			Fail	Safe	1-07-S	-LLOCA	Fail	Safe	Fail	Safe	
Unit 1 Not CD	0	1	Unit	t 1 CD	1	1	1	0		Unit	1 CD		1	1	1	0	Unit	1 CD	1	1	1	0
			Unit 1	Not CD	0	0	0	1		Unit 1 I	Not CD		0	0	0	1	Unit 1	Not CD	0	0	0	1
$\frown$	$\frown \frown \frown \frown \frown \frown \frown$																					
( 1-01-S- DMG ) (	1-02 SB	2-S-	( 1-06- ATW	-s- /s ) (	1-07-5 LLOC	s- )	( 1-0 ML	18-S- OCA	)	( 1-09- SLOC	S- CA	(	10-S-	P)				CPT for 2	L-FPE <sub>LOO</sub>			
$\checkmark$			$\searrow$		$\checkmark$		$\overline{\ }$	-		$\searrow$		/	-		1-	FPE <sub>SLOCA</sub>	Unit	1 CD	Unit 1	SLOCA	Unit 1	Not CD
		$\sim$		$\overline{}$	1			1				/	1		10	-S-LOOP	Fail	Safe	Fail	Safe	Fail	Safe
$\begin{pmatrix} 1^{-} \\ FPE_{DMG} \end{pmatrix} \leftarrow \begin{pmatrix} 1^{-} \\ FPE_{SBO} \end{pmatrix} \leftarrow \begin{pmatrix} 1^{-} \\ FPE_{ATWS} \end{pmatrix} \leftarrow \begin{pmatrix} 1^{-} \\ FPE_{LICCA} \end{pmatrix} \leftarrow \begin{pmatrix} 1^{-} \\ FPE_{MICCA} \end{pmatrix} \leftarrow \begin{pmatrix} 1^{-} \\ FPE_{SICCA} \end{pmatrix} \leftarrow \begin{pmatrix} 1^{-} \\ FPE_{SICCA}$													0									
				/ `				_					¥		Uni	t 1 SLOCA	0	0	1	1	0	0
	CPT for	1-FPE <sub>MIO</sub>	~ <u>^</u>									(			Un	it 1 LOOP	0	0	0	0	1	0
1-FPE	Un	nit 1 CD	Unit 1	Not CD									Unit	1)	Uni	t 1 TRANS	0	0	0	0	0	1
1-08-S-MLOCA	Fail	Safe	Fail	Safe																		
Unit 1 CD	1	1	1	0			CPT fo	or 1-Fl	PE <sub>SLOC</sub>	A							CPT fo	or Unit 1				
Unit 1 Not CD	0	0	0	1	1-1	PE <sub>MLOCA</sub>	l	Jnit 1	CD	Unit 1	Not CD		1-	-FPE <sub>LOOP</sub>	,	Unit 1 CD	Unit 1	SLOCA	Unit :	L LOOP	Unit 1	TRANS
					1-09	-S-SLOC	A Fa	ail	Safe	Fail	Safe		U	nit 1 CD		1	C	)		0		0
					U	nit 1 CD	:	L	1	0	0		Uni	t 1 SLOO	CA	0	1			0		0
					Uni	t 1 SLOCA	A (	)	0	1	0		Uni	it 1 LOO	P	0	C	)		1		0
					Unit	1 Not Cl	D (	)	0	0	1		Unit	t 1 TRAI	NS	0	C	)		0		1

## Figure 5-7 Event tree in Figure 5-1 mapped as a chain-like BN structure with the CPTs for the failure path events and Unit 1

#### 5.2.3 Modeling MUSPRA as a Bayesian Network

At a high level, the BN that models a MUSPRA can be represented by three successive modules that are related to the elements of a seismic PRA. The modules are shown in Figure 5-8 as an object-oriented BN. The objects are described in Sections 5.2.3.1, 5.2.3.2, and 5.2.3.3.



Figure 5-8 BN modules to model a MUSPRA

#### 5.2.3.1 Seismic Demand Module – Spatial Variability of Ground Motion

The hypothetical NPP site is located at a representative Central Illinois site used in the demonstration seismic hazard calculations in NUREG-2115 [48] and the EPRI Ground Motion Model review project [49], which was used in our previous work [64]. The site is assumed to have two identical units separated by 100 m.

The seismic hazard curve obtained from the PSHA process is typically generated for a reference location on the site. However, the ground motion experienced across the site during an earthquake will differ from the ground motion experienced at the reference location [24]–[26], [28], [29]. In this work, the representative location is assumed to be at Unit 1. Consistent with the procedure outlined in [64], the ground motion affecting Unit 2 is then modeled as a function of the ground motion at Unit 1 using a model for spatial variability of ground motion in which variability is a function of the separation distance between units.

The BN that models the ground motion affecting each unit is shown in Figure 5-9, where the nodes 1-GM and 2-GM represent the ground motion events affecting Unit 1 and Unit 2, respectively; node  $P'_{1,2}$  represents the spatial variability of ground motion model; and node EES represents an earthquake of engineering significance (EES) event. An EES event occurs when the ground motion exceeds a minimum threshold. In the context of a seismic PRA, the minimum threshold value is the starting value of the first ground motion bin in the discretized seismic hazard curve. All inferences performed using the BN in this study are conditioned on the occurrence of an EES event, which is the reason why the node EES in Figure 5-9 is shaded. Thus, to estimate the "unconditional" risk, the results provided by the BN must be multiplied by the frequency of an EES event, which is  $\lambda(PGA > 0.08g) = 4.62 \times 10^{-4}/yr$  in this study.

Figure 5-9 also provides the CPTs for all the nodes. The details about generating the CPTs are provided in our previous work [64]. We note that the CPT for  $P'_{1,2}$  has 31 entries and, consequently, the CPT for 2-GM has 1984 (=  $8 \cdot 31 \cdot 8$ ) entries. This makes it impractical to present the entire CPTs for those nodes and we only show a "snippet" of the CPTs in Figure 5-9.

CPT for 1	L-GM	
EES	Yes	No
1-BIN-0; 0g – 0.08g	0	1
1-BIN-1; 0.08g – 0.18g	$7.94  imes 10^{-1}$	0
1-BIN-2; 0.18g – 0.40g	$1.64  imes 10^{-1}$	0
1-BIN-3; 0.40g – 0.56g	$2.07 \times 10^{-2}$	0
1-BIN-4; 0.56g – 0.86g	$1.28 \times 10^{-2}$	0
1-BIN-5; 0.86g – 1.50g	$6.26 \times 10^{-3}$	0
1-BIN-6; 1.50g – 3.32g	$1.99 \times 10^{-3}$	0
1-BIN-7; > 3.32g	$2.19 \times 10^{-4}$	0



CPT for <i>P</i> ' <sub>1,2</sub>	
Pp1; -0.57 – -0.53	$2.50  imes 10^{-3}$
Pp2; -0.530.50	$1.98  imes 10^{-3}$
:	
Pp16; -0.018 - 0.018	$7.71 \times 10^{-2}$
:	
Pp30; 0.50 – 0.53	$1.98 \times 10^{-3}$
Pp31; 0.53 – 0.57	$2.50  imes 10^{-3}$

	"Snippet" of CPT for 2-GM																
1-GM	1-BIN-0	1-BIN-1				1-BIN-4						1-BIN-7					
P'_{1,2}	Pp1 – Pp31	Pp1		Pp16		Pp31		Pp1		Pp16		Pp31	 Pp1		Pp16		Pp31
2-BIN-0; 0g – 0.08g	_	0.5865		0.0031		0		0		0		0	0		0		0
2-BIN-1; 0.08g - 0.18g	entries do not	0.4135		0.9982		0.2423		0		0		0	0		0		0
2-BIN-2; 0.18g - 0.40g	impact the	0		0.0087		0.7577		0.4476		0		0	0		0		0
2-BIN-3; 0.40g - 0.56g	all inferences are	0		0		0		0.5524		0.0092		0	0		0		0
2-BIN-4; 0.56g - 0.86g	performed	0		0		0		0		0.978		0	0		0		0
2-BIN-5; 0.86g - 1.50g	Yes" and, therefore,	0		0		0		0		0.0128		0.9931	0		0		0
2-BIN-6; 1.50g - 3.32g	P(1-BIN-0 EES Yes)	0		0		0		0		0		0.0069	1		0.497		0
2-BIN-7; > 3.32g	15 cquar to zero.	0		0		0		0		0		0	0		0.503		1

Figure 5-9 BN and CPTs used to model the spatial variability of ground motion

#### 5.2.3.2 System Performance Module – Dependent Seismic Failures

The seismic capacity for the failure mode of interest of an SSC is typically modeled as the product of a "best estimate" seismic capacity (usually the median,  $\Theta_m$ ) and two random variables  $E_R$  and  $E_U$  [66]:

$$\Theta = \Theta_{\rm m} \cdot {\rm E}_R \cdot {\rm E}_U \tag{5.2}$$

The random variable  $E_R$  represents the aleatory variability about the median and  $E_U$  represents the epistemic uncertainty in the median value. These two random variables are assumed lognormally distributed with medians of 1 and logarithmic standard deviations  $\beta_R$  and  $\beta_U$ , respectively. Consequently, the seismic capacity  $\Theta$  is a lognormally distributed random variable with median  $\Theta_m$  and logarithmic standard deviation  $(\beta_R^2 + \beta_U^2)^{0.5}$ .

