INL/EXT-15-36741

Light Water Reactor Sustainability Program

Proof-of-Concept Demonstrations for Computation-Based Human Reliability Analysis: Modeling Operator Performance During Flooding Scenarios

September 2015

DOE Office of Nuclear Energy

INL/EXT-15-36741

Light Water Reactor Sustainability Program

Proof-of-Concept Demonstrations for Computation-Based Human Reliability Analysis: Modeling Operator Performance During Flooding Scenarios

Jeffrey C. Joe, Ronald L. Boring, Sarah Herberger, Tina Miyake, Diego Mandelli, and Curtis L. Smith

September 2015

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov/lwrs

Prepared for the
U.S. Department of Energy
Office of Nuclear Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

ABSTRACT

The United States (U.S.) Department of Energy (DOE) Light Water Reactor Sustainability (LWRS) program has the objective to develop technologies and other solutions that can improve the reliability, sustain the safety, and extend the life of the current reactors. To accomplish this program objective, there are multiple LWRS "pathways," or research and development (R&D) focus areas. One LWRS focus area is called the Risk-Informed Safety Margin Characterization (RISMC) Pathway. RISMC R&D primarily focuses on qualitatively and quantitatively characterizing risk specifically in terms of safety margin. The RISMC approach probabilistically combines risk-assessment with multi-physics models of plant physical processes (e.g., thermal-hydraulic models) that govern aging and degradation of systems, structures, and components (SSCs) in order to better optimize plant safety and performance.

Initial efforts to combine probabilistic and plant multi-physics models to quantify safety margins included simplified human reliability analysis (HRA). HRA researchers at Idaho National Laboratory have been collaborating with other risk analysts to develop a computational HRA approach, called the Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER), for inclusion into the RISMC framework. The HUNTER computational HRA method is a hybrid approach that leverages past work from cognitive psychology, human performance modeling, and HRA, but it is also a departure from existing static and even dynamic HRA methods. The basic premise of this research is to leverage applicable computational techniques, namely simulation and modeling, to develop and then, using the Risk Analysis in a Virtual Environment (RAVEN) as a controller, seamlessly integrate virtual operator models created in HUNTER with 1) the Multiphysics Object Oriented Simulation Environment (MOOSE) as a runtime environment that includes a full-scope plant model, and 2) the RISMC risk models already developed.

This report is divided into five chapters that cover the development of an external flooding event example and associated statistical modeling considerations. The first chapter is an overview of RISMC and the HUNTER computational HRA approach. Chapter 2 is a flooding event case study that significantly affected main control room and auxiliary operator performance. Chapter 3 addresses statistical modeling considerations for the development of HUNTER. And finally, Chapter 4 discusses the path forward for the next phase of RISMC research on computation-based HRA.

CONTENTS

FIGURES

TABLES

ACRONYMS

AC	Alternating Current
AI	Artificial Intelligence
ANOVA	Analysis of Variance
BWR	Boiling Water Reactor
CD	Complete Dependence
CROW	A MOOSE based application and C++ library containing probabilistic distributions and the control logic modules used by RAVEN
DC	Direct Current
DOE	Department of Energy
EDGs	Emergency Diesel Generators
EOPs	Emergency Operating Procedures
EPRI	Electric Power Research Institute
ESW	Emergency Service Water
FLEX	Diverse and Flexible Coping Strategies
gPWR	generic Pressurized Water Reactor
HEP	Human Error Probability
HFEs	Human Failure Events
HPCI	High-Pressure Coolant Injection
HRA	Human Reliability Analysis
HSSL	Human Systems Simulation Laboratory
HUNTER	Human Unimodel for Nuclear Technology to Enhance Reliability
IAEA	International Atomic Energy Agency
I & C	Instrumentation and Control
IC	Isolation Condenser
INL	Idaho National Laboratory
INPO	Institute of Nuclear Power Operations
LOSP	Loss of Off-Site Power
LWRS	Light Water Reactor Sustainability
MCR	Main Control Room
MD	Moderate Dependence
MOOSE	Multiphysics Object Oriented Simulation Environment
NPP	Nuclear Power Plants

NUREG	U.S. Nuclear Regulatory Commission Regulation
PPE	Personal Protective Equipment
PRA	Probabilistic Risk Assessment
PSA	Probabilistic Safety Assessment
PSF	Performance Shaping Factor
PWR	Pressurized Water Reactor
R&D	Research and Development
RAVEN	Risk Analysis in a Virtual Environment
RCIC	Reactor Core Isolation Cooling
RCS	Reactor Cooling System
RELAP	Reactor Excursion and Leak Analysis Program
RHR	Residual Heat Removal
RISMC	Risk-Informed Safety Margin Characterization
RPV	Reactor Pressure Vessel
SAMG	Severe Accident Mitigation Guideline
SBO	Station Blackout
SPAR-H	Standardized Plant Analysis Risk- Human Reliability Analysis
SRV	Safety Relief Valves
SSCs	Systems, Structures, and Components
SWS	Service Water System
THERP	Technique for Human Error Prediction
U.S.	United States
V	Volts
ZD	Zero Dependence

Proof-of-Concept Demonstrations for Computation-Based Human Reliability Analysis: Modeling Operator Performance during Flooding Scenarios

1. INTRODUCTION

The United States (U.S.) Department of Energy (DOE) Light Water Reactor Sustainability (LWRS) program has the objective to develop technologies and other solutions that can improve the reliability, sustain the safety, and extend the life of current reactors. To accomplish this objective, the LWRS program has the following program goals:

- Develop the fundamental scientific basis to understand, predict, and measure changes in materials and systems, structures, and components (SSCs) as they age in environments associated with continued long-term operations of the existing reactors

- Apply this fundamental knowledge to develop and demonstrate methods and technologies that support safe and economical long-term operation of existing reactors

- Research new technologies to address enhanced plant performance, economics, and safety

To accomplish these program goals, there are multiple LWRS "pathways," or research and development (R&D) focus areas. One LWRS focus area is called the Risk-Informed Safety Margin Characterization (RISMC) Pathway. Because safety analysis is an important element of sustainability and NPP decision making, a systematic approach to characterize safety margins is needed. The characterization of safety margins and risk is a vital input to the NPP owners and regulator in that it supports business and operational decision-making.

Characterization of risk is not new. Probabilistic risk assessment (PRA) or probabilistic safety assessment (PSA) have been a staple of NPP operations and regulatory oversight. However, the manner in which RISMC characterizes risk, specifically in terms of safety margin, is different from current practices. RISMC safety margin characterization approach dynamically combines PRA with multi-physics models of plant physical processes (e.g., thermal-hydraulic models) that govern aging and degradation in order to better optimize plant safety and performance. This approach is depicted in Figure 1.

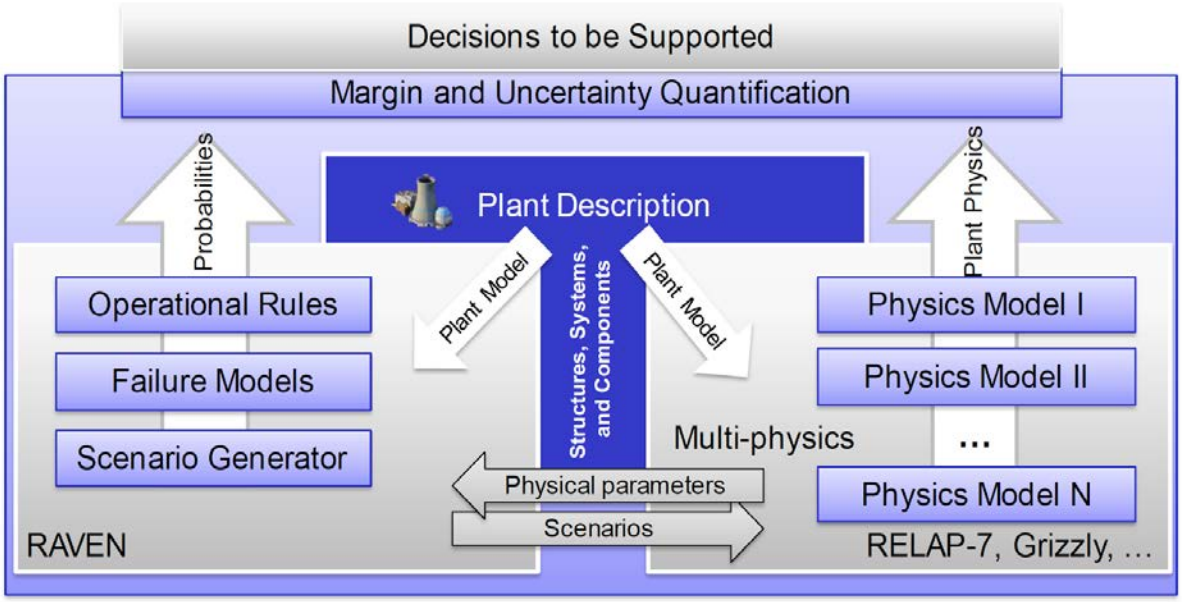

Figure 1. LWRS Approach to RISMC

The RISMC Toolkit uses the Multiphysics Object Oriented Simulation Environment (MOOSE) (Gaston, Hansen, & Newman, 2009) as the runtime environment, and combines 1) multi-physics codes that simulate the thermohydraulics of the plant, including the Reactor Excursion and Leak Analysis Program (RELAP)-7 code (David et al., 2012), and 2) the Risk Analysis in a Virtual Environment (RAVEN) (Alfonsi et al, 2013), which is the controller of the RELAP-7 simulation and generates multiple scenarios by stochastically changing the order and/or timing of events. Figure 2 and Figure 3 (from Alfonsi et al., 2013 & Rabiti et al., 2013) are alternate but basically equivalent representations of the RISMC framework, and they show how both control logic and classical system thermo hydraulic equations are inputs into the MOOSE runtime environment. The control logic equations control parameters such as pump speeds and valve positions, which along with the thermal-hydraulic equations, affect thermal-hydraulic variables such as pressure, temperature, and flow rates. These variables subsequently feed back to the control parameters via monitored variables (e.g., average pressure, delta-t).

The RISMC framework allows for any control logic equation, including ones representing human actions, to be inserted into the models and included in the simulation/analysis. This means that human reliability analysis (HRA) can be easily included into the RISMC framework such that human contributions to overall plant risk and/or dynamic characterization of the relationship between load and capacity can be obtained. It is simply a matter of including HRA in the control logic equations that are part of the overall plant equations and controlled parameters (See Section 1.1.1 for details). Thus, an advanced RISMC toolkit is created through this dynamic interchange of PRA and HRA models and multi-physics codes, which allows NPP owners and regulators to generate an enhanced representation of safety margins, and on how margins can be adjusted to improve operations and economics while still maintaining high levels of safety.

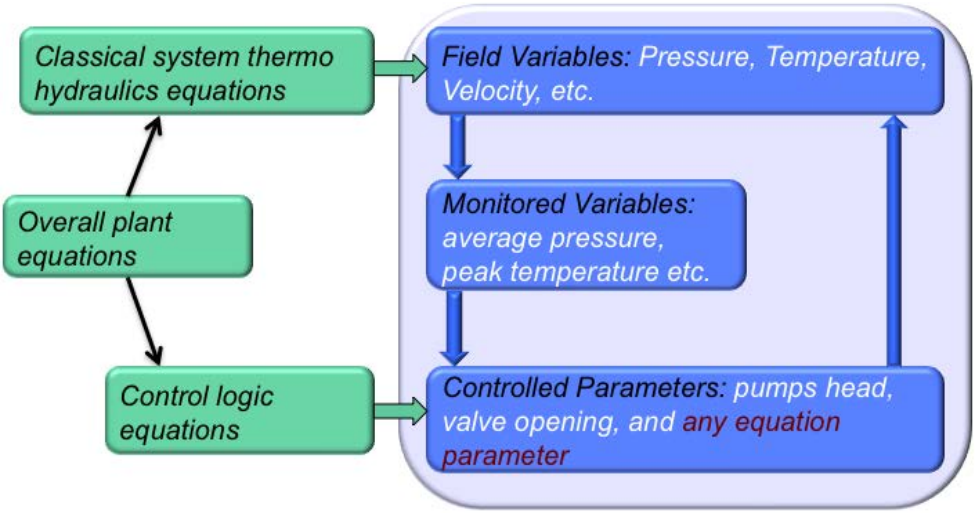

Figure 2. RISMC Control Logic (from Alfonsi et al., 2013)

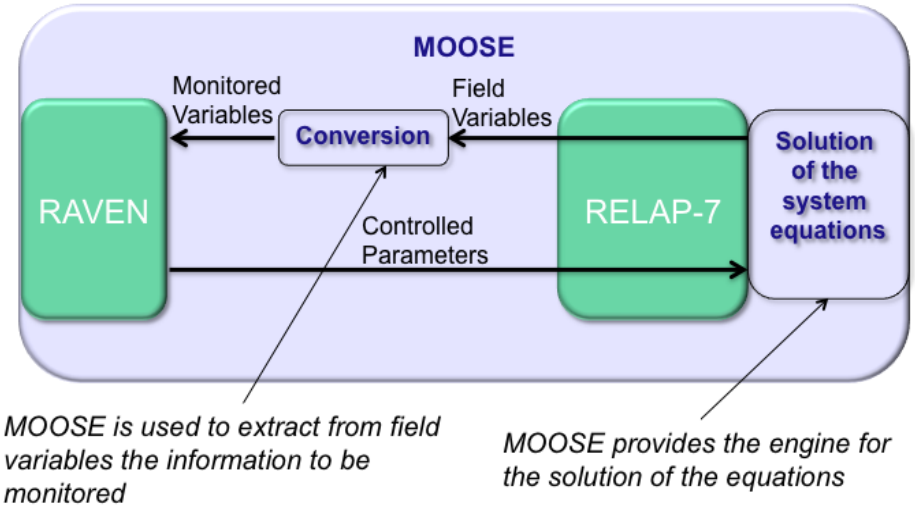

Figure 3. RISMC Control Logic (from Rabiti et al., 2013)

1.1 The Development of Computational HRA for RISMC

The initial efforts to combine probabilistic and plant multi-physics models to quantify safety margins and support business decisions also included HRA, but in a somewhat simplified manner. Figure 4 below depicts the way in which HRA was included into the RISMC framework in these prior studies. Specifically, the LWRS boiling water reactor (BWR) and pressurized water reactor (PWR) station blackout (SBO) demonstration case studies (Mandelli et al., 2013; Smith et al., 2014) were proof-of-concept demonstrations of integrating HRA into the RISMC framework. In these SBO studies, probability density functions (pdfs) based on two performance shaping factors (PSFs), stress and task

complexity, from the Standardized Plant Analysis Risk-HRA (SPAR-H) method (Gertman et al., 2005) were created and then used to further define the limit surface between the failure and success regions for SBO cases when 1) the reactor was at either 100% or 120% power, and 2) the recovery time of the emergency diesel generators (EDGs) varied as a function of a Weilbull distribution found in Eide et al. (2005).

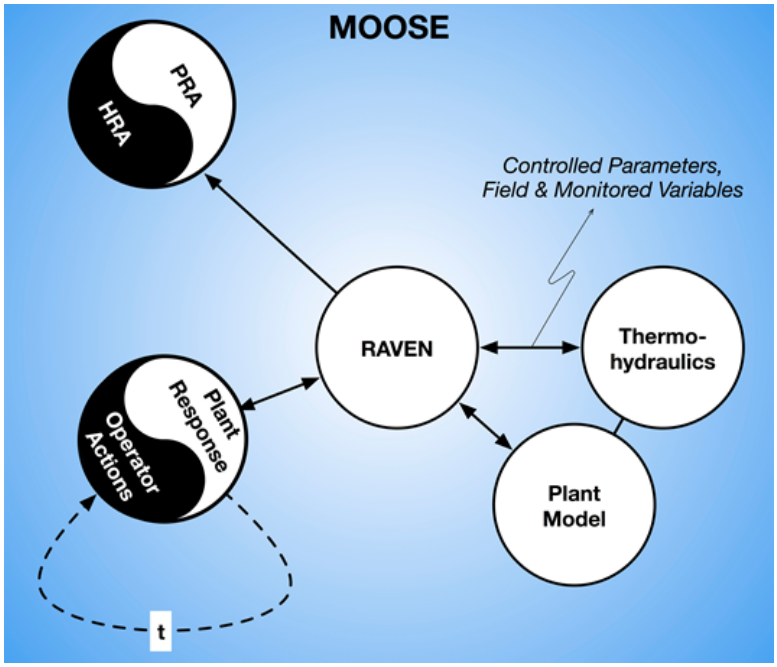

Figure 4. Previous HRA RISMC effort

Since the initial efforts by Mandelli et al. (2013) and Smith et al. (2014), HRA researchers at Idaho National Laboratory (INL) have been collaborating with other risk analysts to develop a computational HRA approach, called the *Human Unimodel for Nuclear Technology to Enhance Reliability* (HUNTER), for inclusion into the RISMC framework (Boring et al., 2014 & 2015). The basic premise of this research is to leverage applicable computational techniques, namely simulation and modeling, to develop and then, using RAVEN as a controller, seamlessly integrate virtual operator models (HUNTER) with 1) the dynamic computational MOOSE runtime environment that includes a full-scope plant model, and 2) the RISMC framework PRA models already in use. Like MOOSE, HUNTER is intended to be a flexible framework for incorporating operator performance models (e.g., cognitive models) into the larger RISMC framework. In this way, the HUNTER computational HRA approach is a hybrid approach that leverages past work from cognitive psychology, human performance modeling, and HRA, but it is also a departure from existing static and even dynamic HRA methods. This departure from existing HRA was also needed because HUNTER needs to factor additional complexities, such as spatial components to the problem, include mechanistic codes, and factor in the topology of the problem space.