When a system consists of multiple SSC, there is correlation in the fragility across the SSCs because of correlation in seismic response and correlation in seismic capacity [53]. This is

modeled in the fragility analysis by separating the components of uncertainty that are common or SSC-specific [15], [56]. Specifically, by taking the natural logarithm of both sides of Equation (5.2), separating the aleatory variability ( $E_U$ ) and epistemic uncertainty ( $E_R$ ) into their common ( $E_{R_{ij}}^*$  and  $E_{U_{ij}}^*$ ) and SSC-specific ( $E_{R_i}'$  and  $E_{U_i}'$ ) parts, and using the following definitions  $\varepsilon_{ij}^* = \ln(E_{ij}^*) = \ln(E_{R_{ij}}^*) + \ln(E_{U_{ij}}^*)$  and  $\varepsilon_i' = \ln(E_i') = \ln(E_{R_i}') + \ln(E_{U_i}')$ , the logarithm of the capacity of SSC *i* is expressed as [83]:

$$\ln(\Theta_i) = \ln(\Theta_{m_i}) + \varepsilon'_i + \sum_{i \neq j} [\varepsilon^*_{ij}]$$
(5.3)

In Equation (5.3),  $\Theta_i$  is the seismic capacity of SSC *i*,  $\Theta_{m_i}$  is the median seismic capacity of SSC *i*,  $\varepsilon_{ij}^r$  represents the SSC-specific capacity uncertainty of SSC *i*, and  $\varepsilon_{ij}^*$  represents the common sources of uncertainty between SSCs *i* and *j* ( $i \neq j$ ). Given the definitions of  $\varepsilon_{ij}^*$  and  $\varepsilon_i'$  in the paragraph above,  $\varepsilon_{ij}^*$  and  $\varepsilon_i'$  are normally distributed random variables with zero mean and standard deviations  $\beta_{C_{ij}}^* = \left[ \left( \beta_{R_{ij}}^* \right)^2 + \left( \beta_{U_{ij}}^* \right)^2 \right]^{0.5}$  and  $\beta_{C_i}' = \left[ \left( \beta_{R_i}' \right)^2 + \left( \beta_{U_i}' \right)^2 \right]^{0.5}$ , respectively. From the BN perspective, the dependent seismic failures are addressed by explicitly modeling nodes that represent the common sources of uncertainties between the SSCs of interest, that is, the  $\varepsilon_{ij}^*$  nodes [83]. The approach to modeling dependent seismic failures relies upon a systematic search for the common sources of uncertainty in the individual factors of safety that result in the seismic capacities of the SSCs of interest [15], [56]. The systematic search is performed by the fragility analyst and (when combined with detailed knowledge about SSCs and expert judgement) results in parameters (logarithmic standard deviations) that represent the common aleatory variability ( $\beta_{R_{ij}}^*$ ) and epistemic uncertainty ( $\beta_{U_{ij}}^*$ ) between each pair of SSCs *i* and *j*.

The works by Campbell et al. [96] and Budnitz et al. [70] provide  $\beta_U$  and  $\beta_R$  parameters for the SSCs within the event and fault trees presented in Section 5.2.1, which are presented in Table 5-1. For simplicity and due to lack of detailed information about the process used to derive the  $\beta_R$  and  $\beta_U$  parameters shown in Table 5-1, in this study, we use an approach adapted from the EPRI MUPRA framework [13] to estimate the  $\beta_{R_{ij}}^*$  and  $\beta_{U_{ij}}^*$  parameters. The process and results of the simplified approach to estimate the  $\beta_{R_{ij}}^*$  and  $\beta_{U_{ij}}^*$  parameters are shown in Table 5-2. Once the  $\beta_{R_{ij}}^*$  and  $\beta_{U_{ij}}^*$  parameters are calculated as  $\beta_{R_i}' =$ 

$$\left[\left(\beta_{R_i}\right)^2 - \sum \left(\beta_{R_{ij}}^*\right)^2\right]^{0.5} \text{ and } \beta_{U_i}' = \left[\left(\beta_{U_i}\right)^2 - \sum \left(\beta_{U_{ij}}^*\right)^2\right]^{0.5}, \text{ respectively [15], [56].}$$

SSC Group	<b>Judged</b> $\rho_{R_{ij}} = \rho_{U_{ij}}$	$\boldsymbol{\beta}_{R_{ij}}^* = \sqrt{\boldsymbol{\rho}_{R_{ij}} \cdot \left(\boldsymbol{\beta}_{R_i}\right)^2}$	$\boldsymbol{\beta}_{U_{ij}}^* = \sqrt{\boldsymbol{\rho}_{U_{ij}} \cdot \left(\boldsymbol{\beta}_{U_i}\right)^2}$
1-01-AUXB, 2-01-AUXB	0.5	0.22	0.18
1-01-NI, 2-01-NI	0.5	0.16	0.30
1-01-TB, 2-01-TB	0.5	0.23	0.21
1-01-CTMT, 2-01-CTMT	0.5	0.18	0.21
1-01-RPV, 2-01-RPV	0.5	0.21	0.28
1-01-PZR, 2-01-PZR	0.5	0.21	0.21
1-01-SG1, 1-01-SG1, 2-01-SG1, 2-01-SG2	0.8	0.36	0.45
1-01-RCP1A, 1-01-RCP1B, 1-01-RCP2A, 1-01-RCP2B, 2-01-RCP1A, 2-01-RCP1B, 2-01-RCP2A, 2-01-RCP2B	0.5	0.21	0.28
1-02-EDGB, 2-02-EDGB	0.5	0.22	0.18
1-02-EDG1A, 1-02-EDG1B, 2-02-EDG1A, 2-02-EDG1B	0.5	0.08	0.16
1-02-SWGR1A, 1-02-SWGR1B, 2-02-SWGR1A, 2-02-SWGR1B	0.5	0.18	0.18
1-02-DCCC1A, 1-02-DCCC1B, 2-02-DCCC1A, 2-02-DCCC1B	0.5	0.16	0.20
1-02-BATT1A, 1-02-BATT1B, 2-02-BATT1A, 2-02-BATT2B	0.8	0.21	0.43
1-06-CEDM, 2-06-CEDM	0.5	0.21	0.28
1-07-PZRSLN, 2-07-PZRSLN	0.5	0.21	0.28
1-08-PZRSPN, 2-08-PZRSPN	0.5	0.21	0.28
1-09-RCSSPI, 2-09-RCSSPI	0.5	0.21	0.28

Table 5-2Common Aleatory and Epistemic Uncertainties

Using the fragility model framework described above, we create BNs for modeling the correlated performance of components contained within each fault tree. We then generate the CPTs for the

seismic performance of each SSC as described in our previous work [83], where the only difference is that the random variable  $\varepsilon'_i$  was integrated into the node representing the seismic performance of the SSC to reduce the number of nodes in the BN and, therefore, reduce the discretization error.

An example BN representing the 1-02-S-SBO fault tree is shown in Figure 5-10. In the BN, all SSCs within Unit 1 (nodes 1-02-EDG1A, 1-02-EDG1B, 1-02-DCCC1A, 1-02-DCCC1B, 1-02-BATT1A, 1-02-BATT1B, 1-02-SWGR1A, 1-02-SWGR1B, 1-02-EDGB) are subject to the same ground motion, as reflected by their common parent node 1-GM. Pairs of SSCs for which the fragility may be correlated have a common " $\varepsilon^*$ -node." For example, the nodes representing the states of the emergency diesel generators (1-02-EDG1A and 1-02-EDG1B) have a common parent node  $\varepsilon^*_{EDG}$ . The nodes representing the MCSs have parents containing their constituent SSCs. The nodes labeled FPE are used to model system performance using the approach described in Section 5.2.2.

Figure 5-11 shows a "snippet" of the CPT for (1) the node representing the state of SSC 1-02-EDG1A, (2) the  $\varepsilon_{EDG}^*$  random variable; (3) a few MCS and FPEs; and (4) the fault tree top event (1-02-S-SBO). The CPTs for MCS1<sub>1-02-S-SBO</sub> through MCS16<sub>1-02-S-SBO</sub> have identical structures representing the AND failure logic of the parent nodes, and the differences are the SSCs that are the parent nodes to the particular MCS node. For the remaining top events in Figure 5-1, we followed the same approach as described above for the 1-02-S-SBO event.





CPT for $\varepsilon_{ED}^*$	G
$\varepsilon^*_{EDG}$ 1; -0.55 – -0.45	$7.05 \times 10^{-3}$
:	
$\varepsilon_{EDG}^{*}6$ ; -0.05 – 0.05	$2.15\times10^{-1}$
:	
$\varepsilon_{EDG}^*$ 11; 0.45 – 0.55	$7.05 \times 10^{-3}$

	"Snippet" of CPT for 1-02-EDG1A (1-02-EDG1B, 2-02-EDG1A, 2-02-EDG1B)																	
1-GM	1-BIN-0				1-BIN-3					1-BIN-4			··· 1-BIN-7					
$\varepsilon_{EDG}^{*}$	$\varepsilon^*_{EDG}1-\varepsilon^*_{EDG}11$		$\varepsilon^*_{EDG}$ 1		$\varepsilon_{EDG}^{*}6$		$\varepsilon^*_{EDG}$ 11	$\varepsilon^*_{EDG} 1$		$\varepsilon^*_{EDG}6$		$\varepsilon^*_{EDG}$ 11		$\varepsilon^*_{EDG} 1$		$\varepsilon^*_{EDG}6$		$\varepsilon^*_{EDG} 11$
1-02-EDG1A Fail	0		0.214		0.0007		0	0.837		0.117		0.0002		1		1		1
1-02-EDG1A Safe	1		0.786		0.9993		1	0.163		0.883		0.9998		0		0		0

CPT for MCS1 <sub>1-02-S-SBO</sub>											
1-02-EDG1A	1-02-EDG1A Fail Safe										
1-02-EDG1B	Fail	Safe	Fail	Safe							
MCS1 <sub>1-02-S-SBO</sub> Fail	1	0	0	0							
MCS1 <sub>1-02-S-SBO</sub> Safe	0	1	1	1							

CPT for 1-FPE1 <sub>SBO</sub>					
MCS1 <sub>1-02-S-SBO</sub>	Fail	Safe			
1-FPE1 <sub>SBO</sub> Fail	1	0			
1-FPE1 <sub>SBO</sub> Safe	0	1			

CPT for MCS9 <sub>1-02-S-SBO</sub>							
1-02-DCCC1A	Fail Safe						
1-02-BATT1B	Fail	Safe	Fail	Safe			
MCS9 <sub>1-02-S-SBO</sub> Fail	1	0	0	0			
MCS9 <sub>1-02-S-SBO</sub> Safe	0	1	1	1			

CPT for 1-FPE(2,, 17) <sub>SBO</sub>							
1-FPE(1,,16) <sub>SBO</sub>	F	ail	Safe				
MCS(2,, 17) <sub>1-02-S-SBO</sub>	Fail	Safe	Fail	Safe			
1-FPE(2,, 17) <sub>SBO</sub> Fail	1	1	1	0			
1-FPE(2,, 17) <sub>SBO</sub> Safe	0	0	0	1			

CPT for 1-0	2-S-SBC	)
1-FPE17 <sub>SBO</sub>	Fail	Safe
1-02-S-SBO Fail	1	0

CPT for MCS17<sub>1-02-S-SBO</sub>

Fail Safe

1 0

0 1

0

1-02-EDGB

MCS17<sub>1-02-S-SBO</sub> Fail

MCS17<sub>1-02-S-SBO</sub> Safe

1-02-S-SBO Safe

Figure 5-11	CPTs for various nodes in the BN representing the 1-02-S-SBO fault tree

### 5.2.3.3 Site-Based Risk Metrics Module

The seismic PRA model for Unit 2 is the same as that of Unit 1 because we assumed that both units are identical. Therefore, we created a BN model for Unit 2 that is identical to the BN model for Unit 1. We then integrated the "unit-specific" BNs into a single multi-unit BN model that accounts for the spatial variability of ground motion between units, the correlation in SSC fragility across units, and includes a node to calculate site-level risk metrics. This process results in the BN shown in Figure 5-12, which includes the spatial variability of ground motion nodes (i.e., nodes EES, 1-GM,  $P'_{1,2}$ , and 2-GM), nodes capturing the dependent seismic failures (e.g., nodes  $\varepsilon^*_{AUXB}$ ,  $\varepsilon^*_{EDG}$ ,  $\varepsilon^*_{CEDM}$ , etc.) and the site state (i.e., node "Site").

Given the four end states for each unit (i.e., core damage (CD), S-SLOCA, S-SLOOP, and TRANS), the node Site has 16 states, and its CPT is shown in Figure 5-13.<sup>21</sup> After performing the inference with the BN, the concurrent core damage probability is the probability of the state labeled "U1CD\_U2CD" state in Figure 5-13. The site core damage probability is estimated by summing the probabilities for states in which any unit is in a core damage state (i.e., the states "U1CD\_U2CD," "U1CD\_U2SLOCA," "U1CD\_U2LOOP," "U1CD\_U2TRANS," "U1SLOCA\_U2CD," "U1LOOP\_U2CD," and "U1TRANS\_U2CD").

<sup>&</sup>lt;sup>21</sup> If the analysis of each unit had been carried out so that the end states are CD and OK, then the CPT for the Site node would be much simpler than the CPT for the Site node shown in Figure 5-13.



Figure 5-12	BN model	of MUSPRA
-------------	----------	-----------

CPT for Site																
Unit 1 (U1)		Core Da	mage (C	D)		SLOCA		LOOP			TRANS					
Unit 2 (U2)	CD	SLOCA	LOOP	TRANS	CD	SLOCA	LOOP	TRANS	CD	SLOCA	LOOP	TRANS	CD	SLOCA	LOOP	TRANS
U1CD_U2CD	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U1CD_U2SLOCA	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U1CD_U2LOOP	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
U1CD_U2TRANS	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
U1SLOCA_U2CD	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
U1SLOCA_U2SLOCA	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
U1SLOCA_U2LOOP	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
U1SLOCA_U2TRANS	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
U1LOOP_U2CD	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
U1LOOP_U2SLOCA	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
U1LOOP_U2LOOP	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
U1LOOP_U2TRANS	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
U1TRANS_U2CD	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
U1TRANS_U2SLOCA	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
U1TRANS_U2LOOP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
U1TRANS_U2TRANS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Figure 5-13	CPT for the Site node
-------------	-----------------------

## 5.3 <u>Results and Discussion</u>

To assess the impact of considering the spatial variability of ground motion and dependent seismic failures in the NPP site-based risk metrics, we developed several cases. The case studies

vary the degree of correlation (dependency) and calculate the site CDF and concurrent CDF risk metrics using BNs. For the ground motion (GM) correlation, we considered three assumptions:

- I. No GM correlation The ground motion experienced at each unit is assumed to be independent of the other. The BN for this alternative is shown in Figure 5-14(a). The CPTs for nodes EES and 1-GM are the same as those shown in Figure 5-9 and, in this case, the CPT for node 2-GM is identical to the CPT for node 1-GM. Given that the EES event is observed, the d-separation property of a common cause applies [92], that is, 1-GM is conditionally independent of 2-GM given an EES. From a site-scale perspective [64], this "no GM correlation assumption" is not realistic and it only serves as a bounding case.
- II. *Partial GM correlation* This uses the partial correlation formulation described in Section 5.2.3.1.
- III. *Perfect GM correlation* This is the typical assumption used in current MUSPRAs (e.g., [13]).We implement it as shown in Figure 5-14(b), which includes the CPT for node 2-GM. The CPTs for nodes EES and 1-GM are the same as those shown in Figure 5-9.



Figure 5-14 (a) BN for the no ground motion correlation assumption. (b) BN for the perfect ground motion correlation assumption and CPT for node 2-GM

With respect to the dependent (correlated) seismic failures, we considered five "fragility correlation" case assumptions and describe them below:

- No SSC correlation The seismic performance of each SSC across the hypothetical site is based on the ground motion affecting the particular unit where the SSC is located and its uncertainty; however, there is no correlation in the performance of SSCs given a GM. From a BN perspective, this assumption means that the only common parent nodes are associated with the seismic demands. This formulation is illustrated in Figure 5-15(a) for an example involving two EDGs.
- ii. *Partial-within unit SSC correlation, and no between units SSC correlation* Under this assumption, the seismic performance of each SSC depends on the ground motion affecting the particular unit where the SSC is located. In addition, there is partial correlation between the performance of the SSCs within a unit and no correlation between the SSCs across units. The BN representing this assumption is shown as Figure 5-15(b), where the EDGs within a unit share a common parent node (i.e.,  $\varepsilon_{EDG}^*$ ).
- iii. *Partial within unit and partial between units SSC correlation* This assumption uses the formulation described Section 5.2.3.2, where there is partial correlation for similar SSCs within and across units.
- iv. *Perfect within unit SSC correlation and partial between units SSC correlation* Under this assumption, the ground motion affects a specific SSC in a particular unit, and the seismic performance of the remaining SSCs in the unit are perfectly correlated with it. This is illustrated in Figure 5-15(c) with EDG1B shown as a perfectly correlated child of EDG1A. The performance of the SSCs across the units that are directly affected by the ground motion are partially correlated and this is shown in Figure 5-15(c) with a common parent node  $\varepsilon_{EDG}^*$ . Figure 5-15(c) also shows the CPT for modeling the perfect correlation within a unit.

*Perfect SSC correlation* – Under this assumption, a single SSC is affected by the ground motion, and the remaining SSCs within the unit and across the site are perfectly correlated with this single SSC. The BN for this assumption is shown in Figure 5-15(d) with the CPT that implements the perfect correlation for the example SSC. This assumption is unrealistic and trivial; however, we consider it as a bounding case.



Figure 5-15 BN models for the (a) none; (b) partial-intra, none-inter; (c) perfect-intra, partialinter; and (d) perfect SSC fragility correlation assumptions

The combination of the three ground motion correlation assumptions with the five correlated seismic fragility assumptions results in 15 risk assessment cases, whose NPP site-based risk metrics are presented in Table 5-3. The results in Table 5-3 were calculated using the GeNIe software [22] and the Roman numerals in the "Fragility" column under "Correlation Assumptions" refer to the items above where these assumptions were described. Results are presented for the case when an EES has occurred as well as under a sensitivity case involving the survival of certain non-redundant SSCs.
Case #	Correlation Assumptions		Risk Metrics <sup>1</sup> given EES P(   EES)				Risk Metrics <sup>1</sup> given EES and Survival of 1-01-S-DMG and 2-01-S-DMG P(  EES, /1-01-S-DMG, /2-01-S-DMG)			
	GM	Fragility	U1 CD	U2 CD	Site CD	Conc CD	U1 CD	U2 CD	Site CD	Conc CD
1	None	i	1.54E-2	1.54E-2	3.06E-2	2.38E-4	2.63E-3	2.63E-3	5.24E-3	6.9E-6
2	None	ii	1.56E-2	1.56E-2	3.10E-2	2.45E-4	3.09E-3	3.09E-3	6.18E-3	9.56E-6
3	None	iii	1.56E-2	1.56E-2	3.10E-2	2.76E-4	3.1E-3	3.1E-3	6.18E-3	1.8E-5
4	None	iv	1.67E-2	1.67E-2	3.30E-2	3.04E-4	4.87E-3	4.87E-3	9.69E-3	3.71E-5
5	None	v	1.67E-2	1.67E-2	1.67E-2	1.67E-2	4.86E-3	4.86E-3	4.86E-3	4.86E-3
6	Partial	i	1.54E-2	1.89E-2	2.4E-2	1.04E-2	1.53E-3	2.31E-3	3.62E-3	2.13E-4
7	Partial	ii	1.56E-2	1.92E-2	2.42E-2	1.06E-2	1.81E-3	2.72E-3	4.23E-3	3.06E-4
8	Partial	iii	1.56E-2	1.92E-2	2.37E-2	1.11E-2	1.98E-3	2.92E-3	4.33E-3	5.76E-4
9	Partial	iv	1.67E-2	2.04E-2	2.51E-2	1.19E-2	3.26E-3	4.73E-3	6.82E-3	1.17E-3
10	Partial	v	1.67E-2	1.67E-2	1.67E-2	1.67E-2	4.86E-3	4.86E-3	4.86E-3	4.86E-3
11	Perfect	i	1.54E-2	1.54E-2	2.04E-2	1.04E-2	1.62E-3	1.62E-3	3E-3	2.43E-4
12	Perfect	ii	1.56E-2	1.56E-2	2.07E-2	1.06E-2	1.93E-3	1.93E-3	3.51E-3	3.53E-4
13	Perfect	iii	1.56E-2	1.56E-2	2.0E-2	1.13E-2	2.14E-3	2.14E-3	3.58E-3	7E-4
14	Perfect	iv	1.67E-2	1.67E-2	2.12E-2	1.21E-2	3.53E-3	3.53E-3	5.64E-3	1.42E-3
15	Perfect	V	1.67E-2	1.67E-2	1.67E-2	1.67E-2	4.86E-3	4.86E-3	4.86E-3	4.86E-3
<sup>1</sup> The risk metrics in this table must be multiplied by $4.62 \times 10^{-4}/yr$ to obtain the appropriate frequencies.										