A representation of the HUNTER approach, similar to what is shown in Figure 5, was included in the Boring et al. (2015) report. This representation has been slightly updated for this report to a) reference where RAVEN's control logic between the probabilistic and plant multi-physics models resides, b) include more explicitly PSFs as a data source for this computational HRA effort, and c) show where the HUNTER approach fits within the MOOSE runtime environment and how it aligns with the overarching RISMC goals of margin and uncertainty quantification.

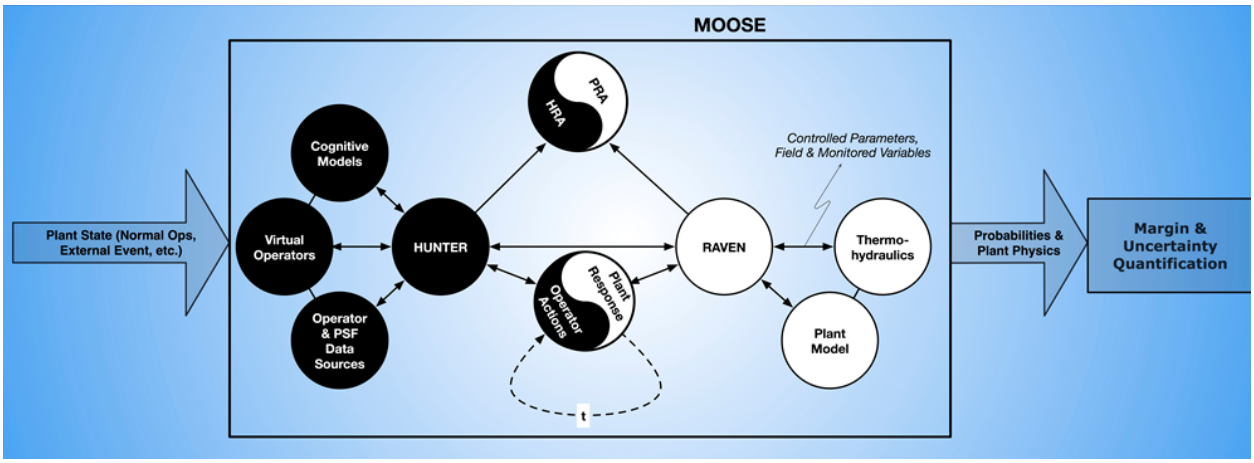

Figure 5. Computational HRA within the Full Context of RISMC Framework

Because HUNTER is event driven, external events such as flooding have a direct effect on the PSFs that govern the operator's behavior. As such, the functionality of HUNTER is that given a specific plant state (e.g., normal operations, external flooding event), a virtual operator (or crew) performs actions to maintain or change the plant state, whereby the thermal-hydraulic and other multi-physics models drive modeling of the plant state. Cognitive models and PSFs influence the virtual operator's actions.

Additionally, for HUNTER to work as planned in the RISMC framework, it needs to have the following characteristics:

- *Small number of PSFs:* A large number of PSFs would impose a significant computational burden on the RISMC framework, and in a large set of PSFs, many will likely have a negligible effect on the overall calculation of the limit surface.

- *Scalable:* HUNTER needs to support the development and testing of simplified virtual operator models for proof-of-concept testing through fully developed virtual models.

- *Not limited to time dynamics:* As mentioned previously, the computational HRA approach developed for RISMC needs to factor in additional complexities in addition to time.

- *Simplified cognitive model:* HUNTER is not reinventing artificial intelligence (AI). Rather, the goal of HUNTER is to find simplified ways of incorporating AI elements in risk modeling.

- *Sensitive to individual differences and crew performance:* Past work by Joe and Boring (2014a, 2014b) noted that existing HRA methods tend to overlook individual differences and emergent crew performance factors which can have a significant effect on the likelihood of operator errors.

- *Able to make use of empirical data:* The use of available data is important for the generation of pdfs of human performance, Bayesian updating (Groth, Smith, & Swiler, 2014), and could serve as a data source to compare and validate the quantification of human performance within the RISMC framework.

This HUNTER computational approach has a number of advantages over static and dynamic HRA in terms of how well it is suited to be integrated into the RISMC framework and address the key issues of:

- **Reducing epistemic uncertainty in modeling within the RISMC Framework.** Epistemic uncertainty is reduced when human contributions to risk are rigorously researched and factored into risk calculations. Plant reliability is a function of operator actions, so not including research on human performance only increases uncertainty that could otherwise be reduced through additional research.

- **Encompassing a greater range of plant dynamics during upsets.** PRA models are often just a predefined set of risks, treating any variability attributable to human actions (e.g., differences in time to complete proceduralized actions) as noise or error variance, when human actions, including errors of omission, can have large effects on overall risk by either significantly exacerbating or attenuating the severity of the failures of SSCs.

- **Enhancing dynamic response to changing conditions.** Human initiated control actions, during normal operations and post-initiator, affect how the plant responds to changing conditions. PRAs that only modeled the control actions of automatic safety systems, or modeled human actions without a robust understanding of the fundamental cognitive underpinnings of those actions, would only have a partial model of how the plant operates safely and efficiently or recovers from transients. Additionally, given the inclusion of HUNTER and RAVEN, the RISMC framework can be used to help plan emergency response actions by looking ahead to consequences of different actions, thereby proving the opportunity to have a real-time risk monitor available to NPP owners and operators.

Boring et al. (2015) further elaborates on the details of HUNTER.

1.1.1 Communication Between HUNTER and RAVEN

For the research being envisioned and performed, HUNTER would be a library of operator models that could be loaded and used within the RAVEN control logic interface. The control logic interface is an ideal environment to create the link between HUNTER and the plant dynamics. In more detail, plant (thermo-hydraulic) dynamics can be seen as a trajectory in the system phase space Rabiti et al. (2013):

$$\frac{\partial \boldsymbol{\theta}}{\partial t} = \boldsymbol{H}(\boldsymbol{\theta}, t) \tag{1}$$

where $\boldsymbol{\theta}$ represent the vector of the system state variables.

When control logic is included in the analysis, it is possible to split the vector $\boldsymbol{\theta}$ in two parts:

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{v} \end{pmatrix} \tag{2}$$

For the scope of this research, the decomposition is carried in such a way that $\boldsymbol{x}$ represents the set of unknowns solved by RELAP-7 while $\boldsymbol{v}$ represents the set of variables (parameters) directly controlled by the control system. The governing equation (2) can now be rewritten as follows:

$$\begin{cases} \dfrac{\partial \boldsymbol{x}}{\partial t} = \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{v}, t) \\ \dfrac{\partial \boldsymbol{v}}{\partial t} = \boldsymbol{V}(\boldsymbol{x}, \boldsymbol{v}, t) \end{cases} \tag{3}$$

As a consequence of this splitting, the components of the phase space in $\boldsymbol{x}$ are now all continuous while $\boldsymbol{v}$ contains both discrete and continuous variables. For example:

- Pressure and temperature in each point of the solution mesh belongs to $\boldsymbol{x}$
- On/off status of a pump (discrete), or the position of the control rods (continuous) belong to $\boldsymbol{v}$

A reasonable assumption is that the function $\mathbf{V}$, representing the control system is not dependent on the whole space spanned by $\mathbf{x}$, but just on a subspace. In fact, it is possible that the control system acts only on a set of signals coming from the plant and not on the whole plant status. Therefore, it is useful to introduce an appropriate subspace of $\mathbf{x}$, i.e., $\boldsymbol{c}$, from which the control logic can be fully derived. Thus, (3) is now re-cast as follows:

$$\begin{cases} \dfrac{\partial \boldsymbol{x}}{\partial t} = \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{v}, t) \\ \boldsymbol{c} = \boldsymbol{G}(\boldsymbol{x}, \mathrm{t}) \\ \dfrac{\partial \boldsymbol{v}}{\partial t} = \boldsymbol{V}(\boldsymbol{x}, \boldsymbol{v}, t) \end{cases} \tag{4}$$

where

- $\mathbf{x}$: set of plant status variables (e.g., temperature, pressure, and velocity on each point of the mesh)
- $\mathbf{F}$: function which describe the temporal evolution of the plant status variables
- $\boldsymbol{c}$: monitored variables; usually they are the result of an integral operator (projection) applied to the plant status variables (e.g., average temperature of a plant component, peak pressure in a pipe)
- $\mathbf{v}$: controlled variables. Variables affected by the control system (e.g., on/off pumps, control rod position, pump head, failure status of components, etc.)
- $\mathbf{V}$: Control logic law

The scope of HUNTER is to create a set of operator modules that expands the set of equations $\boldsymbol{G}$ and $\boldsymbol{V}$. This report does not go into detail on how these set of equations are solved, but an extensive description can be found in Rabiti et al. (2013) and Rabiti et al. (2012).

From a HUNTER point of view, the functions $\boldsymbol{G}$ and $\boldsymbol{V}$ can, for example, be:

- $\boldsymbol{G}$: Computations of PSFs as function of the operators working conditions, set of information that is available through the nuclear plant instrumentation and the human machine interface
- $\boldsymbol{V}$: Operators cognitive model solver, the set of Emergency Operating Procedures (EOPs), and in general any set of operator actions (both deterministic and stochastic)

Note that the set of functions $\boldsymbol{G}$ and $\boldsymbol{V}$ are the link between HUNTER and plant dynamics and PRA information shown in Figure 5. The set of HUNTER operator modules that will be developed will allow the user to customize them depending on the considered scenario.

The actual RAVEN control logic interface is built in `PYTHON` language. This language allows the creation of fairly complex control logic functions that can also, thanks to the set of libraries, be easily imported. In addition, RAVEN interfaces with a `C++` library developed at INL in parallel to RAVEN, called CROW. CROW is a MOOSE based application, which contains the set of probabilistic distributions and the control logic modules used by RAVEN.

The development of the HUNTER modules will be shared between both RAVEN and CROW. In particular, CROW will include the basic operator modules components that will be customized by the user in the RAVEN control logic module.

1.2 Scope of this Report

1.2.1 Phases of Work

As seen in Figure 6, the first year of this research effort to develop the HUNTER computational HRA approach involves three phases:

1. Review existing HRA and human performance modeling approaches to evaluate their applicability and usefulness to this research

2. Formulate human performance modeling (i.e., the creation of a virtual operator) and how it can be incorporated into the RISMC framework

3. Develop an external flooding event test case to explore how a model of a virtual operator would function with the multi-physics models

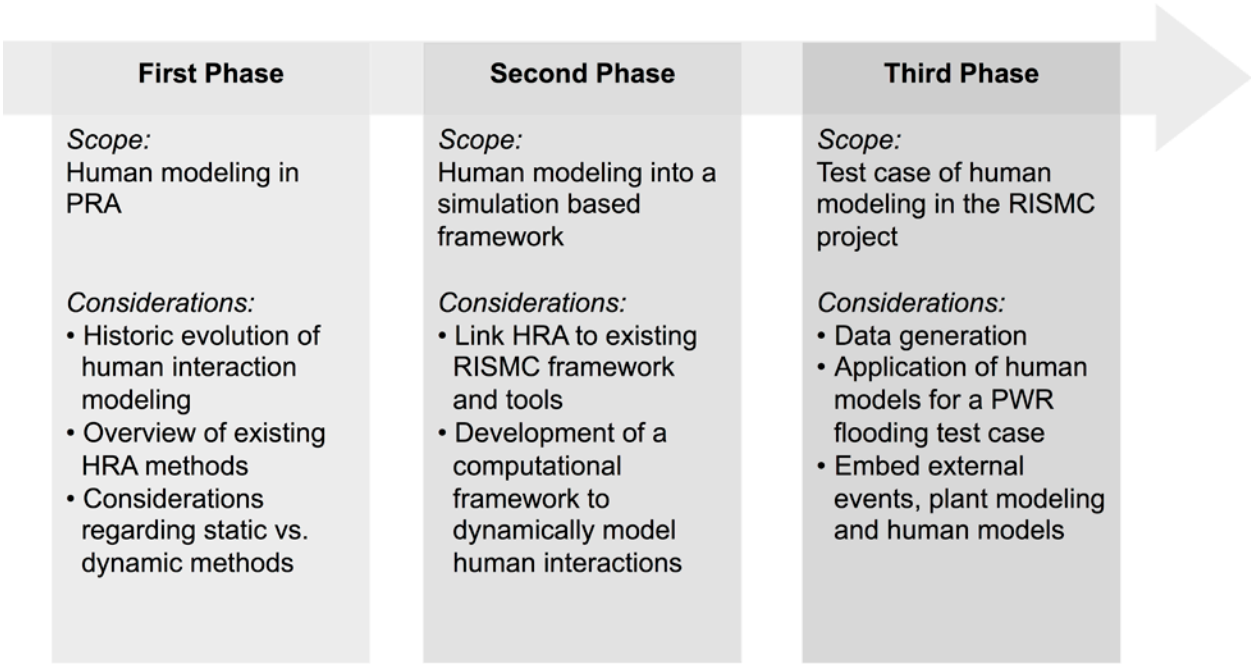

Figure 6. Phases and Scope of Work for HUNTER

Boring et al. (2014) addressed the first phase of this research. Boring et al. (2015) addressed the second phase of this research. This report discusses the work accomplished for the third phase and scope. This report is divided into five chapters that cover the development of an external flooding event test case and associated statistical modeling considerations. The chapters are:

- Chapter 1: The current chapter, which overviews RISMC and the HUNTER computational HRA approach.

- Chapter 2: This chapter presents a case study of a flooding event that significantly affected Main Control Room (MCR) and auxiliary operator performance.

- Chapter 3: This chapter covers statistical modeling considerations for the development of HUNTER.

- Chapter 4: The chapter discusses the path forward for the next phase of RISMC research on computation-based HRA.

2. CASE STUDY: FLOODING

2.1 Introduction

Flooding at an NPP is an important external event to study in the RISMC framework because it exercises multiple aspects of this risk-informed methodology that other internal events would not. The safety system and human response to a flooding event (e.g., tsunami) has spatial and temporal dimensions, which are factors that considered within the computational risk informed approach RISMC is taking (and are rarely considered in traditional PRA approaches). With the events that occurred at the Fukushima Daiichi NPP site, flooding is also an important and timely topic to investigate from a risk management perspective. Flooding also has some similarities to other external events, such as earthquakes and fire (NUREG-1921), and so by studying flooding, this research also generates useful lessons and insights that can be applied to other external events.

2.2 The Effect of the Great East Japan Earthquake and Tsunami on Fukushima Daiichi

On 11 March 2011, a magnitude 9 earthquake occurred off the east coast of Japan, approximately 112 miles from the Fukushima Daiichi Nuclear Power Plant (NPP), resulting in several tsunamis that inundated the Fukushima Daiichi site, and to a lesser degree 4 other nearby NPP sites (IAEA, 2011). Fukushima Daiichi NPP housed 6 Boiling Water Reactors (BWRs). Prior to the earthquake, units 1, 2, and 3 were operating normally with off-site power and were equipped with functioning emergency diesel generators (EDGs) and safety systems (IAEA, 2011; INPO, 2011; Kadota, 2014). Units 4, 5, and 6 were shut down for routine maintenance (INPO, 2011; Kadota, 2014). The earthquake, which lasted 3 minutes, exceeded the protection systems' set points, causing units 1, 2, and 3 to automatically scram by inserting control rods to stop the fission reaction (Kadota, 2014). Additionally, the earthquake caused loss of off-site power (LOSP) to all 6 units. In response to the LOSP, the EDGs started and provided power to the control rooms and the units to maintain reactor pressure, containment pressure, and reactor water level (INPO, 2011). Prior to the series of tsunamis, the safety systems operated as designed, and operators were able to safely maintain the reactors.

The turbine buildings and reactors were located 10 meters above sea-level (Kadota, 2014). The predicted maximum height for a tsunami was 6 meters, and, consequently, the protection for the units were built to that specification. Approximately 40 minutes after the earthquake, the first tsunami arrived at the site. The maximum wave height was approximately 14-15 meters leading to extensive flooding of the site. The tsunamis destroyed the EDGs or their electrical switchgear, and with no off-site AC power, the site entered into station blackout (SBO) conditions. With the loss of all AC power, all safety and non-safety systems driven by AC power became unavailable.

The control room instrumentation and control (I&C) systems and lighting requires DC power to operate, and unfortunately, the tsunami also flooded the DC distribution system and 125V DC batteries at units 1 and 2. The DC power distribution system and 125V DC batteries for unit 3 were available, but only for 30 hours because the battery's charger was flooded and AC power had not yet been restored.

2.3 Summary of Operator Responses

With the loss of AC power, operating units 1, 2, and 3 had lost power to the systems that provided decay heat removal. With the loss of DC power in units 1 and 2, there was no power to the lighting or I&C in the main control room. MCR operators for units 1 and 2, and eventually unit 3 were unable to check the condition of the reactors or the emergency cooling systems. For example, for unit 1, only the

isolation condenser (IC) system was initially available to provide decay heat removal via natural circulation, but because of lack of I&C indications, it was not clear how well it was working. Because of these difficulties, the operators were unable to safely maintain the reactors via normal operating procedures. Additionally, the tsunamis caused considerable damage to the buildings and the site in general. There was considerable debris and damage to infrastructure that made emergency first response and field operations by auxiliary operators more difficult than normal. These extremely poor working conditions were worse at night because of insufficient emergency lighting available, and then further complicated by the hydrogen explosions at units 1, 3, and 4, which created more debris and spread radioactive material. Field workers and auxiliary operators wore full mask respirators and personal protective equipment (PPE) as they performed their recovery actions because they were frequently working in high radiation fields.

Given the tsunami induced SBO conditions, the MCR operators needed to assess the condition of the reactors and the emergency cooling systems, and a way to find a way to cool the reactors (Kadota, 2014). Due to the condition of the control room, MCR operators were forced to assess the condition of the reactors by both physically entering the reactor building and scavenging batteries to power some of the indicators in the control room. Once the condition of the safety systems were ascertained, namely the IC system for unit 1, and the reactor core isolation cooling (RCIC) system and high-pressure coolant injection (HPCI) system for units 2 and 3, the site supervisor, Masao Yoshida, made a request for more fire engines in order to inject water into the reactors since only one fire engine on site remained operable after the tsunami. Supervisors and operators set up a pipeline connected to the Accident Management System, which had its own fire hydrant network. Before setting up the pipeline, supervisors and operators ascertained the pumps were operable. This incursion into the pump room was dangerous given the ongoing tsunami warning.

As previously mentioned, operators had to enter into the reactor building, as radiation levels were increasing, to assess the condition of the plant. Eventually, these operators also had to go into the reactor building to manually open valves to the emergency cooling system in order for water to be fed into the reactors via the pipeline. While their actions were deemed heroic (Kadota, 2014) and timely, the injection of the water into the reactors was delayed by 90 minutes, unfortunately, because the Prime Minister of Japan visited the Daiichi station during the nuclear emergency. Additionally, the pressure in the containment vessel for unit 1 had increased beyond the acceptable maximum level. Operators determined they were going to have to carry out a vent, which would release some radiation into the environment, but would help avoid a possible explosion of the containment vessel. In order to carry out the vent, operators would again need to go into the reactor building to open more valves. The first incursion into the reactor building was successful; however, radiation levels at the next set of valves exceeded the allowable limits, and the attempt was aborted to keep the operators safe. The day after the tsunami struck the NPP, the first water was injected into the reactor. Unfortunately, a hydrogen explosion occurred in unit 1 as well. INPO report (2011) indicates that hydrogen built up in the core and leaked into the containment vessel. The resulting explosion damaged the portable generator as well as hoses being set up to inject water into unit 1 and unit 2. Eventually, hydrogen explosions would also occur at unit 3 and unit 4.

2.4 HRA Characterization of MCR Operator Response to SBO and Flooding

Many aspects of operator response to seismic events have been characterized in HRA (Park et al., 2015; EPRI, 2012). Presley et al. (2013a & 2013b) have performed some additional analyses using the EPRI method. Both Park et al. (2015) and the preliminary HRA approach developed by the Electric Power Research Institute (EPRI; 2012) focused on the increased response times in operator actions as a result of a seismic event. Park et al. (2015) in particular attempted to quantify the appropriate time delays

for diagnosis and action for main control room and field operators under four different levels of seismic damage. Their quantification factored in the operator's stress level and degree to which reliable and valid information about the plant's state could be ascertained (e.g., level of information ambiguity given the availability of functioning I&C).

While the Park et al. (2015) analysis is insightful and instructive, its applicability and overlap with the focus of this research is somewhat limited. For example, the first 3 levels of seismic severity in the Park et al. (2015) analyses examine operator actions under less severe accident conditions than our situation. Additionally, the Park et al. (2015) analyses focuses primarily on the delay in time to complete required actions and primarily includes just one PSF – the operator's stress level – as a contributor to how effectively the operator can cogitate (e.g., detect, diagnose, decide, act, and monitor the plant's response) as a necessary aspect to their ability to perform their required duties successfully. The focus of this research effort is to analyze MRC and auxiliary operator actions under the severe accident conditions of external flooding. To do this, a review of reports on the Fukushima Daiichi accident was performed with a specific focus on the unique challenges operators faced and how methods such as SPAR-H (Gertman et al., 2005) and its performance shaping factors (PSFs) could be used to further characterize the challenges the operators faced. In general, in a flooding scenario where there is a loss of a) all AC power, b) DC batteries and/or the DC distribution system, and c) EDGs and/or their electrical switchgear, there will be significant delays in operators' detection, diagnosis, decision-making, and actions. These delays can be appropriately modeled for these conditions. The following list shows how an information processing based HRA method like SPAR-H would characterize the main human performance challenges for MCR operators:

1. Diagnosis becomes much harder for control room operators when indications from the I&C normally relied upon to obtain status on the plant are not readily available. This would be the case immediately after the flooding event, and subsequently during the SBO until I&C systems could be restored. As one operator stated, "At that point we had no idea how it [the SBO] could have happened. None of us had actually seen the water from the tsunami. All we knew was that it had happened. The generators had started up, had run properly, and then suddenly the power was gone. We could hardly believe it. It was something that just wasn't supposed to happen." (Kadota, 2014). Recovery of the I&C would require finding an alternative source of power besides the existing DC batteries (e.g., scavenging batteries from vehicles and wiring to power the control panels). This would lead to longer than normal diagnoses for cause(s) of the SBO and state of safety systems designed to remove decay heat.

2. Decision making becomes much harder when indications normally relied upon to obtain status on the plant are not available, and alternate communication channels are inadequate to provide the data needed to make an informed decision. This will generally force operators to make decisions in situations with greater than normal amounts of uncertainty. Additionally, depending on the significance of the decision, other decision makers who are not normally in the control room, but nevertheless in important leadership positions, may insert themselves into the situation to offer their opinion or dictate what to do. Their input can be valuable, but could also disrupt the normal leadership hierarchy and lead to the inappropriate stripping the operator's authority and responsibility to make key decisions. For example:

    a. For Fukushima Daiichi, deciding to vent steam from the reactor's primary containment vessel to the atmosphere. Automatic systems should vent automatically at 8 atmospheres (800 kilopascals) of pressure, but according to Kadota (2014), decision makers needed to have "irrefutable reasons" that the vent was the only way to resolve the situation.

b. The decision to use diesel fire engine pumps to provide reactor pressure vessel (RPV) makeup cooling, after loss of IC, RCIC, and HPCI was delayed by 90 minutes because the Prime Minister visited the Fukushima Daiichi site to question whether they were taking the appropriate actions to address the situation.

3. Actions become significantly more difficult to perform. Actuation of controls that normally occurs from main control room now requires deployment of operators into the field to perform the actions manually. These operators would be wearing full PPE that would restrict movement and interfere with normal verbal face-to-face communication. Operators would also be carrying extra equipment, including tools, radiation detectors, and flashlights which if carried in their hands means any action they have to perform with manual dexterity becomes more challenging.

    a. Actuation of valves that would normally occur with the push of a button in the control room now involves operators entering the irradiated reactor building in PPE with hand carried equipment. For unit 1, controls in MCR to align valves in the emergency cooling (core spray) system to allow external water source into containment were not available, requiring operator entry into the reactor building.

    b. Use of diesel fire engine pumps to provide RPV makeup cooling ideally requires availability of DC power and compressed nitrogen or air to actuate safety relief valves (SRV) to depressurize RPV. When these were not readily available, MCR and field operators tried to find alternative ways to open the SRVs.

As we integrate HUNTER into the flooding simulation research, we have the opportunity to create more realistic scenarios describing operator actions. For example, as shown in Figure 7, since water is leaking under the door and is spraying in the room, if an action was required to move from this room into the adjacent room, PSFs would likely be judged not favorable, resulting in an increased human error probability (HEP).

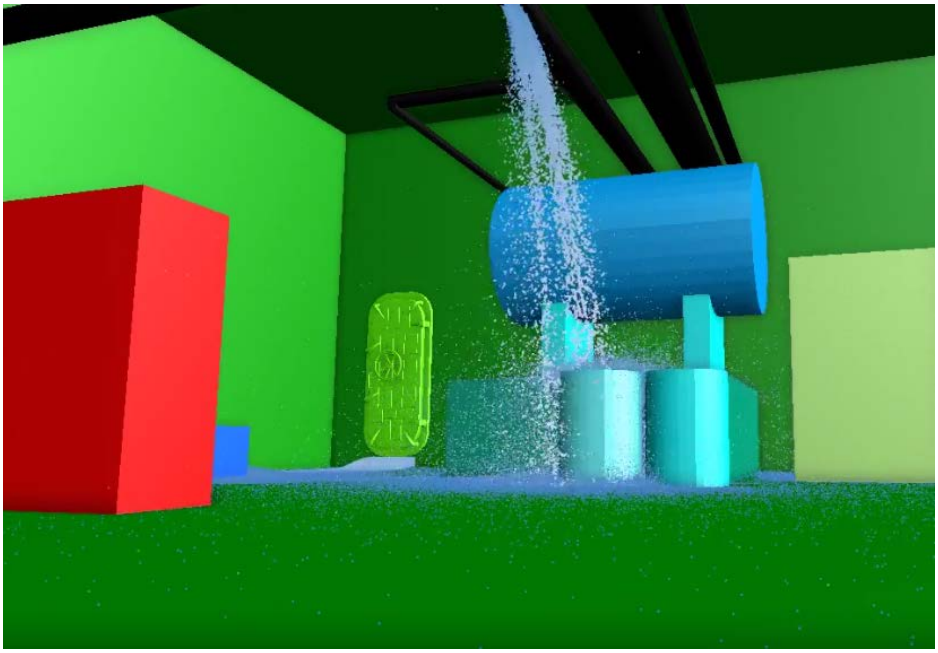

Figure 7. Illustration of the challenges posed by simulated flooding events in a facility

2.5 HRA Characterization of Auxiliary Operator Response to SBO and Flooding

The focus of the analysis in Section 2.4 was on MCR operators, but in an external flooding event, there will also be significant effects on auxiliary operators and their ability to perform their duties successfully. Using the SPAR-H method as a framework for our analysis of the auxiliary operator's response to a flooding induced SBO, it becomes apparent that the same generic tasks that the operators perform: namely a) make diagnoses, b) make decisions, and c) take action are similarly impaired by external flooding. However, the nature of the auxiliary operators' generic tasks, and the environment in which they need to perform those tasks, is drastically different than it is for MCR operators. Indeed, there were a few unique contextual factors that affected the performance of operators during their response to the SBO. This implies that, along with a need to understand how the task goals of auxiliary operators are often more sub-system specific (e.g., restore EDGs) versus the big picture goals MCR operators have (e.g., avoid core damage and/or the release of radiation by removing decay heat), there are a number of unique PSFs that need to be considered.

For example, two important sub-system specific tasks auxiliary operators at Fukushima Daiichi had to perform were:

A. Restore EDGs and/or their electrical switchgear

B. Operate diesel fire engine pumps to provide coolant to the RCS

The steps required to restore EDGs, including being able to correctly diagnose the cause(s) of their failure, what actions are required to recover them, and deciding how long that would take versus finding and installing replacement EDGs, would have been difficult to perform given the conditions of the site. To operate the diesel fire engine pumps, diagnosing whether they had sufficient flow was challenging due to the unavailability of I&C indications in the MCR. The following excerpt from Kadota (2012) highlights the specific challenges here:

> *At 19:54 water injection commenced. But there was still one thing they had to confirm. Had the safety valve opened, venting steam, thus lowering the pressure and allowing water to flow in, or was there another reason? As Yoshida pointed out, the instruments could not be trusted. "I immediately had the men beside the fire engine check that water was actually flowing into the reactor. I mean, it didn't matter how hard they pumped if the valve was still closed and water was not actually going in. Could they tell if water was actually flowing? I had them check the fire engine's flow gauge and feel the hoses by hand to make sure." Water was flowing and the hoses were throbbing. Even from the outside you could tell that the water was pulsing through the hoses. That was what Yoshida had wanted them to confirm. "I told them to look at the flow gauge and feel the flow of water in the hoses manually. After a while they reported back, first that the flow indicator was up, and then that they could also feel the flow in the hoses. You've no idea how relieved I was."*

In addition, the following factors and PSFs had a significant impact on the auxiliary operators' ability to perform these actions successfully.

- Degraded field conditions (e.g., damaged infrastructure and debris piles). Large groups of people were needed to remove this debris. Additionally, malfunctions with security infrastructure inhibited auxiliary operator performance in a number of unexpected ways.

Specifically, two auxiliary operators got stuck in "no man's land" (i.e., between the outer and inner security doors of a building) because the security system thought they were intruders. Physical security measures that were designed to thwart terrorist attacks (e.g., fences) were destroyed in the tsunami and created large debris piles that affected access to the site.

- The presence of radiation a) prohibiting free range and access to certain facilities (e.g., the reactor building), and b) requiring operators to wear PPE that restricted movement and interfered with verbal face-to-face communication. Additional issues with the PPE include:
    - The fact that the correct PPE that was needed was not always available (e.g., the correct radiation protection suits and radiation monitors).
    - Lead-lined PPE being heavy and increasing fatigue, and lead-lined gloves limiting manual dexterity.
    - Other equipment such as flashlights and satellite phones were in short supply. When emergency equipment was available, it often created the problem of requiring the operators to hold too many tools at once, thereby affecting their ability to perform actions (e.g., opening large valves) requiring two hands.
    - Over the course of the event, the supply of clean PPE ran low because operators needed to change their PPE every time they came from a radiation area and entered into a clean area. It was a constant challenge to keep clean areas free of radioactive contamination when others were required to repeatedly go back and forth between clean and irradiated areas.

2.5.1 Comparison with Other External Flooding Events and Latent Organizational Factors

INL researchers evaluated other external flooding events for relevance to the analyses and simulations of NPPs in the hours and days post-flooding that we will perform in the near future. Specifically, Hurricane Katrina and the effect of the flooding of the Missouri river on the Fort Calhoun NPP were studied. In general, however, it was determined that there were too many important differences between these flooding events and the Fukushima Daiichi tsunami flooding for them to provide additional insights into the human performance challenges of NPP operators. Specifically, with Hurricane Katrina, while there was extensive flooding and damage to infrastructure, there was no NPP site that was flooded and experienced an SBO. With Fort Calhoun, a NPP site was flooded, but it never experienced an SBO. It is also important to note that the single unit at Fort Calhoun was also in a refueling outage when Fukushima Daiichi units 1, 2, and 3, were in operation at full power. Another critical difference for Fort Calhoun was the relatively slow development of the flooding danger, caused by rising seasonal floodwater, versus a 14-meter high tsunami wave inundating the Fukushima Daiichi site. As a result of this difference in available time to react to the impending flood, personnel at Fort Calhoun were able to put in effective countermeasures to prevent any significant damage to important SSCs at the site. They deployed numerous quickly erectable temporary dams and berms to protect key areas of the site (e.g., power block, switchyard, transformers), brought in extra fuel for emergency equipment on-site, and additional pumps in the event the on-site emergency equipment failed. For the plant staff, extra food and water was stored on-site and additional satellite phone were issued to key personnel (NRC, 2011). None of this would have been possible for the Fukushima Daiichi event. Furthermore, because the infrastructure surrounding Fort Calhoun was more or less intact, if Fort Calhoun needed additional resources, it would have also been relatively easy to resupply them. For Fukushima Daiichi, because the magnitude 9 earthquake and tsunami caused extensive damage to the surrounding infrastructure, it was

considerably more difficult to resupply the station when they needed additional resources (e.g., additional diesel fire trucks).

Given these key differences between the Fukushima Daiichi flood induced SBO, Hurricane Katrina and Fort Calhoun, there are no additional insights to be gained with respect to the immediate, post-initiator actions operators needed to take in response to these other flooding events. There are, however, a number of latent organizational challenges related to 1) the degree of disaster preparedness pre-event, and 2) the long-term capability to respond to the flooding event. For example, one parallel between Fukushima and Hurricane Katrina is the degree to which experts underestimated the extent of destruction nature can wrought. Decisions as to what type of flood protection should be built in New Orleans were influenced by cost considerations (Rogers, Kemp, Bosworth, & Seed, 2015). Furthermore, misinterpretation of data resulted in a decrease in the reliability of the floodwalls surrounding New Orleans. Similar arguments could be made about the height of the sea wall protecting Fukushima Daiichi. Additionally, similar to Hurricane Katrina, the long-term emergency response to the tsunami was complicated by several factors. According to the 2012 report by the National Diet of Japan (Japan's bicameral legislature), emergency procedures and SAMGs for an SBO were not well developed due to the perceived low probability of a tsunami of that magnitude occurring, and procedures from other countries (e.g., United States) could have been implemented prior to the earthquake. Having said this, the examination of latent organizational factors on human performance is outside the current scope of the RISMC modeling framework. The near-term focus is on modeling the human performance of MCR and auxiliary operators in the hours and days post-flooding, and the challenges they have to maintain adequate safety margin for the NPP.