Table 5-3Comparison of Results for NPP Site-Based Risk Metrics from 15 Cases

Comparing (a) Cases 6 and 11, (b) Cases 7 and 12, and (c) Cases 8 and 13, we observe that assuming a perfect correlation in ground motions (typical assumption) leads to an underestimation of the site CD probability [P(Site CD | EES)] while having a negligible effect on the concurrent CD probability [P(Concurrent CD | EES)] results. The underestimation of P(Site CD | EES) when using a perfect GM correlation assumption (relative to the more realistic partial GM correlation) is about 16%.<sup>22</sup> Physically, the perfect GM correlation assumption neglects the possibility that Unit 2 experiences a higher ground motion than Unit 1 given the occurrence of an EES event. Mathematically, as the correlation increases, the failure probability of a series system (akin to the site CD probability) decreases. The negligible effect on the P(Concurrent CD | EES) results is due to the existence of single SSC failures (i.e., single SSC MCSs), such as those in the

<sup>&</sup>lt;sup>22</sup> The percent differences mentioned in this discussion were calculated using the following equation:  $|x_1 - x_2|/[(x_1 + x_2)/2]$ .

1-01-S-DMG and 2-01-S-DMG fault trees, dominating the CD probability contribution from each unit to the P(Concurrent CD | EES) results.

To further investigate the impact of "single failures" on results, we ran another set of calculations assuming that the 1-01-S-DMG and 2-01-S-DMG fault trees (fault trees where any SSC failure results in the occurrence of the fault tree top event) are in their safe state to test whether there would be appreciable differences in the results. We note that to run these calculations, the CPTs did not need to be recalculated; rather, the "information updating" capabilities of BNs were leveraged.

For the same cases that were compared above and assuming the survival of the identified fault trees, we observe that, under the perfect GM correlation assumption, the site CD probability [P(Site CD | EES, /1-01-S-DMG, /2-01-S-DMG)] is being underestimated by a slightly greater amount (about 19%), suggesting the presence of single SSC MCSs may reduce the magnitude of underestimation associated with the perfect GM correlation assumption. Using the perfect GM correlation assumption overestimates the concurrent CD probability [P(Concurrent CD | EES, /1-01-S-DMG)] by about 13% to 19%. This overestimation is expected because the computed failure probability of a parallel system (akin to the concurrent CD probability) increases as the correlation increases. However, this sensitivity case shows how single failure MCSs may "quantitatively mask" this effect.

Next, we explore the impact of the fragility correlation assumptions by looking within each GM correlation group; i.e., comparing Cases 1 through 5 (No GM correlation), Cases 6 through 10 (Partial GM correlation), and Cases 10 through 15 (Perfect GM correlation). As expected, the

concurrent CD probability increases as the fragility correlation increases within each GM correlation group.

There is no clear overall trend in the P(Site CD | EES) results because of the complex interplay between the site-level risk metric (series system) and the within unit properties (redundant SSCs behaving as a parallel system). However, when breaking down the cases, we see some trends. For example, there is an increase in P(Site CD | EES) for the following set of cases in which the within unit SSC correlation increases and the between units SSC correlation remains the same: (a) 1 to 2, (b) 3 to 4, (c) 6 to 7, (d) 8 to 9, (e) 11 to 12, and (f) 13 to 14. In each set of cases, the higher within unit correlation causes an increase in the CD probability of *each unit*, which is composed of redundant SSCs. In turn, this leads to the observed increase in the P(Site CD | EES) results. Also, there is an expected decrease in P(Site CD | EES) in the following set of cases where the within unit SSC correlation stays the same and the between units SSC correlation increases: (a) 4 to 5, (b) 7 to 8, (c) 9 to 10, (d) 12 to 13, and (e) 14 to 15. There is a numerically negligible difference when moving from Case 2 to 3.

There is an increasing trend in P(Site CD | EES, /1-01-S-DMG, /2-01-S-DMG), which may appear counterintuitive to the knowledge that a series system's failure probability decreases as the correlation increases. This trend in P(Site CD | EES, /1-01-S-DMG, /2-01-S-DMG) is because, as the fragility correlation increases, the minimal cut sets with multiple SSCs (i.e., redundant SSCs) contribute to the increase in the CD probability of each unit. Also, in the cases where the within unit SSC correlation remains the same and the between units SSC correlation increases (i.e., Cases 2 to 3, 7 to 8, and 12 to 13), the individual unit CD probabilities increase

because there is a slight increase in the probability of the mid GM value bins (e.g., 1-BIN-3, 1-BIN-4, and 1-BIN-5).

Finally, we compare Case 8 (which reflects the approach outlined in this study), to the spectrum of cases that capture the state of practice, that is, Cases 11, 12, and 13. Comparing Cases 8 and 11, we observe that Case 11 underestimates all risk metrics, which can be attributed to the decrease in assumed correlation within unit and between units SSC fragility correlation, leading to a decrease in CD probability in each unit. Another contributing factor is the assumed perfect GM correlation under Case 11 neglecting the potential for higher GM values in Unit 2 compared to Unit 1. Comparing Cases 8 and 12, we observe that Case 12 also underestimates all risk metrics mainly due to neglecting the potential for higher GM values in Unit 2 than in Unit 1 (due to the perfect GM assumption) and, to a lesser degree, the decrease in the between units SSC fragility correlation. The comparison of risk metrics from Cases 8 and 13 follows the expected trends. That is, given that Case 13 has a higher degree of GM correlation than Case 8, Case 13 underestimates the site CD probability results and overestimates the concurrent CD probability results. Case 8 provides more realistic results than Cases 11, 12, and 13 because it appropriately considers the GM correlation between units (consistent with the spatial variability of ground motion [24]–[26], [28], [29]) and the fragility correlation within a unit and between units.

### 5.4 <u>Summary and Conclusions</u>

We proposed a BN-based approach to model a MUSPRA that integrates our previous works on the spatial variability of ground motion and modeling dependent seismic failures. The approach proposed herein is complementary to the typical event tree/fault tree approach used in the current SPRAs. The proposed BN approach to model a MUSPRA integrates all elements of a seismic

PRA into a coherent and probabilistically rigorous framework. Similar to our previous work [83], BNs offer the following advantages: (1) graphical representation that enables transparency and facilitates communication of modeling assumptions; (2) they are effective at modeling complex dependencies and can accommodate differing probability distribution assumptions; and (3) they facilitate information updating, which allows for the efficient calculation of joint and conditional probability distributions for all random variables in the BN. In this paper, we demonstrated advantages of the proposed approach that include: (1) the BNs shown in multiple figures provide a graphical representation of the MUSPRA and the underlying event trees and fault trees, (2) allowing computation of the NPP site-based risk metrics which include complex dependencies represented by the spatial variability of ground motion nodes and the  $\varepsilon_{ij}^*$  nodes, and (3) ability to calculate the NPP site-based risk metric conditioned on the success of most of the major structures at the hypothetical NPP site by using the updating aspects of the BN without recalculating the CPTs. We implemented the proposed approach for a realistic NPP subsystem.

- The computed site CDF ("series system logic" assuming either unit experiences core damage) decreases as the assumed correlation in ground motion between units is increased. As a result, the site CDF is underestimated by using the perfect ground motion correlation assumption. This supports our previous work's conclusion that the perfect ground motion correlation assumption may not necessarily be conservative [64].
- The concurrent CDF ("parallel system logic" assuming both units experience core damage) may not show appreciable differences between the partial and perfect ground motion correlation assumptions where single SSC failures dominate the core damage probability contribution from each unit to the concurrent CDF results. However, in cases

where the single SSC failures are "removed" (assumed to not fail), the concurrent CDF shows appreciable differences, with the perfect ground motion correlation overestimating concurrent CDF relative to the partial GM correlation assumption.

- When holding ground motion correlation assumptions constant, the site CDF shows a counterintuitive trend, that is, the site CDF increases as the fragility correlation increases.
- Case 8, which reflects the approach outlined in this study, provides more realistic results in all considered NPP site-based risk metrics compared to Cases 11, 12, and 13 (i.e., the spectrum of cases that capture the state of practice).

#### 5.5 <u>Disclaimer and Acknowledgements</u>

This paper was prepared, in part, by an employee of the U.S. Nuclear Regulatory Commission (NRC) on his own time apart from his regular duties. The NRC has neither approved nor disapproved its technical content. The views expressed in this paper are not necessarily those of the NRC.

The authors thank Dr. Taotao Zhou for sharing the MUSPRA model used in his work (i.e., [40]), which was instrumental in understanding the concept of seismic CCF and large fault trees in the context of a MUSPRA.

THIS PAGE IS INTENTIONALLY LEFT BLANK

# Chapter 6 Conclusion, Summary of Contributions, and Future Research

## 6.1 <u>Conclusion</u>

The conclusions specific to each research objective identified in Section 1.2 are presented in Sections 3.5, 4.5, and 5.4, respectively. These conclusions are summarized as follows:

- Conclusion for Objective 1: The method for modeling spatial variability of ground motion presented in this dissertation is a more realistic alternative to the conventional assumption that ground motions across an NPP site are perfectly correlated. For the example case, the probability distribution of the ground motion at the non-reference unit is more dispersed and slightly shifted to the right relative to the distribution of the ground motion at the reference unit, allowing for the possibility of higher levels of ground motion at the non-reference location. This observation should remain valid regardless of the model of spatial variability of ground motion that is used because any model should estimate that there is a non-negligible probability that the ground motion hazard at the non-reference unit may be different than that of the reference unit.
- Conclusion for Objective 2: The method for modeling dependent seismic performance of SSCs presented in this dissertation improves upon the Reed-McCann method (as implemented in NUREG/CR-7237 [15]) as judged by the comparison against the theoretical lower and upper bounds of the system fragility. Also, the method can be used in both single-unit and multi-unit seismic PRAs.
- Conclusion for Objective 3: The MUSPRA approach presented in this dissertation integrates all elements of a seismic PRA into a coherent and probabilistically rigorous

framework. Based on the results from the case study in Chapter 5, this dissertation provides the following insights:

- The computed site CDF ("series system logic" assuming either unit experiences core damage) decreases as the assumed correlation in ground motion between units is increased. As a result, the site CDF is underestimated by using the perfect ground motion correlation assumption. This supports the conclusion that using the perfect ground motion correlation assumption in a MUSPRA is not conservative when the site CDF is the risk metric of interest.
- The concurrent CDF ("parallel system logic" assuming both units experience core damage) may not show appreciable differences between the partial and perfect ground motion correlation assumptions where single SSC failures dominate the core damage probability contribution from each unit to the concurrent CDF results. However, in cases where the single SSC failures are "removed" (i.e., assumed to not fail), the concurrent CDF shows appreciable differences, with the perfect ground motion correlation overestimating concurrent CDF relative to the partial GM correlation assumption.
- When holding ground motion correlation assumptions constant, the site CDF shows a counterintuitive trend, that is, the site CDF increases as the fragility correlation increases. This is because the core damage probability of each unit (where each unit is akin to a parallel system) increases as the fragility correlation increases.
- The proposed approach (labeled Case 8 in Chapter 5), which accounts for the spatial variability of ground motion and assumes that the seismic performance of SSCs within a unit and between units is partially correlated, provides more realistic results

when considering multiple NPP site-based risk metrics compared to the spectrum of cases that capture the state of practice.

Overall, this dissertation presents a BN perspective on the elements of a MUSPRA by outlining a MUSPRA approach that explicitly accounts for the dependencies across NPP reactor units. BNs are used because they offer the following advantages: (1) graphical representation that enables transparency and facilitates communication of modeling assumptions; (2) they are effective at modeling complex dependencies and can accommodate differing probability distribution assumptions; and (3) they facilitate information updating, which allows for the efficient calculation of joint and conditional probability distributions for all random variables in the BN. Specifically, the MUSPRA approach considers the spatial variability of the ground motions (seismic hazard), dependent seismic performance of SSCs (fragility evaluation), and efficient BN modeling of systems (systems (plant response) analysis). Accounting for these dependencies in a systematic manner makes the MUSPRA more realistic and, therefore, should provide confidence in its results and risk insights.

### 6.2 <u>Summary of Contributions</u>

This dissertation makes the contributions that are listed below.

• This dissertation introduced a method with two model formulations for modeling the spatial variability of ground motion within a single unit and across multiple units. The method allows for a practical extension of an existing PSHA at a hard-rock site. These model formulations represent an improvement over the typical assumption that the ground motions across a NPP site are perfectly correlated while not requiring that the PSHA be repeated for the site.

- This dissertation developed a method for modeling dependent seismic performance of SSCs. The method can be used in single-unit seismic PRA and MUSPRA. This method is an improvement over the current "perfectly dependent or independent" approach for dependent seismic performance and provides system failure probability results that better comply with theoretical bounds than the Reed-McCann method. Also, because of its graphical nature, the proposed method for modeling the dependent seismic performance of SSC is easier to understand and communicate than the Reed-McCann method.
- This dissertation described a new approach to integrate the models of spatial variability
  of ground motion and dependent seismic performance of SSCs into a MUSPRA. The
  MUSPRA approach uses efficient BN modeling to complement the typically used event
  trees and fault trees by explicitly considering the dependencies that play an important role
  in the estimation of NPP site-based risk metrics. The proposed approach allows for the
  assessment of site-based risk metric considering more realistic models of dependencies.

#### 6.3 <u>Future Research</u>

This dissertation considered seismic PRA for NPPs located on hard-rock sites. However, the majority of NPP sites in the United States are soil sites. While the methods outlined for modeling dependent seismic performance of SSCs and estimating NPP site-based risk metrics remain valid for all site types, an approach is needed to consider the spatial variability of ground motions at soil sites. Expanding the proposed hard-rock spatial variability of ground motion model to practically model the spatial variability of ground motion at soil sites requires different model formulations, depending on the information available from the existing PSHA. While reserved as an area requiring further research, several potential modeling formulations are briefly outlined below.

The first proposed model formulation would be used if the soil seismic hazard is developed using Approach 4 in NUREG/CR-6728 [98], which estimates the soil seismic hazard using ground motion models that are a function of the earthquake magnitude, source-to-site distance, other explanatory variables, and the specific soil conditions at the site of interest. The model formulation would be similar to that of the model formulation for hard-rock sites discussed in Chapter 3. The Abrahamsom and Sykora  $\Delta SA_{1,2}$  model for soil sites [24] may be used if the magnitude deaggregation information is available. Alternatively, data from dense accelerograph arrays at soil sites may be used to develop a spatial variability of ground motion model similar to that from Kawakami and Mogi [25] (which was reflected by the node  $P'_{1,2}$  throughout the BNs presented in this dissertation).

Another proposed model formulation would be used if the soil seismic hazard is developed using Approach 3B in NUREG/CR-6728 [98], which estimates the soil seismic hazard as a function of the bedrock ground motion and a site-specific amplification factor. In this case, the method for soil sites can use the method for hard-rock sites as a first step because the seismic hazard at soil sites is typically estimated in three high-level steps [99]:

1. Estimate the seismic hazard at the top of bedrock, that is, the hard-rock seismic hazard curve.

- 2. From the site-response analysis, develop an amplification factor<sup>23</sup> that considers the different layers of the soil column (and their uncertain properties) between the top of bedrock and the foundation of the structure of interest. The amplification factor is represented as a probability distribution.
- 3. Convolve the hard-rock seismic hazard curve with the probability distribution of the amplificatory factor.

Then, based on the soil column underneath the reference unit, a horizontal correlation of shearwave velocities model [100], and models of spatial variability of soil properties [101]–[104], a correlated soil column underneath non-reference unit would be developed. Subsequently, the soil hazard at the non-reference unit would be estimated using the bedrock ground motion and the correlated soil column. The bedrock ground motion may or may not consider the spatial variability of ground motion depending on the available information.

Finally, the spatial variability of ground motion model for soil sites would be used in the same manner as described in Chapter 5 to estimate the NPP site-based risk metrics and develop risk insights.

This dissertation focused on the seismic fragility methodology that is currently used in the nuclear industry and expanded it to explicitly consider dependencies between SSCs using BNs. In the BN implementation of the dependent seismic performance of SSCs, it was assumed that

<sup>&</sup>lt;sup>23</sup> The amplification factor is defined as the ratio of the spectral acceleration at the soil surface to the spectral acceleration at the top of bedrock [99].

only two SSCs formed the SSC group. However, in some cases there were more than two SSCs in the SSC group. Research is needed to explore the expansion of SSC groups to more than two SSCs and its effect on the calculated risk metrics and risk insights.

This dissertation focused on seismic failures that may lead to adverse events (e.g., core damage). In the PRAs developed by the nuclear industry, the seismic failures of SSCs are added to the full power, internal events PRAs by making logic adjustments in the fault trees (e.g., using flag sets). Given the significant investment throughout several decades in the event tree/fault tree PRA approach, research and development of hybrid risk tools is needed to integrate the BN approach and other modeling approaches into existing PRA software tools, such as the Systems Analysis Programs for Hands-on Integrated Reliability Evaluations software [105]. These hybrid risk tools would be useful to assess the risk of NPPs from external hazards (e.g., seismic and external flooding). THIS PAGE IS INTENTIONALLY LEFT BLANK