3. STATISTICAL MODELING CONSIDERATIONS FOR COMPUTATIONAL HUMAN RELIABILITY ANALYSIS

3.1 Introduction

Human reliability can be greatly impacted by beyond-design-basis accidents like severe flooding or seismic events. As noted in Boring, St. Germain, et al. (2015), many HRA methods do not properly account for human actions during severe accidents. Unique aspects of human activity include the use of severe accident mitigation guidelines (SAMGs), use of FLEX equipment that are a part of the Diverse and Flexible Coping Strategies implemented by NPPs post-Fukushima, unavailability of key equipment, use of extended teams as part of the emergency operating center, and potentially issues of fitness for duty due to severe environmental conditions and prolonged work periods.

The legacy of HRA is that almost all methods to date have been static (Boring, Mandelli, et al., 2015). Just as the adaptation from design-basis to beyond-design basis is difficult for static methods, the problem is made more complex when developing dynamic HRA methods, which look at the emergent evolution of an event instead of analyzing a prescripted set of scenarios. The promise of dynamic methods is their ability to model performance more completely than the expert judgment processes required for completing static HRAs. The downside of dynamic methods is the increased methodological and implementation complexity. The general challenge of making HRA dynamic is increased multifold when dynamic methods must tackle the inherent uncertainty of severe accidents. Not only is the method complexity increased, but also the modeling complexity.

Static methods are based on analyzing human performance for a pre-defined set of tasks that are generally clustered as human failure events (HFEs). The challenge in extrapolating from these HFE snapshots to dynamic models is that many of the basic assumptions of these methods have not been validated for dynamic applications. For example, as depicted hypothetically in Figure 8, a sequence of events can be parsed in many ways. The horizontal axis divides the event along a chronological progression, in this case in terms of minutes. The dotted vertical lines demark subtasks during the sequence of events. Finally, the blue boxes denote HFEs. Each minute reveals a different outcome in terms of the dynamic HEP calculation. Similarly, the subtasks and HFEs track the changing HEP. Yet, HRA methods are not designed to track at all three levels of delineation. An HRA method that is applied successfully to three sequential HFEs as part of an event progression may not adequately cover further delimiting the HFE into 9 subtasks or 10 minute-long time slices. To model the event progression, however, it is necessary to model the HFE at a finer granularity corresponding to the 9 subtasks or 10 time slices. The static HRA method may not lend itself to these different units of analysis. Moreover, the error quantification approach used may not prove accurate for the different unit of analysis.

To frame the event progression in Figure 8 differently, consider the case of a major flooding incident. Major damage to the plant is sustained around the 4-minute mark along the timeline. HFE1 corresponds to the pre-initiator, HFE2 encompasses the initiating event, and HFE3 spans the post-initiator recovery. As can be seen, the HEP remains low during the pre-initiator period, surges during the initiating event, and remains high during the recovery period. Static HRA methods, which would tend to analyze the event in terms of the three HFEs may not fully model the changes to operator performance within each HFE. For example, the surge in error during HFE2 (likely caused by sudden increases in stress) actually consists of three different slopes of the error plot—an initial relatively flat period, a rapidly rising period, and a plateau that shows signs of gradually declining. The flooding has differing effects on the plant and the operators, but conventional static parsing of the event may not fully map the dynamic progression of the event and the equally dynamic error curve associated with different tasks and time slices.

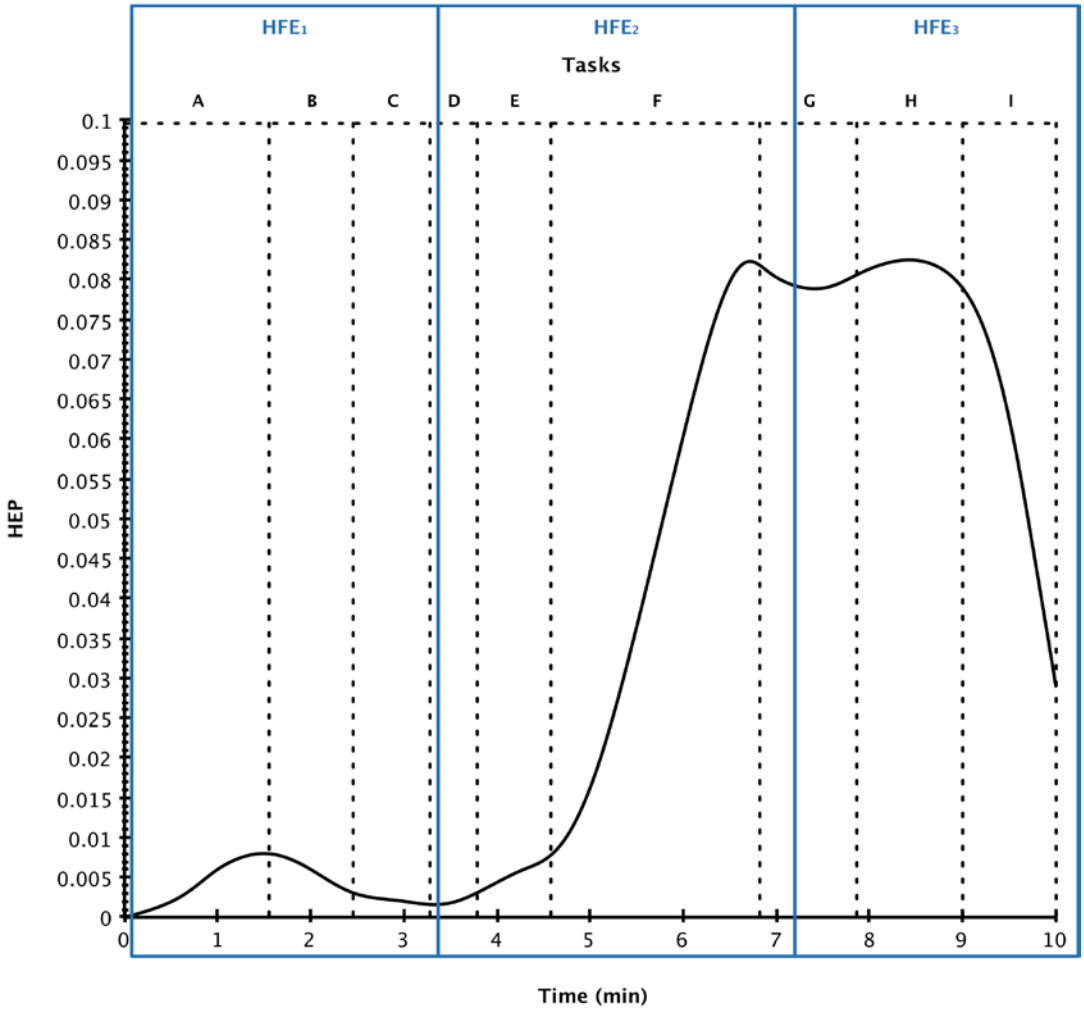

Figure 8. Human event progression according to time slices, subtasks, and HFEs.

This chapter reviews what happens to HRA when the unit of analysis is changed from an HFE to a unit of analysis suitable for dynamic modeling. Underlying this discussion is the key assumption that dynamic HRA requires a finer grain of modeling precision than the HFE. Ideally, the HFE represents a thorough human factors subtask analysis (EPRI, 1992; IEEE, 1997; Kolaczkowski et al., 2005). The human reliability analyst will then quantify the event at the appropriate level of aggregation. HRA methods treat the unit of quantification differently. For example, the original HRA method, the Technique for Human Error Prediction (THERP; Swain and Guttman, 1983) quantifies at the subtask level. In contrast, SPAR-H (Gertman et al., 2005) analyzes events at the HFE level, despite being derived from THERP (Boring & Blackman, 2007). Ideally, the quantification approach should transfer between different framings of the event space.

This chapter reviews three areas of HRA quantification as they are translated from HFEs to subtasks or time slices. These areas are:

- Uncertainty quantification (see Section 3.2),
- Conditional Probability Quantification (see Section 3.3), and
- Basic Probability Quantification (see Section 3.4).

Swain and Guttman (1983) discuss three phases of quantification in THERP, which are largely mirrored in subsequent HRA methods:

- *Nominal HEP*—which is the default error rate corresponding to a particular task type,
- *Basic HEP*—which is the nominal HEP modified for influences like PSFs on the basic performance, and
- *Conditional HEP*—which is the basic HEP modified for dependence, the assumption that the error likelihood increases when an error condition is already underway.

These three areas map also well to the three sections covered in this chapter, although in a slightly different order. It is also important to note that this chapter will not review assumptions about the nominal HEP. Rather, it is assumed that the nominal error rate tables provided in a method such as THERP are valid at the subtask level for which they were originally created. However, associated with each HEP is also a measure of uncertainty. The uncertainty discussion centers on statistical considerations associated with propagating uncertainty over a large number of units of analysis.

This chapter does not focus exclusively on flooding events. Flooding, as discussed earlier, is only one of a number of different kinds of external events, but the findings should generalize to other types of events, from normal occurrences to off-normal and severe accidents. The important point to note is that when modeling dynamic HRA, the granularity of analysis is finer than most static HRA methods. This chapter serves to test how conventional HRA methods scale to this level of precision, whether for flooding or other events. In order to determine whether HRA methods can be use to model flooding effects dynamically, they first need to model general phenomena dynamically. As such, this chapter serves as a modeling proof for the transferability of static HRA quantification to dynamic applications.

3.2 Uncertainty Quantification

3.2.1 Basic Equations

To calculate the total human failure in a NPP and the corresponding uncertainty bounds, the following set of equations were outlined in Appendix A of THERP (Swain and Guttman, 1983). The probability of total failure, $P(F_t)$, is the sum of all of the individual failure probabilities, $P(F_i)$. These individual failures, F_i, correspond to tasks as defined in THERP. Additionally, each F_i is composed of several subtasks, $F_{(i,j)}$, which do not individually lead to a failure event but only result in failure when a certain combination of $F_{(i,j)}$ occur together (see Figure 9 for illustration). The probability distribution of $P(F_{(i,j)})$ is assumed to be log-normal, and the relationship between $P(F_i)$ and $P(F_{(i,j)})$ is described in THERP Appendix A in the following way:

$$P(F_i) = \prod_{j}^{n_i} P(F_{(i,j)}) \tag{5}$$

where i is a task, j is a subtask of i, and n_i is the number of subtasks in each task. For example, as seen in Figure 9, there are there tasks, or i's, such that we model $P(F_1)$, $P(F_2)$, $P(F_3)$. Also in Figure 9, there are three subtasks in $P(F_1)$, or j's such that we have $P(F_{(1,1)})$, $P(F_{(1,2)})$, $P(F_{(1,3)})$. The probability of a basic or conditional HEP is $P(F_{(i,j)})$ and has a known or expertly defined median (M_{ij}), upper (U_{ij}), and lower (L_{ij}) bound. From these three values of log-normal $F_{(i,j)}$ the following system of equations are derived to describe the behavior of the uncertainty bounds on F_t. Thus, for a log-normal distribution of $P(F_{(i,j)})$ the following is defined:

$$\mu_{ij}ln = \ln(M_{ij}) \tag{6}$$

$$\sigma_{ij}ln = \frac{1}{3.29} ln\left(\frac{U_{ij}}{L_{ij}}\right) \tag{7}$$

$$U_{ij} = e^{\mu_{ij}ln + 1.645 * \sigma_{ij}ln} \tag{8}$$

$$L_{ij} = e^{\mu_{ij}ln - 1.645 * \sigma_{ij}ln} \tag{9}$$

where $\mu_{ij}ln$ is the log-normal mean of $P(F_{(i,j)})$, not to be confused with μ_{ij} which is the mean calculated assuming a normal distribution. $\sigma_{ij}ln$ is the log-normal standard deviation of $P(F_{(i,j)})$ and 1.645 is the standard normal ($Z_{0.95}$) for a 95% confidence interval. Equation (7), describing log-normal standard deviation ($\sigma_{ij}ln$) was included so the origin of Equation (11), for $\sigma_i ln$, can be derived easily. Since we know that $P(F_{(i,j)})$ is log-normal, we are also aware therefore that $P(F_i)$ is additionally a log-normal distribution with mean ($\mu_i ln$) and standard deviation ($\sigma_i ln$). As such, we see the following behaviors of log-normal mean ($\mu_i ln$), and standard deviation ($\sigma_i ln$):

$$\mu_i ln = \sum_{j}^{n_i} \ln(M_{ij}) \tag{10}$$

$$\sigma_i ln = \frac{1}{3.29} \sqrt{\sum_{j}^{n_i} \left(ln\left(\frac{U_{ij}}{L_{ij}}\right)\right)^2} \tag{11}$$

Additionally the log-normal variance ($\sigma_i{}^2 ln$) is considered in the following equation:

$$\sigma_i{}^2 ln = \frac{1}{3.29^2} \sum_{j}^{n_i} \left(ln\left(\frac{U_{ij}}{L_{ij}}\right)\right)^2 \tag{12}$$

From the values of the mean and standard deviation for each individual failure event, the following equations define the upper (U_i) and lower (L_i) uncertainty bounds on $P(F_i)$:

$$U_i = e^{\mu_i ln + 1.645 * \sigma_i ln} \tag{13}$$

$$L_i = e^{\mu_i ln - 1.645 * \sigma_i ln} \tag{14}$$

To calculate the mean (μ_i) and variance ($\sigma_i{}^2$) on $P(F_i)$, without a log-normal bias, and for further use in the system of equations:

$$\mu_i = exp\left[\mu_i ln + \frac{\sigma_i{}^2 ln}{2}\right] \tag{15}$$

$$\sigma_i{}^2 = exp[\sigma_i{}^2 ln + 2 * \mu_i ln](exp[\sigma_i{}^2 ln] - 1) \tag{16}$$

Note that based on Equation A12 in THERP in appendix A, there is a superscripted 2 before the $\mu_i ln$. It is assumed here that this should be a multiple of 2 rather than a square on the preceding lognormal. And, finally, we have the mean and variance without a log-normal bias of $P(F_t)$:

$$\mu_T = \sum_i^n \mu_i \tag{17}$$

$$\sigma_T{}^2 = \sum_i^n \sigma_i{}^2 \tag{18}$$

where μ_T is the mean and $\sigma_T{}^2$ is the variance of $P(F_t)$ assuming a normal distribution. Additionally, the normal variance of $P(F_t)$ is used for further calculations rather than the standard deviation previously used, as seen in Equations (19) and (20).

$$\mu_T ln = ln \frac{\mu_T}{\sqrt{1 + \frac{\sigma_T{}^2}{\mu_T{}^2}}} \tag{19}$$

$$\sigma_T{}^2 ln = ln \frac{\sigma_T{}^2}{\mu_T{}^2} \tag{20}$$

where $\mu_T ln$ is the log-normal mean and $\sigma_T{}^2 ln$ is the log-normal variance of of $P(F_t)$. Finally, $P(F_t)$ is calculated using Equation (21). Furthermore, median (M_T), upper bound (U_T), and lower bound (L_T) of the probability of total failure ($P(F_t)$) are provided in the following:

$$P(F_t) = \sum_i^n P(F_i) \tag{21}$$

$$M_T = e^{\mu_T ln} \tag{22}$$

$$L_T = e^{\mu_T ln - 1.645 * \sigma_T{}^2 ln} \tag{23}$$

$$U_T = e^{\mu_T ln + 1.645 * \sigma_T{}^2 ln} \tag{24}$$

3.2.2 Simulation of Uncertainty Bounds

A simulation was created using the statistical software package R (R Core Team, 2015) that personified the behavior described in THERP Appendix A, which was captured in the previous section's system of equations. This simulation takes the structure of Figure 9 with three failure paths, whereby each path has multiple $F_{(i,j)}$, where i is defined as the number of failure paths, such that for Figure 9, i=3, and $P(F_i)$ is the probability of an HFE, which sums together to equal the probability of total failure, $P(F_T)$ as seen in Equation (21).

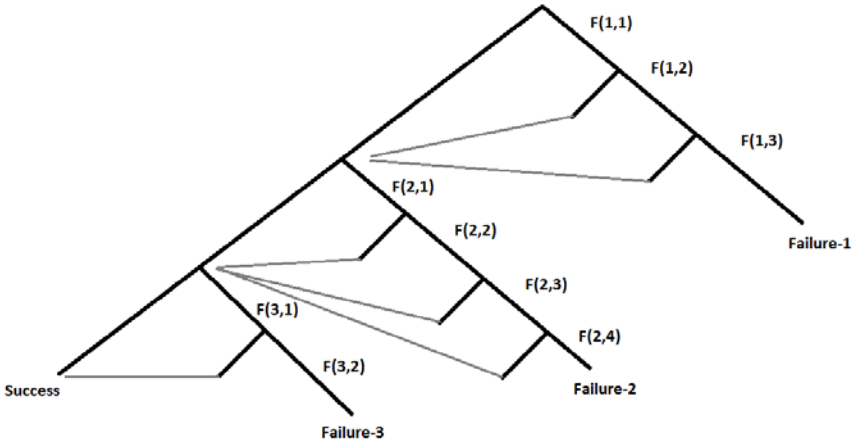

Figure 9. THERP HRA event tree with 3 failure paths.

For the purposes of simulation, one value of $P(F_{(i,j)})$ was randomly selected from a generation of a random log-normal distribution of 100 observations with a log-normal mean of -5.8 and log-normal standard deviation of 0.3, which is a log-normal probability distribution centered around 0.003 (i.e., 3E-3), a common nominal HEP used in THERP. This was done so that the calculation could be empirically derived rather than based on expert opinion. The median (M_{ij}) and standard deviation (σ_{ij}) were calculated for each ($F_{(i,j)}$) using Equations (25) and (26). This was done, since an expertly defined upper bound and lower bound of $P(F_{(i,j)})$ for a subtask were not available.

$$\sigma_{ij}ln = \sqrt{\log\left(\frac{1 + variance}{mean^2}\right)} \tag{25}$$

$$\mu_{ij}ln = \ln\left(\frac{mean}{\sqrt{1 + \frac{variance}{mean^2}}}\right) \tag{26}$$

From the example in Figure 9, there are three HFEs: 3 $P(F_1)$ [i.e. $P(F_{(1,1)})$, $P(F_{(1,2)})$, $P(F_{(1,3)})$], 4 $P(F_2)$, and 2 $P(F_3)$. Thus, there are 900 $P(F_{(i,j)})$ generated in groups of 100, with each group of 100 having the lognormal standard deviation (σ_{ij}ln) calculated from either Equation (7) or (25). The log-normal mean of $P(F_{(i,j)})$ can be calculated using Equation (6) or (26) and used to calculate the upper bound (U_{ij}), and lower bound (L_{ij}). A $P(F_{(i,j)})$ is randomly selected from each of the groups of 100 to represent a single $P(F_{(1,1)})$, $P(F_{(1,2)})$, … $P(F_{(3,2)})$. The results of these calculations are summarized in Table 1.

Table 1. Calculations for all events in one iteration of the THERP HRA event tree. Equations (6)–(11) were applied to the calculations pertaining to $P(F_{(i,j)})$, resulting in the necessary values for $P(F_i)$ as displayed in Table 2.

$F_{(i,j)}$	i	j	M_{ij}	$\sigma_{ij}ln$	U_{ij}	L_{ij}	$P(F_{(i,j)})$
$F_{(1,1)}$	1	1	0.00318	0.30307	0.00537	0.00198	0.00282
$F_{(1,2)}$	1	2	0.00312	0.29078	0.00509	0.00196	0.00386
$F_{(1,3)}$	1	3	0.00300	0.30307	0.00489	0.00180	0.00241
$F_{(2,1)}$	2	1	0.00278	0.35420	0.00498	0.00155	0.00324
$F_{(2,2)}$	2	2	0.00300	0.30340	0.00496	0.00183	0.00335
$F_{(2,3)}$	2	3	0.00310	0.26600	0.00482	0.00201	0.00270
$F_{(2,4)}$	2	4	0.00304	0.32809	0.00516	0.00175	0.00325
$F_{(3,1)}$	3	1	0.00300	0.25446	0.00451	0.00195	0.00306
$F_{(3,2)}$	3	2	0.00283	0.29332	0.00468	0.00178	0.00140

Table 2. Computations for each failure in the THERP HRA event tree, $P(F_i)$, using Equations (10)–(16). These values are calculated off the values in Table 1 for $P(F_{(i,j)})$.

F_i	$\mu_i ln$	$\sigma_i ln$	$P(F_i)$	U_i	L_i	μ_i	σ_i^2
F_1	-17.3304	0.5179	2.62E-08	6.97E-08	1.27E-08	3.40E-08	3.56E-16
F_2	-23.2651	0.6292	9.51E-11	2.22E-10	2.80E-11	9.60E-11	4.47E-21
F_3	-11.6769	0.3883	4.29E-06	1.61E-05	4.48E-06	9.15E-06	1.36E-11

Equations (17) through (22) were applied to the calculated values of $P(F_i)$, as seen in Table 2, resulting in an upper bound (U_T) and lower bound (L_T) estimate of $P(F_t)$. This is then simulated 5,000 times so that there exist 5,000 calculations of total failure probability ($P(F_t)$), upper bound (U_T), lower bound (L_T), and median (M_T). This is done so that their behavior can be clearly viewed with sufficient clarity. Additionally, the upper bound (U_T), lower bound (L_T), median (M_T) and total failure probability ($P(F_t)$) distribution are graphed in Figure 10. The golden colored portion of the image is the distribution as would be seen in a histogram, mirrored, while the black portion of each component follows the pattern of a box and whisker plot. The inner thicker black bar is the interquartile range, which contains 50% of the data. The whiskers are the outer quartiles of the data, and the white dot is the median.

The distribution of $P(F_{(i,j)})$ and $P(F_t)$ is based on Figure 10, and the equations modeled are log-normal. The lower bound and median have a normal behavior, while the upper bound retains the log-normal behavior. After consideration, starting with HEP of 0.003 for subtasks $P(F_{(i,j)})$ and having a total failure probability $P(F_t)$ around 1e-5 does not make sense. The model appears to stabilize when larger distributions of lognormal are applied to $P(F_{(i,j)})$; however, these is a limit on enlarging $P(F_{(i,j)})$ as the range of $P(F_t)$ will go beyond 100% when $P(F_{(i,j)})$ approaches but does not exceed 1.

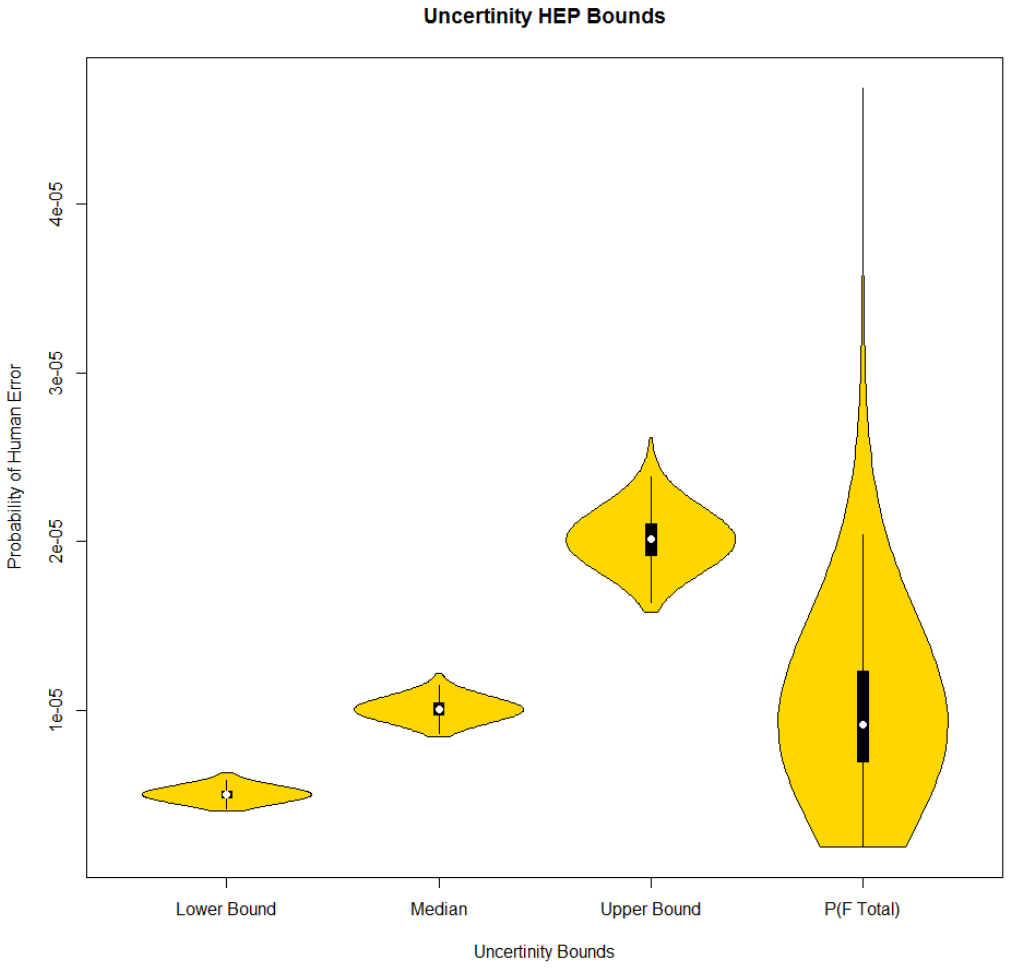

Figure 10. A violin plot of the lower bound (L_T) of $P(F_t)$, median (M_T) of $P(F_t)$, upper bound (U_T) of $P(F_t)$.

3.3 Conditional Probability Quantification

Dependence, as commonly treated in HRA, is the relationship between two or more sequential human errors. Almost all dependence modeling is based on the approach first presented in THERP (Swain & Guttman, 1983). THERP breaks dependence down into five conditional equations, corresponding to five levels of dependence. The modeling exercise presented in this section is purely exploratory with an approach toward a continuous distribution desired that can be simulated in a dynamic HRA. For the purposes of this section, we assume Task (or Event) A precedes Task B, and both involve human actions. Given Task A occurs, and then Task B and a discrete dependency level, we have the following equations crafted in THERP:

$$P(B|A|ZD) = P(B) \tag{27}$$

$$P(B|A|LD) = \frac{1 + (19 * P(B))}{20} \tag{28}$$

$$P(B|A|MD) = \frac{1 + (6 * P(B))}{7} \tag{29}$$

$$P(B|A|HD) = \frac{1 + P(B)}{2} \tag{30}$$

$$P(B|A|CD) = 1 \tag{31}$$

where *ZD* is zero dependence, *LD* is low dependence, *MD* is moderate dependence, *HD* is high dependence, and *CD* is complete dependence, as selected by the analyst. These equations produce the conditional probability of human error on Task B given Task A and the dependence level. When an HRA is completed, a dependence level for Task B is assigned by an expert, and Task A is not taken directly into consideration in the calculation of the conditional probability. Task A is a prerequisite but not a calculated contributor to the conditional probability of Task B. In Figure 11 each of the dependence equations are graphed assuming Event B has a random uniform distribution from 0 to 1.

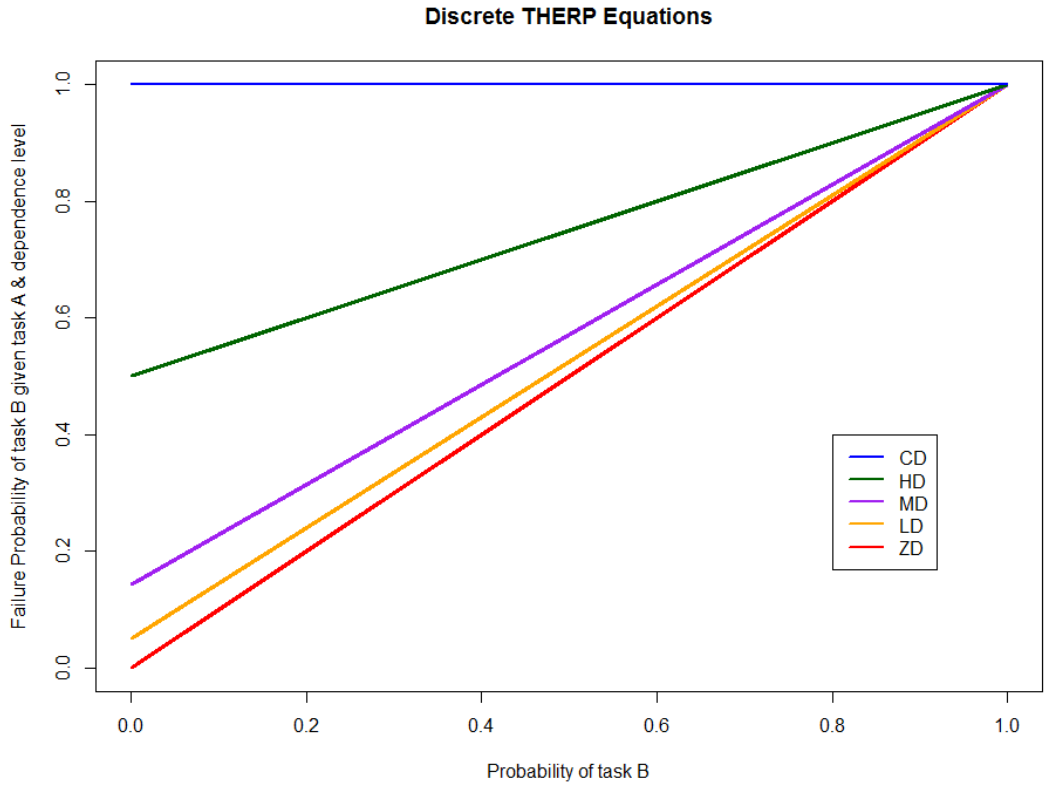

Figure 11. Failure probability of Task B given dependence levels and Task A.

As seen in Figure 11, at low probabilities, high dependence is midway between complete dependence and zero dependence. The widest difference between dependence levels occurs at low probabilities; with convergence occurring at the probability of 1 for Task B, or 100%. Dependence effectively serves to set a lower bound on the HEP: $^1/_{20}$ (i.e., 0.05) for low dependence, $^1/_7$ (i.e., 0.14) for moderate dependence, and $^1/_2$ (i.e., 0.5) for high dependence. This is commonly referred to as a human performance limiting value.

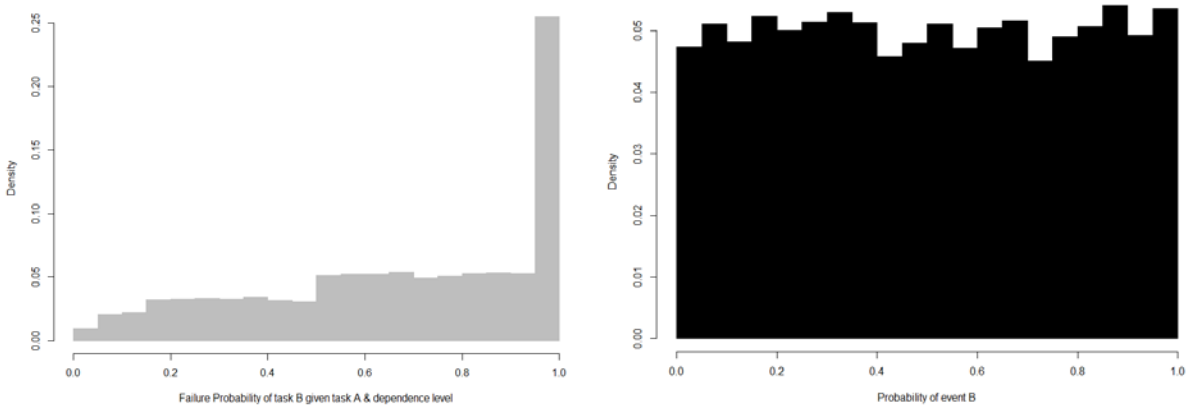

Figure 12. Distribution of HEP of Task B given all dependence levels, Equations (27)-(31), given Task B, is a uniform distribution (left) and HEP as a random uniform distribution of Task B (right).

The probability distribution function of Task B is randomly uniform from 0 to 1 as seen in Figure 12right. The distribution of the HEP for Task B given it has a uniform distribution and all dependence levels, is not what we currently see in reality. Figure 12 left shows that a majority of events are operating at a complete dependence level because the other dependence levels are distributed over several bins while CD inhabits its own bin. Because of this a higher, than what is considered acceptable, rate of $P(B|A|\text{Dependence}) = 1$, which we know to be false. Additionally we know that the distribution of human failure events in a NPP has a log-normal distribution, and the log-normal curve is located at very low levels, as failure incidents do not frequently occur. As such, for further simulations, human failure of Task B is given a random lognormal distribution centered on 0.003, as indicated by the red vertical line in Figure 13 right.

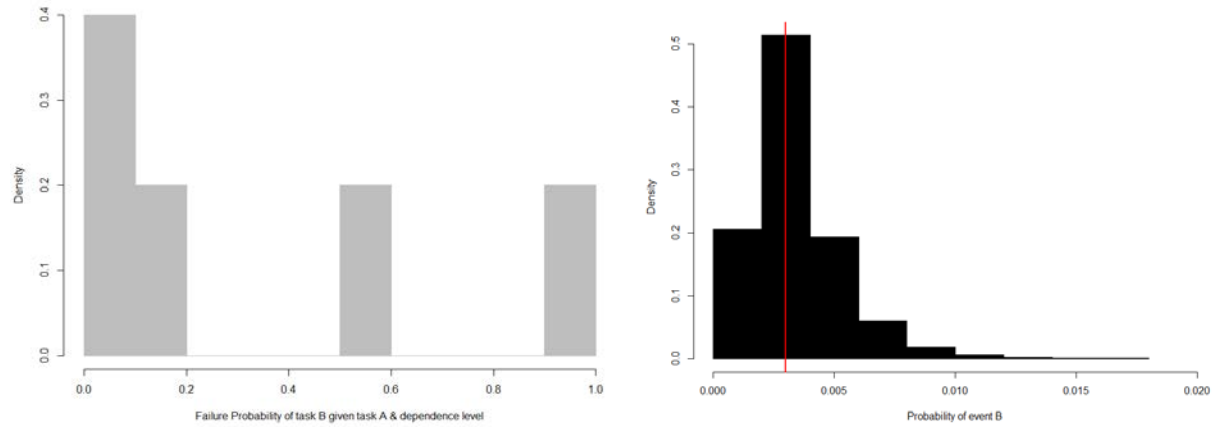

Figure 13. Distribution of HEP of Task B given all dependence levels, Equations (27)-(31), given task B, is a log-normal distribution (left) and random log-normal distribution of Task B centered around 0.003 as indicated by the red line (right).