# Bibliography

- [1] "Nuclear Power Proves its Vital Role as an Adaptable, Reliable Supplier of Electricity during COVID-19," Jun. 24, 2021. https://www.iaea.org/newscenter/news/nuclear-power-proves-its-vital-role-as-an-adaptable-reliable-supplier-of-electricity-during-covid-19 (accessed Jul. 16, 2021).
- U.S. Energy Information Administration, "Glossary Capacity Factor." https://www.eia.gov/tools/glossary/index.php?id=Capacity%20factor (accessed Jun. 15, 2021).
- [3] "IAEA Power Reactor Information System." https://pris.iaea.org/PRIS/home.aspx (accessed Jun. 06, 2021).
- [4] World Nuclear News, "Tianwan 6 enters commercial operation." https://www.worldnuclear-news.org/Articles/Tianwan-6-enters-commercial-operation (accessed Jun. 06, 2021).
- [5] U.S. Energy Information Administration, "Capacity Factors for Utility Scale Generators Primarily Using Fossil Fuels." https://www.eia.gov/electricity/monthly/epm\_table\_grapher.php?t=epmt\_6\_07\_a (accessed Jun. 15, 2021).
- [6] U.S. Energy Information Administration, "Capacity Factors for Utility Scale Generators Primarily Using Non-Fossil Fuels." https://www.eia.gov/electricity/monthly/epm\_table\_grapher.php?t=epmt\_6\_07\_b (accessed Jun. 15, 2021).
- U.S. Nuclear Regulatory Commission, "Probabilistic Risk Assessment (PRA)." https://www.nrc.gov/about-nrc/regulatory/risk-informed/pra.html (accessed Jun. 15, 2021).
- [8] S. Kaplan and B. J. Garrick, "On the Quantitative Definition of Risk," *Risk Analysis*, vol. 1, no. 1, pp. 11–27, Mar. 1981, doi: 10.1111/j.1539-6924.1981.tb01350.x.
- [9] C. Miller, A. Cubbage, D. Dorman, J. Grobe, G. Holahan, and N. Sanfilippo, "Recommendations for Enhancing Reactor Safety in the 21st Century: The Near-Term Task Force Review of Insights from the Fukushima Dai-ichi Accident," U.S. Nuclear Regulatory Commission, Jul. 2011. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML1118/ML111861807.pdf
- [10] M. L. Corradini *et al.*, "Fukushima Daiichi: ANS Committee Report, A Report by The American Nuclear Society Special Committee on Fukushima," American Nuclear Society, Jun. 2012. Accessed: Jun. 16, 2021. [Online]. Available: https://www.ans.org/pubs/reports/fukushima/report/
- [11] International Atomic Energy Agency, "The Fukushima Daiichi Accident, Report by the Director General," International Atomic Energy Agency, Vienna, Austria, 2015. Accessed: Jun. 16, 2021. [Online]. Available: https://www.iaea.org/publications/10962/the-fukushima-daiichi-accident

- [12] V. M. Andersen, L. K. Lee, P. H. Patel, and E. T. Burns, "Seismic Probabilistic Risk Assessment Implementation Guide," Electric Power Research Institute, Palo Alto, CA, EPRI 3002000709, Dec. 2013. Accessed: Jun. 16, 2021. [Online]. Available: https://www.epri.com/#/pages/product/00000003002000709/
- [13] Electric Power Research Institute, "Framework for Assessing Multi-Unit Risk to Support Risk-Informed Decision-Making: Phase 1 & 2: General Framework and Application-Specific Refinements," Electric Power Research Institute, EPRI 3002018229, Sep. 2020. Accessed: Jun. 16, 2021. [Online]. Available: https://www.epri.com/research/products/00000003002018229
- [14] A. Der Kiureghian, "A Coherency Model for Spatially Varying Ground Motions," *Earthquake Engng. Struct. Dyn.*, vol. 25, no. 1, pp. 99–111, Jan. 1996, doi: 10.1002/(SICI)1096-9845(199601)25:1<99::AID-EQE540>3.0.CO;2-C.
- [15] R. J. Budnitz, G. S. Hardy, D. L. Moore, and M. K. Ravindra, "Correlation of Seismic Performance in Similar SSCs (Structures, Systems, and Components)," U.S. Nuclear Regulatory Commission, NUREG/CR-7237, Dec. 2017. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML1734/ML17348A155.pdf
- [16] M. Modarres, M. Kaminskiy, and V. Krivtsov, *Reliability Engineering and Risk Analysis: A Practical Guide*, Third edition. Boca Raton: CRC Press, Taylor & Francis Group, 2017.
- [17] U. B. Kjærulff and A. L. Madsen, *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*. Springer, 2008. [Online]. Available: https://doi.org/10.1007/978-0-387-74101-7
- [18] M. T. Bensi, A. Der Kiureghian, and D. Straub, "A Bayesian Network Methodology for Infrastructure Seismic Risk Assessment and Decision Support," Pacific Earthquake Engineering Research Center, PEER 2011/02, Mar. 2011. Accessed: Jun. 16, 2021. [Online]. Available: https://peer.berkeley.edu/sites/default/files/webpeer-2011-02michelle\_t.\_bensi\_armen\_der\_kiureghian\_and\_daniel\_straub.pdf
- [19] D. Kelly and C. Smith, *Bayesian Inference for Probabilistic Risk Assessment: A Practitioner's Guidebook*. London: Springer, 2011.
- [20] S. L. Lauritzen and D. J. Spiegelhalter, "Local Computations with Probabilities on Graphical Structures and their Application to Expert Systems," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 50, no. 2, pp. 157–224, 1988.
- [21] F. V. Jensen, S. L. Lauritzen, and K. G. Olesen, "Bayesian updating in causal probabilistic networks by local computations," *Computational Statistics Quarterly*, vol. 4, pp. 262–282, 1990.
- [22] *GeNIe* (2.2 Academic). BayesFusion, LLC (https://www.bayesfusion.com), 2018. Accessed: Jun. 16, 2021. [Online]. Available: https://bayesfusion.com
- [23] S. Schroer and M. Modarres, "An event classification schema for evaluating site risk in a multi-unit nuclear power plant probabilistic risk assessment," *Reliability Engineering & System Safety*, vol. 117, no. Supplement C, pp. 40–51, Sep. 2013, doi: 10.1016/j.ress.2013.03.005.

- [24] N. Abrahamson and D. Sykora, "Variation of Ground Motions Across Individual Sites," in *Proc. of the Fourth U.S. Department of Energy Natural Phenomena Hazards Mitigation Conference, LLNL CONF-9310102*, Atlanta, GA, USA, Oct. 1993, pp. 192– 198. Accessed: Jun. 16, 2021. [Online]. Available: http://www.iaea.org/inis/collection/NCLCollectionStore/\_Public/25/052/25052560.pdf?r =1
- [25] H. Kawakami and H. Mogi, "Analyzing Spatial Intraevent Variability of Peak Ground Accelerations as a Function of Separation Distance," *Bulletin of the Seismological Society of America*, vol. 93, no. 3, pp. 1079–1090, Jun. 2003, doi: 10.1785/0120020026.
- [26] A. Zerva, Spatial Variation of Seismic Ground Motions: Modeling and Engineering Applications. CRC Press, 2009.
- [27] K. Goda and G. M. Atkinson, "Intraevent Spatial Correlation of Ground-Motion Parameters Using SK-net Data," *Bulletin of the Seismological Society of America*, vol. 100, no. 6, pp. 3055–3067, Dec. 2010, doi: 10.1785/0120100031.
- [28] N. Abrahamson, "Program on Technology Innovation: Effects of Spatial Incoherence on Seismic Ground Motions," Electric Power Research Institute, EPRI 1015110, Dec. 2007. Accessed: Jun. 16, 2021. [Online]. Available: https://www.epri.com/#/pages/product/0000000001015110/
- [29] E. Vanmarcke, "Local spatial variation of earthquake ground motion," in *Proceedings of the Tenth World Conference on Earthquake Engineering*, Madrid, Spain, Jul. 1994, pp. 6615–6621. Accessed: Jun. 16, 2021. [Online]. Available: http://www.iitk.ac.in/nicee/wcee/article/10\_vol11\_6615.pdf
- [30] K. Goda and H. P. Hong, "Spatial Correlation of Peak Ground Motions and Response Spectra Spatial Correlation of Peak Ground Motions and Response Spectra," *Bulletin of the Seismological Society of America*, vol. 98, no. 1, pp. 354–365, Feb. 2008, doi: 10.1785/0120070078.
- [31] N. Jayaram and J. W. Baker, "Correlation model for spatially distributed ground-motion intensities," *Earthquake Engng. Struct. Dyn.*, vol. 38, no. 15, pp. 1687–1708, Dec. 2009, doi: 10.1002/eqe.922.
- J. DeJesus Segarra, M. Bensi, and M. Modarres, "Framework for Modeling Ground Motion Variability at a Nuclear Power Plant Site for Use in a Seismic Multi-Unit Probabilistic Risk Assessment," Los Angeles, CA, Sep. 2018. Accessed: Jun. 16, 2021.
   [Online]. Available: https://www.iapsam.org/psam14/proceedings/paper/paper\_212\_1.pdf
- [33] J. W. Shea, "Tennessee Valley Authority's Seismic Hazard and Screening Report (CEUS Sites), Response to NRC Request for Information Pursuant to 10 CFR 50.54(f) Regarding Recommendation 2.1 of the Near-Term Task Force Review of Insights from the Fukushima Dai-ich Accident (ML14098A478)," Mar. 31, 2014. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML1409/ML14098A478.pdf
- [34] J. A. Ventosa, "Entergy Seismic Hazard and Screening Report (CEUS Sites), Response to NRC Request for Information Pursuant to 10 CFR 50.54(f) Regarding Recommendation 2.1 of the Near-Term Task Force Review of Insights from the Fukushima Dai-ich

Accident, Indian Point Unit No. 2 and No. 3 (ML14099A110 and ML14099A111)," Mar. 31, 2014.

- [35] J. W. Baker, "An Introduction to Probabilistic Seismic Hazard Analysis (PSHA)." 2008.
   Accessed: Jun. 16, 2021. [Online]. Available: https://web.stanford.edu/~bakerjw/Publications/Baker\_(2008)\_Intro\_to\_PSHA\_v1\_3.pdf
- [36] N. M. Kuehn and N. A. Abrahamson, "Spatial Correlations of Ground Motion for Non-Ergodic Seismic Hazard Analysis," *Earthquake Engineering & Structural Dynamics*, vol. 49, no. 1, pp. 4–23, Jan. 2020, doi: 10.1002/eqe.3221.
- [37] J. J. Bommer and N. A. Abrahamson, "Why Do Modern Probabilistic Seismic-Hazard Analyses Often Lead to Increased Hazard Estimates?," *Bulletin of the Seismological Society of America*, vol. 96, no. 6, pp. 1967–1977, Dec. 2006, doi: 10.1785/0120060043.
- [38] R. J. Budnitz *et al.*, "Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts," U.S. Nuclear Regulatory Commission, Washington, D.C., NUREG/CR-6372, Apr. 1997. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/reading-rm/doccollections/nuregs/contract/cr6372/index.html
- [39] International Atomic Energy Agency, "Probabilistic Safety Assessment for Seismic Events," International Atomic Energy Agency, Vienna, Austria, IAEA-TECDOC-724, Oct. 1993. Accessed: Jun. 16, 2021. [Online]. Available: https://www.iaea.org/publications/986/probabilistic-safety-assessment-for-seismicevents
- [40] T. Zhou, M. Modarres, and E. López Droguett, "An improved multi-unit nuclear plant seismic probabilistic risk assessment approach," *Reliability Engineering & System Safety*, vol. 171, no. Supplement C, pp. 34–47, Mar. 2018, doi: 10.1016/j.ress.2017.11.015.
- [41] R. K. McGuire, "Probabilistic seismic hazard analysis and design earthquakes: Closing the loop," *Bulletin of the Seismological Society of America*, vol. 85, no. 5, pp. 1275– 1284, Oct. 1995.
- [42] P. Bazzurro and C. A. Cornell, "Disaggregation of Seismic Hazard," *Bulletin of the Seismological Society of America*, vol. 89, no. 2, pp. 501–520, Apr. 1999.
- [43] J. F. Schneider, J. C. Stepp, and N. A. Abrahamson, "The Spatial Variation of Earthquake Ground Motion and Effects of Local Site Conditions," in *Proc. of the Tenth World Conference on Earthquake Engineering*, Madrid, Spain, 1992, pp. 967–972. Accessed: Jun. 16, 2021. [Online]. Available: http://www.iitk.ac.in/nicee/wcee/article/10\_vol2\_967.pdf
- [44] T. D. Ancheta, J. P. Stewart, and N. A. Abrahamson, "Engineering Characterization of Earthquake Ground Motion Coherency and Amplitude Variability," Aug. 2011. Accessed: Jun. 16, 2021. [Online]. Available: https://escholarship.org/uc/item/3mr1w33t
- [45] T. Lin and J. Baker, "Probabilistic Seismic Hazard Deaggregation of Ground Motion Prediction Models," presented at the 5th International Conference on Earthquake Geotechnical Engineering, Santiago, Chile, Jan. 2011. Accessed: Jun. 16, 2021. [Online].

Available:

https://web.stanford.edu/~bakerjw/Publications/Lin\_Baker\_(2011)\_GMPM\_deagg,\_5ICE GE.pdf

- [46] J. DeJesus Segarra, M. Bensi, and M. Modarres, "Incorporation of Spatial Variability of Ground Motions in a Seismic Multi-Unit Probabilistic Risk Assessment," in *Proceedings* of the 2019 International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2019), Charleston, SC, May 2019, pp. 292–299.
- [47] S. L. Kramer, *Geotechnical Earthquake Engineering*. Pearson, 1996.
- [48] K. J. Coppersmith *et al.*, "Central and Eastern United States Seismic Source Characterization for Nuclear Facilities," U.S. Nuclear Regulatory Commission, U.S. Department of Energy, and Electric Power Research Institute, NUREG-2115, DOE/NE-0140, EPRI 1021097, Jan. 2012. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr2115/index.html
- [49] L. A. Salomone *et al.*, "EPRI (2004, 2006) Ground-Motion Model (GMM) Review Project," Electric Power Research Institute, EPRI 3002000717, Jun. 2013. Accessed: Jun. 16, 2021. [Online]. Available: https://www.epri.com/research/products/00000003002000717
- [50] Tennessee Valley Authority, "Watts Bar Nuclear Plant, Units 1 and 2 Seismic Probabilistic Risk Assessment in Response to 50.54(f) Letter with Regard to NTTF 2.1 Seismic - Summary Report (ML17181A485)," CNL-17-071 (ML17181A485), Jun. 2017. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML1718/ML17181A485.pdf
- [51] M. Bensi, A. Der Kiureghian, and D. Straub, "Bayesian network modeling of correlated random variables drawn from a Gaussian random field," *Structural Safety*, vol. 33, no. 6, pp. 317–332, Sep. 2011, doi: 10.1016/j.strusafe.2011.05.001.
- [52] S. Ferson *et al.*, "Dependence in Probabilistic Modeling, Dempster-Shafer Theory, and Probability Bounds Analysis," Sandia National Laboratories, SAND2004-3072, Oct.
   2004. Accessed: Jun. 16, 2021. [Online]. Available: https://www.osti.gov/biblio/1427286
- [53] M. P. Bohn and J. A. Lambright, "Procedures for the External Event Core Damage Frequency Analyses for NUREG-1150," U.S. Nuclear Regulatory Commission, NUREG/CR-4840, Nov. 1990. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML0634/ML063460465.pdf
- [54] P. D. Smith *et al.*, "Seismic Safety Margins Research Program Phase I Final Report— Overview," U.S. Nuclear Regulatory Commission, NUREG/CR-2015, Vol. 1, Apr. 1981. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML1309/ML13093A113.pdf
- [55] G. E. Cummings, "Summary Report on the Seismic Safety Margins Research Program," U.S. Nuclear Regulatory Commission, NUREG/CR-4431, Jan. 1986.
- [56] J. W. Reed, M. W. McCann, J. Iihara, and H. Hadidi-Tamjed, "Analytical Techniques for Performing Probabilistic Seismic Risk Assessment of Nuclear Power Plants," in

Proceedings of the 4th International Conference on Structural Safety and Reliability (ICOSSAR '85), Kobe, Japan, May 1985, vol. III, pp. 253–263.

- [57] T. Zhou, M. Modarres, and E. López Droguett, "Issues in Dependency Modeling in Multi-Unit Seismic PRA," in *Proceedings of the 2017 International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2017)*, Pittsburgh, PA, Sep. 2017, pp. 668–675.
- [58] J. W. Baker, "Introducing correlation among fragility functions for multiple components," presented at the The 14th World Conference on Earthquake Engineering, Beijing, China, Oct. 2008. Accessed: Jun. 16, 2021. [Online]. Available: https://web.stanford.edu/~bakerjw/Publications/Baker\_(2008)\_Correlated\_fragilities,\_W CEE.pdf
- [59] O. Ditlevsen, Uncertainty modeling with applications to multidimensional civil engineering systems. New York ; London: McGraw-Hill International Book Co, 1981.
- [60] Y. Watanabe, T. Oikawa, and K. Muramatsu, "Development of the DQFM method to consider the effect of correlation of component failures in seismic PSA of nuclear power plant," *Reliability Engineering & System Safety*, vol. 79, no. 3, pp. 265–279, Mar. 2003, doi: 10.1016/S0951-8320(02)00053-4.
- [61] S. Kwag, J. Park, and I.-K. Choi, "Development of efficient complete-sampling-based seismic PSA method for nuclear power plant," *Reliability Engineering & System Safety*, vol. 197, p. 106824, May 2020, doi: 10.1016/j.ress.2020.106824.
- [62] Electric Power Research Institute, "Seismic Fragility and Seismic Margin Guidance for Seismic Probabilistic Risk Assessment," Electric Power Research Institute, EPRI 3002012994, Sep. 2018. Accessed: Jun. 16, 2021. [Online]. Available: https://www.epri.com/research/products/00000003002012994
- [63] K. Ebisawa, K. Abe, K. Muramatsu, M. Itoh, K. Kohno, and T. Tanaka, "Evaluation of response factors for seismic probabilistic safety assessment of nuclear power plants," *Nuclear Engineering and Design*, vol. 147, no. 2, pp. 197–210, Mar. 1994, doi: 10.1016/0029-5493(94)90206-2.
- [64] J. DeJesus Segarra, M. Bensi, T. Weaver, and M. Modarres, "Extension of probabilistic seismic hazard analysis to account for the spatial variability of ground motions at a multiunit nuclear power plant hard-rock site," *Structural Safety*, vol. 85, p. 101958, Jul. 2020, doi: 10.1016/j.strusafe.2020.101958.
- [65] International Atomic Energy Agency, "Technical approach to probabilistic safety assessment for multiple reactor units," International Atomic Energy Agency, Vienna, Austria, IAEA Safety Report Series No. 96, May 2019. Accessed: Jun. 16, 2021. [Online]. Available: https://www.iaea.org/publications/12228/technical-approach-toprobabilistic-safety-assessment-for-multiple-reactor-units
- [66] R. P. Kennedy and M. K. Ravindra, "Seismic Fragilities for Nuclear Power Plant Risk Studies," *Nuclear Engineering and Design*, vol. 79, no. 1, pp. 47–68, May 1984, doi: 10.1016/0029-5493(84)90188-2.