The HEP for Task B is now behaving closer to what we know as reality in Figure 13 left, with a majority of the data below 0.2. The resulting distribution from all discrete dependence distributions of the human error from Task B is pictured in Figure 13 left. The discrete nature of this distribution is now apparent, where the bar between 0.9-1 is indicative of complete dependence and the bar 0.5-0.6 is clearly the high dependence being expressed.

Swain and Guttman (1983) described human error dependence as a continuum, not something to be discretized; unfortunately, discrete behavior is clearly exemplified in Figure 13 left. Swain and Guttman originally described the frequency of zero dependence as "uncommon between human tasks," (p. 10-15) and complete dependence as "unusual" (p. 10-18). Further descriptions state, "Complete dependence between performances on two tasks is rare, but it not as rare as zero dependence" (Bell & Swain, 1980). Low dependence is assigned, "if there is any doubt as to an assessment of ZD." This leaves the speculation that low, moderate, and high dependencies occur more frequently.

The small range as well as low probability of human error for Task B in Figure 13 right is closer to the reality of a working NPP. In Figure 13 left, the results of each dependence level can easily be distinguished in the distribution of conditional THERP equations. The suggested form the THERP equations imply is to move away from the original discrete approach, with the addition of a continuous variable C to the equation:

$$P(B|A|Dependence) = \frac{(1 + ((C - 1) * P(B)))}{C} \tag{32}$$

where *P(B)* is the probability of human error of Task B, and the value of *C* is expertly assigned on a range of 1 to 20. The value of *C*=20 behaves like low dependence, and a value of *C*=1 behaves like complete dependence in this scenario. For our purposes, we know that *CD* and *ZD* occur infrequently. It is also postulated that low, moderate, and high dependency occurs more commonly. Thus the distribution of *C* is assumed to be normal centered on 7 (moderate dependence) with truncation at 1 and 20. The continuous form of the dependence calculation does not consider zero dependence, which is essentially a case where *C* is infinitely large. For zero dependence, no correction equation is applied such that:

$$P(B|A|ZD) = P(B|A) \tag{33}$$

The performance of continuous dependence is illustrated in a series of graphs below. Figure 14 shows a depiction of dependence level, *C*, as a uniform distribution. Next, Figure 15 also shows the behavior of Equation (32), except the dependence levels, *C*, have a normal distribution. Finally, Figure 16 shows that *C* is a log-normal distribution centered on 7 (moderate dependence) and B is log-normally distributed and centered on an HEP of 0.003.

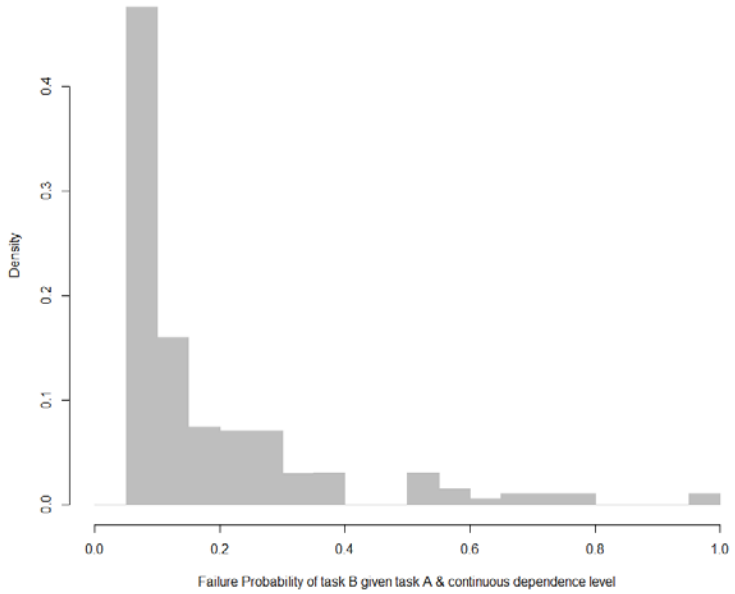

Figure 14. The distribution of the conditional THERP coefficient from Equation (32) with a continuous uniform distribution for dependence level (C_{min}=1, C_{max}=20), and log-normal distribution of Task B.

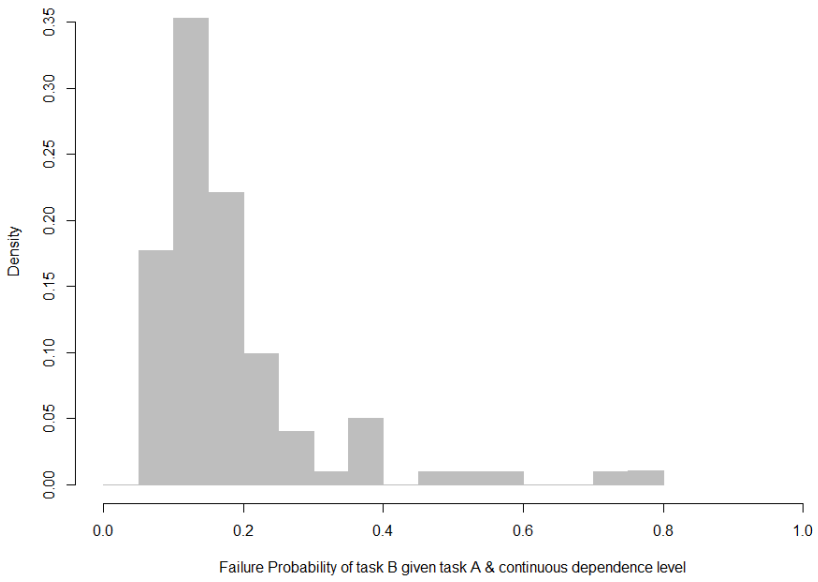

Figure 15. Distribution of the conditional THERP coefficient from Equation (32) with a continuous normal dependence level (C) and log-normal distribution of Task B.

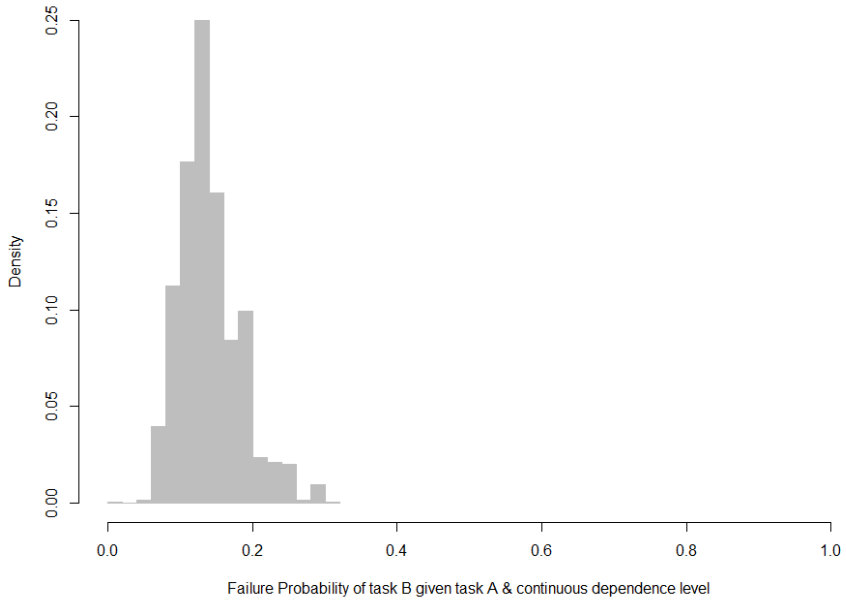

Figure 16. Distribution of the conditional THERP coefficient from Equation (32) with a continuous lognormal dependence level (C) and log-normal distribution of Task B.

3.3.1 Joint Distribution

Joint distribution is the behavior that is described when Task (or Event) A and Task B occur at the same time. This is usually difficult to characterize when A and B are dependent upon one another. The event of zero dependence, or independent events, only occur when the performance of Task B is unaffected by the performance of Task A, which is infrequent. Bayes rule postulates the probability of A given that we know the probability of B:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \tag{34}$$

Another form of Bayes' equation incorporates the joint distribution of A and B such that:

$$P(A|B) * P(B) = P(A,B) = P(B|A) * P(A) \tag{35}$$

If we were to apply this well-known law to the THERP equations, such as Čepin (2007) has previously done, we see the following relationship for the joint probability of human error dependence between Tasks A and B:

$$P(A,B)_{ZD} = P(B) * P(A) \tag{36}$$

$$P(A,B)_{LD} = P(A) * \frac{1 + (19 * P(B))}{20} \tag{37}$$

$$P(A,B)_{MD} = P(A) * \frac{1 + (6 * P(B))}{7} \tag{38}$$

$$P(A,B)_{HD} = P(A) * \frac{1 + P(B)}{2} \tag{39}$$

$$P(A,B)_{CD} = P(A) \tag{40}$$

Assuming that probability of human error on Tasks A and B retains a random log-normal behavior centered on an error rate of 0.003, the behavior of the discrete dependence levels can be seen below in Figure 17. Figure 17 shows that the discrete dependence levels clearly clump into levels as intuitively expected. Zero dependence inhabits the lowest joint probability, and complete dependence takes the higher values of joint probability.

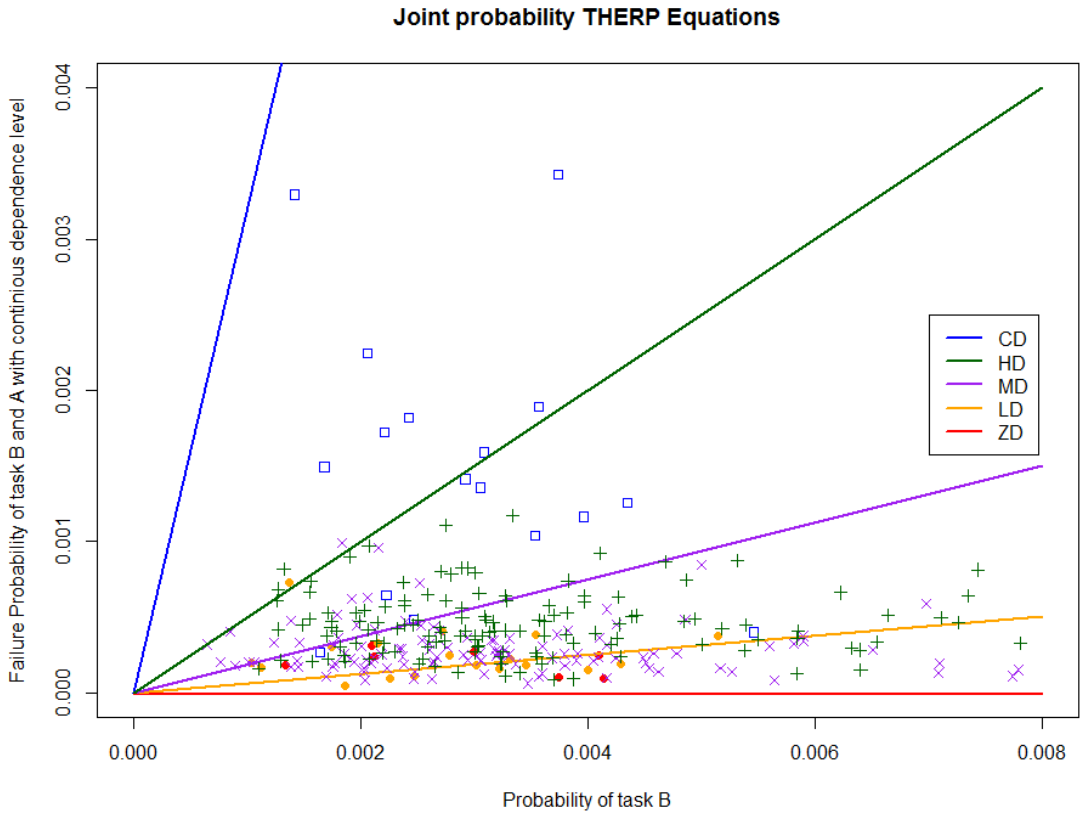

Figure 17. Joint dependence calculations after Čepin (2007).

The next step was to allow dependence to move fluidly in a continuous nature. This was completed using Equation (41). Yet again when C is 1 the behavior is complete dependence, and when C is 20 the dependence is defined as low.

$$P(A,B)_C = P(A) * \frac{1 + ((C-1) * P(B))}{C} \tag{41}$$

An exploratory visualization of model Equation (41), in different ranges of *C*, other than 0-20, can be seen in Figure 18.

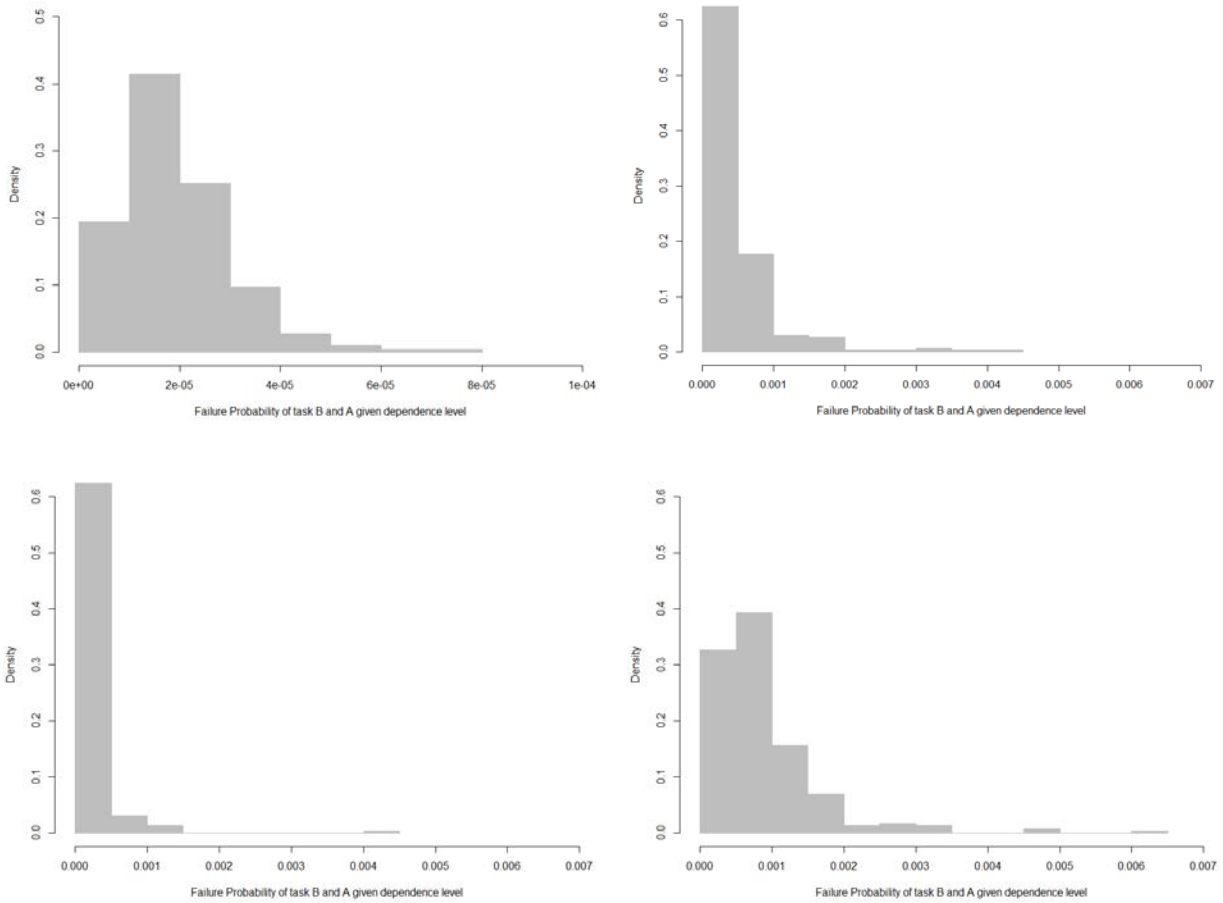

Figure 18. Log-normal human error distribution of Tasks A and B centered on an HEP of 0.003, with a normal distribution of *C* dependence truncated at 1-10 (top left), 1-20 (top right), 1-100 (bottom left) and 1-1000 (bottom right).

Figure 18 shows the behavior of dependence as the values of C are increased using Equation (41). While the distribution shape remains distinctly log-normal with a long tail, the range it inhabits decreases significantly as each order of C increases. Again, the smaller the value of C, the closer to complete dependence; and the larger C becomes, the closer the dependence tends toward zero. In terms of dynamic dependence, this finding suggests that a continuous dependency level needs to be used so that more advanced and accurate simulations can model dependent HRA. Additional theoretical considerations of dependence are discussed in Boring (2015), reiterating the importance of the dynamics of dependence beyond traditional static modeling.