- [67] S. Kaplan, V. M. Bier, and D. C. Bley, "A note on families of fragility curves—is the composite curve equivalent to the mean curve?," *Reliability Engineering & System Safety*, vol. 43, no. 3, pp. 257–261, Jan. 1994, doi: 10.1016/0951-8320(94)90029-9.
- [68] G. Hardy *et al.*, "Seismic Evaluation Guidance: Screening, Prioritization and Implementation Details (SPID) for the Resolution of the Fukushima Near-Term Task Force Recommendation 2.1: Seismic," Electric Power Research Institute, Palo Alto, CA, EPRI 1025287, Feb. 2013. Accessed: Jun. 16, 2021. [Online]. Available: https://www.epri.com/research/products/00000000001025287
- [69] J. W. Reed *et al.*, "A Methodology for Assessment of Nuclear Power Plant Seismic Margin," Electric Power Research Institute, EPRI NP-6041-SL, Rev. 1, Aug. 1991. Accessed: Jun. 16, 2021. [Online]. Available: https://www.epri.com/research/products/NP-6041-SLR1
- [70] R. J. Budnitz *et al.*, "A Methodology for Analyzing Precursors to Earthquake-Initiated and Fire-Initiated Accident Sequences," U.S. Nuclear Regulatory Commission, NUREG/CR-6544, Apr. 1998. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML0716/ML071650470.pdf
- J. W. Baker, "Introduction to Probabilistic Seismic Hazard Analysis." 2013. Accessed: Jun. 16, 2021. [Online]. Available: http://www.stanford.edu/~bakerjw/Publications/Baker\_(2013)\_Intro\_to\_PSHA\_v2.pdf
- U.S. Nuclear Regulatory Commission, "A Performance-Based Approach to Define the Site-Specific Earthquake Ground Motion," U.S. Nuclear Regulatory Commission, Regulatory Guide 1.208 (ML070310619), Mar. 2007. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML0703/ML070310619.pdf
- [73] M. Bensi, A. D. Kiureghian, and D. Straub, "Efficient Bayesian network modeling of systems," *Reliability Engineering and System Safety*, vol. 112, no. Supplement C, pp. 200–213, Apr. 2013, doi: 10.1016/j.ress.2012.11.017.
- [74] R. P. Kennedy, "Overview of Methods for Seismic PRA and Margin Analysis Including Recent Innovations," Aug. 1999. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML0429/ML042960158.pdf
- [75] A. Der Kiureghian, "First- and Second-Order Reliability Methods," in *Engineering Design Reliability Handbook*, 1st ed., E. Nikolaidis, D. M. Ghiocel, and S. Singhal, Eds. Boca Raton, FL: CRC Press, 2005, p. 24. [Online]. Available: https://doi.org/10.1201/9780203483930
- [76] *Mathcad (15.0).* Parametric Technology Corporation (PTC) (https://www.mathcad.com/en/), 2010. Accessed: Jun. 16, 2021. [Online]. Available: https://www.mathcad.com/en/
- [77] B. Gencturk et al., FERUM Finite Element Reliability Using Matlab. Department of Civil and Environmental Engineering, University of California Berkeley (http://projects.ce.berkeley.edu/ferum/). Accessed: Jun. 16, 2021. [Online]. Available: http://projects.ce.berkeley.edu/ferum/

- [78] *Matlab (R2018b)*. MathWorks (https://www.mathworks.com/), 2018. Accessed: Jun. 16, 2021. [Online]. Available: https://www.mathworks.com/
- [79] ASME and ANS, "Addenda to ASME/ANS RA-S–2008 Standard for Level 1/Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plant Applications," American Society of Mechanical Engineers and American Nuclear Society, ASME/ANS RA-Sa–2009, Feb. 2009.
- [80] P.-L. Liu, H.-Z. Lin, and A. Der Kiureghian, "CalREL User Manual," Department of Civil Engineering, University of California Berkeley, UCB/SEMM-89/18, Aug. 1989. Accessed: Jun. 16, 2021. [Online]. Available: https://bitbucket.org/sanjayg0/calrel/src/master/
- [81] Nuclear Energy Institute, "Diverse and Flexible Coping Strategies (FLEX) Implementation Guide," Nuclear Energy Institute, NEI 12-06 [Rev 5] (ML18120A300), Apr. 2018. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML1812/ML18120A300.pdf
- [82] M. Modarres, T. Zhou, and M. Massoud, "Advances in multi-unit nuclear power plant probabilistic risk assessment," *Reliability Engineering & System Safety*, vol. 157, no. Supplement C, pp. 87–100, Jan. 2017, doi: 10.1016/j.ress.2016.08.005.
- [83] J. DeJesus Segarra, M. Bensi, and M. Modarres, "A Bayesian Network Approach for Modeling Dependent Seismic Failures in a Nuclear Power Plant Probabilistic Risk Assessment," *Reliability Engineering & System Safety*, vol. 213, p. 107678, Sep. 2021, doi: 10.1016/j.ress.2021.107678.
- [84] K. Fleming, "Summary Report of the International Workshop on Multi-Unit Probabilistic Safety Assessment," Canadian Nuclear Safety Commission, Ottawa, Ontario, Canada, Edoc reference #4704298, Nov. 2014. Accessed: Jun. 16, 2021. [Online]. Available: http://crarisk.com/wp-content/uploads/2015/09/Summary-Report-of-International-Workshop-on-Multi-Unit-PSA-2014-Ottawa-Canada-Karl-Fleming.pdf
- [85] T. Hakata, "Seismic PSA method for multiple nuclear power plants in a site," *Reliability Engineering & System Safety*, vol. 92, no. 7, pp. 883–894, Jul. 2007, doi: 10.1016/j.ress.2006.04.022.
- [86] K. Ebisawa, T. Teragaki, S. Nomura, H. Abe, M. Shigemori, and M. Shimomoto, "Concept and methodology for evaluating core damage frequency considering failure correlation at multi units and sites and its application," *Nuclear Engineering and Design*, vol. 288, pp. 82–97, Jul. 2015, doi: 10.1016/j.nucengdes.2015.01.002.
- [87] M. P. Bohn *et al.*, "Application of the SSMRP Methodology to the Seismic Risk at the Zion Nuclear Power Plant," U.S. Nuclear Regulatory Commission, NUREG/CR-3428, Jan. 1984.
- [88] W. S. Jung, K. Hwang, and S. K. Park, "A new methodology for modeling explicit seismic common cause failures for seismic multi-unit probabilistic safety assessment," *Nuclear Engineering and Technology*, vol. 52, no. 10, pp. 2238–2249, Oct. 2020, doi: 10.1016/j.net.2020.03.023.

- [89] S. Kwag and A. Gupta, "Probabilistic risk assessment framework for structural systems under multiple hazards using Bayesian statistics," *Nuclear Engineering and Design*, vol. 315, pp. 20–34, Apr. 2017, doi: 10.1016/j.nucengdes.2017.02.009.
- [90] S. Kwag, A. Gupta, and N. Dinh, "Probabilistic risk assessment based model validation method using Bayesian network," *Reliability Engineering & System Safety*, vol. 169, pp. 380–393, Jan. 2018, doi: 10.1016/j.ress.2017.09.013.
- [91] Y. Heo and S. J. Lee, "Development of a multi-unit seismic conditional core damage probability model with uncertainty analysis," *Reliability Engineering & System Safety*, vol. 207, p. 107353, Mar. 2021, doi: 10.1016/j.ress.2020.107353.
- [92] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*. Cambridge, MA: MIT Press, 2009.
- [93] Korea Electric Power Corporation and Korea Hydro & Nuclear Power Co., LTD,
   "APR1400 Design Control Document Tier 2, Chapter 19, Probabilistic Risk Assessment and Severe Accident Evaluation," APR1400-K-X-FS-14002-NP, Revision 3 (ML18228A666), Aug. 2018. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML1822/ML18228A666.pdf
- [94] U.S. Nuclear Regulatory Commission, "Design Certification Rule for the APR1400 Design," Appendix F to Title 10 of the Code of Federal Regulations Part 52. Accessed: Apr. 05, 2021. [Online]. Available: https://www.nrc.gov/reading-rm/doccollections/cfr/part052/part052-appf.html
- [95] Korea Electric Power Corporation and Korea Hydro & Nuclear Power Co., LTD,
   "APR1400 Design Control Document Tier 2, Chapter 1, Introduction and General Description of the Plant," APR1400-K-X-FS-14002-NP, Revision 3 (ML18228A648), Aug. 2018. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML1822/ML18228A648.pdf
- [96] R. D. Campbell, M. K. Ravindra, and R. C. Murray, "Compilation of Fragility Information from Available Probabilistic Risk Assessments," Lawrence Livermore National Laboratory, UCID-20571, Rev. 1, Sep. 1988.
- [97] A. Bobbio, L. Portinale, M. Minichino, and E. Ciancamerla, "Improving the analysis of dependable systems by mapping fault trees into Bayesian networks," *Reliability Engineering & System Safety*, vol. 71, no. 3, pp. 249–260, Mar. 2001, doi: 10.1016/S0951-8320(00)00077-6.
- [98] R. K. McGuire, W. J. Silva, and C. J. Costantino, "Technical Basis for Revision of Regulatory Guidance on Design Ground Motions: Hazard- and Risk-Consistent Ground Motion Spectra Guidelines," U.S. Nuclear Regulatory Commission, NUREG/CR-6728, Oct. 2001. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML0131/ML013100232.html
- [99] P. Bazzurro and C. A. Cornell, "Nonlinear Soil-Site Effects in Probabilistic Seismic-Hazard Analysis," *Bulletin of the Seismological Society of America*, vol. 94, no. 6, pp. 2110–2123, Dec. 2004.

- [100] G. R. Toro, "Probabilistic Models of Site Velocity Profiles for Generic and Site-Specific Ground-Motion Amplification Studies' in 'Description and Validation of the Stochastic Ground Motion Model (PE&A 94PJ20)' — See page 1030 of the pdf file," Nov. 1995. Accessed: Jun. 16, 2021. [Online]. Available: https://www.nrc.gov/docs/ML0428/ML042800294.pdf
- [101] A. L. Jones, S. L. Kramer, and P. Arduino, "Estimation of Uncertainty in Geotechnical Properties for Performance-Based Earthquake Engineering," Pacific Earthquake Engineering Research Center, Berkeley, CA, USA, PEER 2002/16, Dec. 2002. Accessed: Jun. 16, 2021. [Online]. Available: https://peer.berkeley.edu/sites/default/files/0216\_a.\_jones\_s.\_kramer\_p.\_arduino.pdf
- [102] G. B. Baecher and J. T. Christian, *Reliability and Statistic in Geotechnical Engineering*. West Sussex, England: John Wiley & Sons, Inc., 2003.
- [103] G. A. Fenton and D. V. Griffiths, *Risk Assessment in Geotechnical Engineering*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2008.
- [104] H. Hu and Y. Huang, "PDEM-based stochastic seismic response analysis of sites with spatially variable soil properties," *Soil Dynamics and Earthquake Engineering*, vol. 125, p. 105736, Oct. 2019, doi: 10.1016/j.soildyn.2019.105736.
- [105] C. L. Smith and S. T. Wood, "Systems Analysis Programs for Hands-on Integrated Reliability Evaluations (SAPHIRE) Version 8," U.S. Nuclear Regulatory Commission, NUREG/CR-7039, Jun. 2011. Accessed: Jun. 18, 2021. [Online]. Available: https://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr7039/