3.4 Basic Probability Quantification

3.4.1 Introduction to SPAR-H

SPAR-H is a widely accepted method to determine the HEP based on expert estimation using calculation worksheets. Estimations are carried out using weighted PSFs and a standard diagnosis failure probability. In many HRA methods, including SPAR-H, context-specific probabilities are generated by multiplying a nominal HEP by multipliers representing the effect of specific context elements (generally

represented by PSFs) which were deemed relevant to the problem by the method developers. This has resulted in the following equation:

$$DFP = BR * PSF \tag{42}$$

where *DFP* is the diagnosis failure probability, *BR* is the base rate which is assumed to be 1E-3 (after the Action worksheet in SPAR-H), and *PSF* is the product of all eight PSFs in the method (Gertman et al., 2005). PSFs come in many flavors, with SPAR-H defining: available time, stress, complexity, experience, procedures, ergonomics, fitness for duty, and work process. Each PSF has different levels with a corresponding multiplier for diagnosis and action as seen in Table 3.

PSFs	PSF Level	Multiplier for Action	Multiplier for Diagnosis
Available Time	Inadequate Time	P(failure)=1	P(failure)=1
	Time Available ≈ Time Required	10	10
	Nominal Time	1	1
	Extra Time Available	0.1	0.1
	Expansive Time Available	0.01	0.1 to 0.01
	Insufficient Information	1	1

Table 3. The PSF available time with its respective levels and the associated action and diagnosis multipliers.

The application of the PSF levels and multipliers produce the following equation:

$$DFP = BR * \text{available time} * \text{stress} * \text{complexity} * \text{experience} * \text{procedures} * \text{ergonomics} * \text{fitness for duty} * \text{work process} \tag{43}$$

where each PSF is substituted with the respective PSF level's multiplier. Of course, each level of a PSF is not equally likely. As such, the frequency of PSF level assignments was taken from Boring et al. (2006). Additionally, for the purposes of this exploratory analysis, only the SPAR-H Action worksheet PSF multipliers are used. A small excerpt of the data used for this simulation can be seen in Table 4.

PSFs	PSF Level	Multiplier for Action	Action Frequency	Action Probability
Available Time	Inadequate Time	P(failure)=1	5	0.009
	Time Available ≈ Time Required	10	36	0.065
	Nominal Time	1	500	0.898
	Extra Time Available	0.1	10	0.018
	Expansive Time Available	0.01	4	0.007
	Insufficient Information	1	2	0.004

Table 4. Shown is the PSF 'available time' with its respective levels, action multiplier, action frequency, and action probability.

3.4.2 Human Failure Event Simulation

The simulation of human failure event simulation are based on the probabilities of a PSF level in Table 4 and Equation (43), a simulation of 5,000 data points are run to represent the distribution of a single task. This is then repeated for Tasks A, B and C, so that there are a total of 15,000 data points.

Without taking into consideration the frequencies provided by Boring et al., (2006) and assuming that each PSF level is equally likely, the distributions of Tasks A, B, and C tend toward the probability of 100%, as seen in Figure 12 right. Verifying the results from the simulation, a one-way analysis of variance (ANOVA) could be used to compare means of three or more groups. However, the distributions of the Tasks and HFE are clearly not normally distributed, thus a non-parametric approach, Kruskal-Wallis, is suggested for comparison purposes. Tasks A, B and C were compared using a Kruskal-Wallis analysis and received p-value of 0.6186 with 2 degrees of freedom. This is what is expected as one task is generated from the same data as the other and does not differ much from another (see Figure 19).

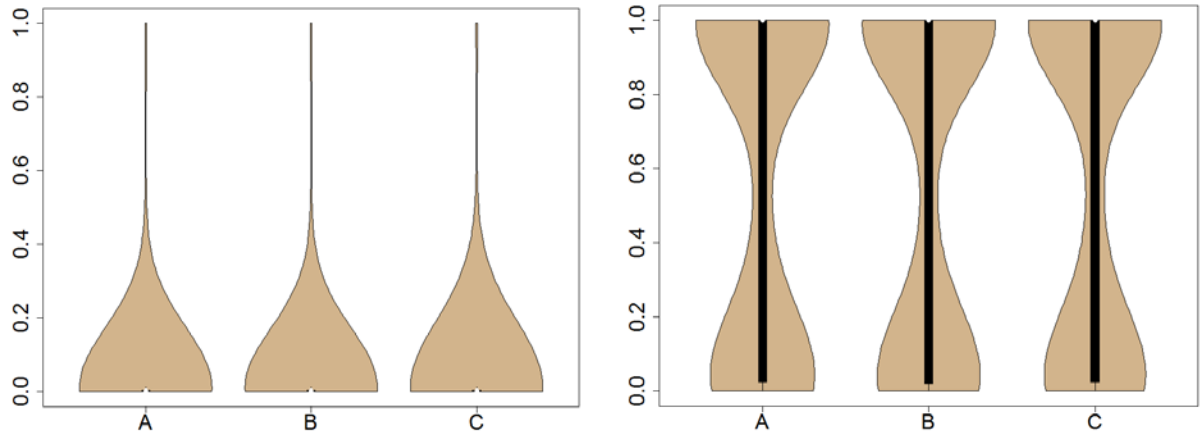

Figure 19. Tasks A, B and C taking into consideration PSF frequencies from Boring et al. (2006) (left). Tasks A, B and C assuming each PSF level are equally likely (right). Because A, B, and C are generated in the same manner, for 5,000 iterations A, B, and C are expected to have the same distributions.

Multiple tasks are often grouped as HFEs. THERP provides an explicit way to map the tasks and subtasks to HFEs; SPAR-H assumes the unit of analysis is the HFE. If HFE1 is comprised of Tasks A, B, and C (see Figure 8), there are then several ways to calculate the HFE based on a PSF multiplier or group of PSF multipliers. The Maximum HFE calculation selects the largest values across Tasks A, B and C. The assumption is that the analysis should capture the strongest manifestation of the PSF, even if the PSF changes across the evolution of the HFE. Median HFE selects the median value of the three tasks, and Average HFE calculates the average of the three tasks. The respective distributions for the different HFEs can be seen in Figure 20.

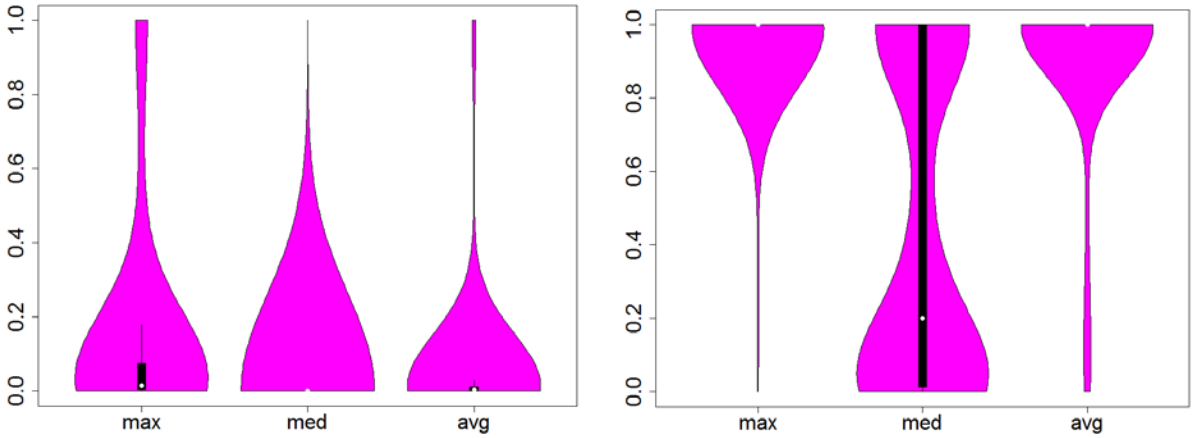

Figure 20. Violin plot of HFEs calculated three different ways from Task A, B, and C. The Maximum (max) calculation selects the largest of the three tasks. Median (med) selects the median value of the three tasks. Average (avg) calculates the average of the three tasks. The left is calculated using frequencies from Boring et al. (2006), while the right is calculated assuming a uniform frequency for all PSF levels.

Again, Tasks A, B, and C and Maximum HFE were compared using a Kruskal-Wallis analysis and received p-value < 0.001 with 3 degrees of freedom. Tasks A, B, and C and Average HFE were compared using a Kruskal-Wallis analysis and received p-value < 0.001 with 3 degrees of freedom. Both of these p-values indicate that Maximum HFE and Average HFE are significantly different from Tasks A, B, and C (Figure 21). Additionally, Tasks A, B, and C and Median HFE were compared using a Kruskal-Wallis analysis and received p-value < 0.001 with 3 degrees of freedom. While still significant, visually Median HFE is the closest in distribution to the three tasks, and the graphical representation can be seen in Figure 21. Generally, Maximum HFE overestimate Task A, B, and C and average underestimates Task A, B, and C. Again all 6 distributions, Task A, B, C, Max HFE, Median HFE and Average HFE can be seen in Figure 21.

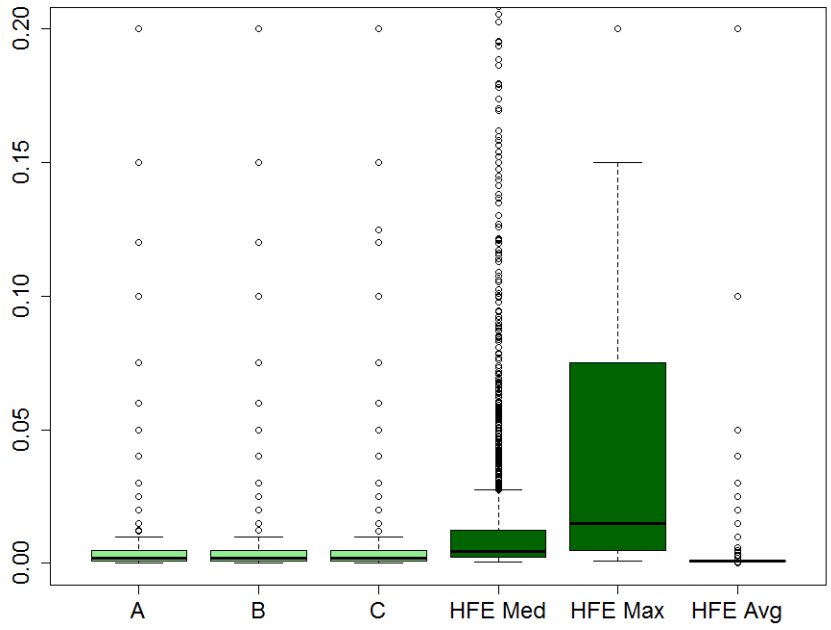

Figure 21. Task A, B, C, HFE Median, HFE Maximum, and HFE Average. Each task was sampled 5,000 times from each PSF with frequencies.

3.4.3 Joint THERP Dependency Simulation

Once again, Tasks A, B, and C were generated using SPAR-H with 5,000 observations per task, for a total of 15,000 observations generated. This exploration applied the THERP joint distribution equations for ZD, MD, and CD. This applied Equations (44), (45), and (46) respectively.

$$P(A,B,C)_{ZD} = P(Task\ A) * P(Task\ B) * P(Task\ C) \tag{44}$$

$$P(A,B,C)_{MD} = P(Task\ A) * \frac{1 + (6 * P(Task\ B))}{7} * \frac{1 + (6 * P(Task\ C))}{7} \tag{45}$$

$$P(A,B,C)_{CD} = P(Task\ A) \tag{46}$$

A Kruskal-Wallis test was again employed, as none of the distributions are a normal distribution. ZD and MD are significant from Task A, B and C with a p-value <0.001 and 3 degrees of freedom each. CD is not significant with a p-value of 0.936, which is to be expected, as complete dependence is the value of the first task, which in our case is Task A. Distributions of the three dependency levels can be seen in Figure 22.

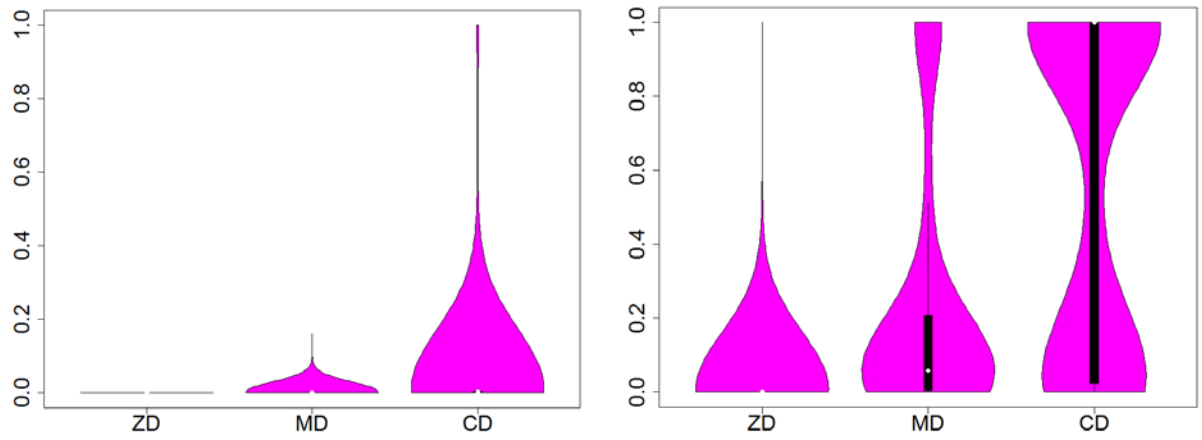

Figure 22. A violin plot of Zero Dependence (ZD), Moderate Dependence (MD), and Complete Dependence (CD) calculated for joint THERP dependence and frequencies from Boring et al. (2006) applied to the PSF levels for Tasks A, B, and C (left). Joint THERP dependence equations applied assuming PSF level is equally likely for Tasks A, B, and C.

As seen in figure 15 and the results from the Kruskal-Wallis test, zero dependence is very low such that its estimate is unrealistic. Medium dependence appears to be reasonable, and within the same range of the tasks. And complete dependence is the same distribution as Task A, and as such is exactly the same distribution as the Tasks.

3.4.4 Further Simulations

SPAR-H in a dynamic simulation can be seen in Figure 8. Viewing the situation from a more simplified position, how to define a Task or HFE into subtasks becomes burdensome if calculations change depending on the resolution defined, and whether the data is point estimates or units of time. The summation of Tasks is usually limited in publications, as indefinite summation will certainly lead to infinity. Contrarily, if all tasks have their product taken of one another, while they would never reach a calculated 0, they would asymptotically approach 0. Thus, reality exists in the space between the summation of all tasks and their product. A simulation was run varying the rate of calculation, given that each task is sampled from a 0.003 centered log-normal distribution. The results of which can be viewed in Figure 23.

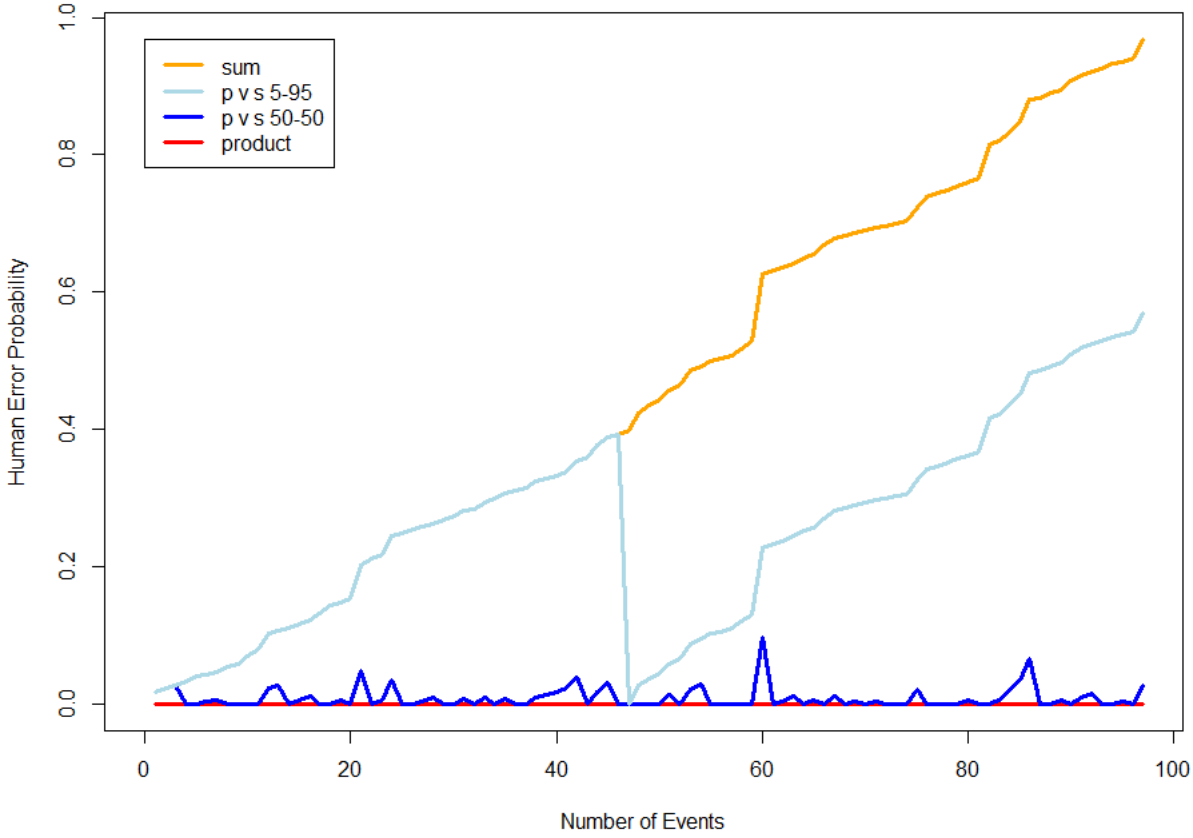

Figure 23. HEP for number of events. Each event was taken from a random log-normal distribution centered on 0.003.

Each event is randomly sampled from a log-normal distribution centered on 0.003. Defining these events as a task or sub-task is irrelevant as sub-tasks are multiplied together to calculate a task, and tasks are summed together to receive the total failure probability. In Figure 23 we can see that the sum of HEP of all events (orange) will eventually hit 100% at roughly 100 events. We know that in a nuclear power plant setting humans try to avoid making mistakes, so a 100% HEP scenario is very unlikely. The product of all of the same tasks (red) is also unlikely as each human does make mistakes, at some point. Then we have the two blue lines – the dark blue is a 50% chance of an "and" or product calculation and a 50% of an "or" or sum calculation with a distribution as seen in Figure 24 left. While the light blue line is a 5% chance of an "and" or product calculation and a 95% of an "or" or sum calculation with a distribution as seen in Figure 24 right. Simulation was repeated until the summation (Figure 16 orange) had a human error probability of 100% or 1.

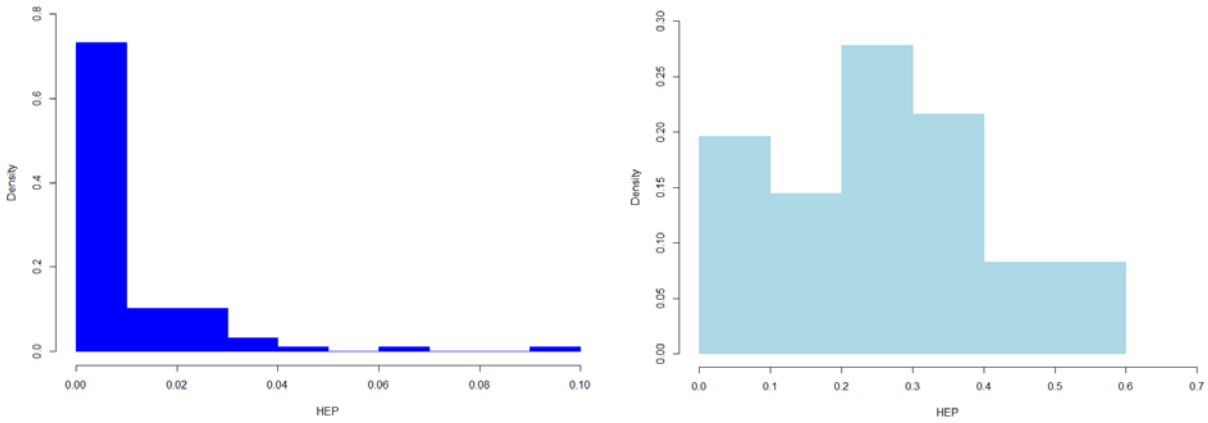

Figure 24. The distribution of a 50% chance of an "and" or product calculation and a 50% of an "or" or sum calculation of events (left). The distribution of a 5% chance of an "and" or product calculation and a 95% of an "or" or sum calculation of events (right).

This chapter has reviewed basic assumptions of quantification and extrapolated from static modeling to dynamic modeling. HRA in support of dynamic or computation-based HRA must necessarily model a range of human actions, typically at a finer resolution than is accounted for by the HFE. As the HFE is parsed into finer units of analysis, it is important to consider the mathematical underpinnings of the HRA methods that will be used for quantification. Two HRA methods, THERP and SPAR-H, were considered as part of this chapter. THERP, which was originally based on subtask analysis, generally translates from static to dynamic HRA. SPAR-H, which is based on HFEs, may require further refinement before its quantification approach can be employed in dynamic HRA. This may not be entirely unexpected, since SPAR-H is by design a simplified method. While the use of SPAR-H gives the method flexibility to model a wide variety of scenarios, including flooding events, the use of the HFE as a unit of analysis is problematic when modeling the dynamic evolution of the event. This lack of fine resolution can result in spurious HEPs when using SPAR-H. Conversely, the subtasks modeled in THERP may not generalize to novel scenarios such as flooding, making the quantification impossible based on the subtask lookup tables in the method. Clearly, a hybrid approach should be employed in dynamic HRA. Future efforts within HUNTER will seek to refine these static methods for better application in dynamic contexts. Additional candidate HRA methods will also be explored.

4. CONCLUSION

The purpose of this report has been to summarize INL's latest efforts to develop a computational HRA method for inclusion into the RISMC framework. Recognizing that existing static HRA methods, and even dynamic HRA methods, cannot be easily incorporated into the RISMC multi-model simulation based approach, INL has and continues to develop a computational HRA method called HUNTER. This R&D effort is exploring how HUNTER, using simulation and modeling to create a virtual human operator, can qualitatively and quantitatively describe how human performance affects and is affected by NPP behavior as well as external events, such as flooding.

HRA methodological issues that were specifically investigated include:

- How HRA is affected by changing the unit of analysis from an HFE to something more suitable for dynamic modeling
- What happens to various aspects of quantification within HRA when it is treated dynamically, including uncertainty, conditional probability, and basic probability quantification
- Why is dependence among actions important, and how might it be treated quantitatively

This exploration of statistical modeling for computational HRA also investigated what happens to the calculations of dependency when the unit of analysis is further subdivided from tasks into sub-tasks. In addition, recognizing that severe accidents, in particular external flooding events, are of particular interest to model and understand more completely from a risk perspective, INL has and will continue to focus on applying HUNTER to a range of severe accidents, including SBOs, seismic events, and SBOs induced by external flooding.

4.1 Next Steps

Boring et al. (2015) identified a number of next steps, and many of them are still applicable components of our research path forward. Namely:

4.1.1 Continue to develop the HUNTER framework

To further develop HUNTER, more HRA elements need to be incorporated into its structure. For example, the crew activity submodels, PSF models for crew activities, and NPP control actions (as described in Boring et al., 2015), need to be further developed, as well as the use of dynamic Bayesian networks. Once these tasks are completed, the HUNTER framework will be applied to various demonstration tests (e.g., a scenario involving decision making by the operators while under the influence of PSFs), so that it can be evaluated and further refined. This work will also continue to integrate HUNTER with RAVEN, following the approach as described in Section 1.1.1. of this report.

4.1.2 Conduct a Proof-of-Concept Demonstration of HUNTER

One early, proof-of-concept activity that we have also started and will continue to investigate is using the Human Systems Simulation Laboratory (HSSL) at INL to test the emergent interplay between plant conditions and operator actions. Specifically, researchers have and will continue to explore the use of generic Pressurized Water Reactor (gPWR) simulator in the HSSL as a platform to study the interaction between virtual models of human operators, multi-physics codes, and a specific plant model. The gPWR is a full scope plant model based on RELAP-5, and because it was originally designed as a

training simulator, it has the ability to script a full range of scenarios, including insertion of operator actions and omissions. The researchers thus plan to use the gPWR to run various accident scenarios while varying key aspects of the operator's actions and responses to plant conditions (e.g., delays in making correct diagnoses, delays in decision-making and actions, etc.). It is important to note, however, that the gPWR is a real time auto-run of scenarios, and does not have a faster than real-time simulation capability using massively parallel supercomputing. As such, this initial proof-of-concept testing is a stepping-stone to eventual integration with RELAP-7 and other aspects of the RISMC framework.

Despite this limitation, researchers have begun to investigate the feasibility of using the gPWR to model and simulate how safety systems and operators would respond to a flooding incident. To simulate conditions similar to Fukushima Daiichi, we identified safety systems that would be vulnerable to flooding that would also have a significant impact on the safety of the plant needed. We are also in the process of investigating how human actions would be necessary to maintain cooling (i.e., remove decay heat) post-trip affect plant behavior. In particular, the first safety system in the gPWR the team is investigating for this proof-of-concept testing is the residual heat removal (RHR) system. Based on conversations with a gPWR subject matter expert (i.e., gPWR simulator instructor), having the RHR fail will lead to challenges in removing decay heat, but that the effect will also take some time. In a flooding event, the RHR pumps are outside of containment, and could be susceptible to failure when inundated with water. The gPWR subject matter expert reported that there could also be leaks in the emergency service water that affect the service water system and cause the RHR pumps to fail. The expected outcome of investigating this scenario would be to see how the resulting decay heat temperature changes as a function of the operator being able to perform their actions successfully versus failing at their actions.

4.1.3 Long Term Research Needs

In addition to these near term activities, there is a need to:

- Take the findings from these early demonstrations of the HUNTER method to evaluate and further refine it.
- Perform addition proof-of-concept demonstrations, possibly including aspects of FLEX, to further evaluate the robustness of HUNTER across multiple scenarios of interest.
- Validation of HUNTER, particularly the aspects of HEP quantification, through various means, including:
    - Performing a Bayesian update on legacy performance data and comparing the results from the HUNTER proof-of-concept demonstrations to those data.
    - Comparing the results of the HUNTER proof-of-concept demonstrations to data from actual crews running scenarios in the HSSL.

These HRA R&D activities will allow the RISMC framework to encompass a greater range of plant dynamics during upsets, enhance how the framework dynamically models responses to changing conditions, and will reduce epistemic uncertainty in modeling, thereby creating a technical basis for including models of human performance into the RISMC framework.

5. REFERENCES

Alfonsi, A., Rabiti, C., Mandelli, D., Cogliati, J. & Kinoshita, R. (2013). RAVEN as a tool for Dynamic Probabilistic Risk Assessment: Software Overview. International Conference on Mathematics and Computational Methods Applied to Nuclear Science & Engineering (M&C 2013). Sun Valley, ID.

Boring, R.L. (2015). A dynamic approach to modeling dependence between human failure events. In L. Podofillini et al. (Eds.), Safety and Reliability of Complex Engineered Systems, 2845-2851. London, UK: Taylor & Francis Group.

Boring, R.L., & Blackman, H.S. (2007). The origins of the SPAR-H method's performance shaping factor multipliers. Official Proceedings of the Joint 8th IEEE Conference on Human Factors and Power Plants and the 13th Annual Workshop on Human Performance/Root Cause/Trending/Operating Experience/Self Assessment, 177-184.

Boring, R., Mandelli, D., Joe, J., Smith, C., & Groth, K. (2015). A Research Roadmap for Computation-Based Human Reliability Analysis, INL/EXT-15-36051. Idaho Falls, ID: Idaho National Laboratory.

Boring, R., Shirley, R., Joe, J.C., Mandelli, D., & Smith, C. (2014). Simulation and Non-Simulation Based Human Reliability Analysis Approaches, INL/EXT-14-33903, Idaho Falls, ID: Idaho National Laboratory.

Boring, R.L., Whaley, A.M., Tran, T.Q., McCabe, P.H., Blackwood, L.G., & Buell, R.F. (2006). Guidance on Performance Shaping Factor Assignments in SPAR-H, INL/EXT-06-11959. Idaho Falls, ID: Idaho National Laboratory.

Čepin, M. (2007). DEPEND-HRA – A method for consideration of dependency in human reliability analysis. Reliability Engineering & System Safety, 93, 1452-1460.

David, R., Gaston, D., Martineau, R., Peterson, J., Zhang, H., Zhao, H. & Zou, L. (2012). RELAP-7 Level 2 Milestone Report: Demonstration of a Steady State Single Phase PWR Simulation with RELAP-7. Idaho Falls, ID: Idaho National Laboratory.

Eide, S., Gentillon, C., Wierman, T., & Rasmuson, D. (2005). Reevaluation of station blackout risk at nuclear power plants: Analysis of Loss of Offsite Power Events: 1986-2004, NUREG/CR-6890, Vol. 1. Washington, DC: U. S. Nuclear Regulatory Commission.

Electric Power Research Institute (EPRI). (2012). A preliminary approach to human reliability analysis for external events with a focus on seismic, 1025294, Palo Alto, CA.

Electric Power Research Institute (EPRI). (1992). SHARP1 – A Revised Systematic Human Action Reliability Procedure, EPRI-101711, Palo Alto, CA.

Gaston, D., Hansen, G. & Newman, C. (2009). MOOSE: A Parallel Computational Framework for Couples Systems for Nonlinear Equations. International Conference on Mathematics, Computational Methods, and Reactor Physics, Saratoga Springs, NY.

Gertman, D., Blackman, H., Marble, J., Byers, J., & Smith, C. (2005, August). The SPAR-H Human Reliability Analysis Method, NUREG/CR-6883. Washington, DC: U.S. Nuclear Regulatory Commission.

International Atomic Energy Agency (IAEA). (2011). IAEA international fact finding expert mission of the Fukushima Dai-Ichi NPP accident following the great east Japan earthquake and tsunami. Tokyo, Japan.

Institute of Electrical and Electronics Engineers (IEEE). (1997). Guide for Incorporating Human Action Reliability Analysis for Nuclear Power Generating Stations, IEEE 1082. New York, NY.

Institute of Nuclear Power Operations (INPO). (2011). Special report on the nuclear accident at the Fukushima Dai-ichi nuclear power station, INPO 11-005. Atlanta, GA.

Kadota, R. (2014). On the brink: The inside story of Fukushima Daiichi. Chuo-Ku, Fukuoka, Japan: Kurodahan Press.

Kolaczkowski, A., Forester, J., Lois, E., and Cooper, S. (2005). Good Practices for Implementing Human Reliability Analysis (HRA), Final Report, NUREG-1792. Washington, DC: U.S. Nuclear Regulatory Commission.

Mandelli, D., Smith, C., Ma, Z., Riley, T., Schroeder, J., Rabiti, C., Alfonsi, A., Nielsen J., Maljovec, D., Wang, B. & Pascucci, V. (2013). Risk-Informed Safety Margin Characterization (RISMC) BWR Station Blackout Demonstration Case Study, INL/EXT-13-30203. Idaho Falls, ID: Idaho National Laboratory.

Park, J., Kim, Y., Kim, J. H., Jung, W., & Jang, S. C. (2015). Estimating the response times of human operators working in the main control room of nuclear power plants based on the context of a seismic event: A case study. Annals of Nuclear Energy, 85, 36-46.

Presley, M., Julius, J. Grobbelaar, J., & Kohlheepp, K. (2013a). A preliminary approach to human reliability analysis for external events with a focus on seismic HRA, Proceedings of the ANS PSA 2013 International Topical Meeting on Probabilistic Safety Assessment and Analysis, Columbia, SC.

Presley, M., Julius, J. Grobbelaar, J., & Kohlheepp, K. (2013b). A review of seismic operating experience with implications for human reliability, Proceedings of the ANS PSA 2013 International Topical Meeting on Probabilistic Safety Assessment and Analysis, Columbia, SC.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rabiti, C., Alfonsi, A., Cogliati, J., Mandelli, D. & Kinoshita, R. (2013). Deployment and Overview of RAVEN Capabilities for a Probabilistic Risk Assessment Demo for a PWR Station Blackout, INL/EXT-13-29510. Idaho Falls, ID: Idaho National Laboratory.

Rabiti, C., Alfonsi, A., Cogliati, J., Mandelli, D. & Kinoshita, R. (2012). Reactor analysis and virtual control environment (RAVEN) FY12 report, INL/EXT-12-27351, Idaho Falls, ID: Idaho National Laboratory.

Rabiti, C., Mandelli, D., Alfonsi, A., Cogliati, J. & Kinoshita, R. (2013). Mathematical framework for the analysis of dynamic stochastic systems with the RAVEN code. In Proceeding of M&C2013 International Topical Meeting on Mathematics and Computation. LaGrange Park, IL: American Nuclear Society.

Rogers, J. D., Kemp, G. P., Bosworth, H. J., Jr., & Seed, R. B. (2015). Interaction between the US Army Corps of Engineers and the Orleans Levee Board preceding the drainage canal wall failures and catastrophic flooding of New Orleans in 2005. Water Policy, 17, 707-723.

Smith, C., Mandelli, D., Prescott, S., Alfonsi, A., Rabiti, C., Cogliati, J. & Kinoshita, R. (2014). Analysis of Pressurized Water Reactor Station Blackout Caused by External Flooding Using the RISMC Toolkit, INL/EXT-14-32906. Idaho Falls, ID: Idaho National Laboratory.

Swain, A.D., & Guttman, H.E. (1983). Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications. Final report. NUREG/CR-1278. Washington, DC: U.S. Nuclear Regulatory Commission.

The National Diet of Japan. (2012). The official report of the Fukushima nuclear accident independent investigation commission. Tokyo, Japan.

U.S. NRC (2011). The rising river puts flood preparation to the test. http://public-blog.nrc-gateway.gov/2011/06/22/the-rising-river-puts-flood-preparations-to-the-test/. URL retrieved 30 September 2015.